Deep Residual Pyramid-Based Skin Lesion Segmentation and Feature Extraction Mechanism*

DONG Chunxu[1], LI Xue[2], CHEN Siguang[1*]
(1.School of Internet of Things, Nanjing University of Posts and Telecommunications, Nanjing Jiangsu 210003;
2.Department of Dermatology, Women's Hospital of Nanjing Medical University (Nanjing Woman and Children's healthcare Hospital), Nanjing Jiangsu 210004)

Abstract: The lesion appearance of skin diseases is characterized by small inter-class variance, large intra-class variance and imbalanced sample distribution, which results in a high misdiagnosis rate of intelligent diagnosis. A deep residual pyramid-based skin lesion segmentation and feature extraction mechanism is proposed to address the above problem. Specifically, compared with the mainstream single-scale network output, a deep residual pyramid multi-scale coding network is constructed to increase the accuracy of feature extraction and training outcomes, and the bottleneck layer is divided into multi-scale coding networks to segment and extract multi-scale features. Meanwhile, a loss gradient supervision mechanism based on focal loss is designed to solve the problem of sample imbalance. That is, the focal loss is used to increase the model's attention to the hard-to-separate samples, and the gradient coordination mechanism is developed to reduce the impact of hard-to-separate samples and outliers, so as to reduce the impact of sample imbalance on the diagnosis results. Finally, the experiment results demonstrate that the proposed mechanism of segmentation and extraction obtains competitive accuracy compared with existing representative works and the Jaccard index is 3%-10% higher, reaching 82.3%.

Key words: skin lesion; deep residual network; feature pyramid; loss function
EEACC: 6200 **doi**: 10.3969/j.issn.1004-1699.2025.02.009

基于深度残差金字塔的皮肤病变分割与特征提取研究*

董春序[1], 李 雪[2], 陈思光[1*]
(1.南京邮电大学物联网学院, 江苏 南京 210003;
2.南京医科大学附属妇产医院(南京市妇幼保健院)皮肤科, 江苏 南京 210004)

摘 要: 由于皮肤病的类内差异大、类间差异小及样本分布不均衡导致恶性皮肤病智能诊断误诊率极高, 因此提出一种基于深度残差金字塔的皮肤病变分割与特征提取机制。具体地, 对比主流的单一尺度网络输出, 为了提高特征提取和训练结果的准确度, 构建了一个深度残差金字塔多尺度编码网络, 将瓶颈层划分为多尺度编码网络, 通过提取多尺度特征, 实现网络分割与提取结果的输出。进一步, 为了解决样本不均衡问题, 设计了基于焦点损失的损失梯度监督机制, 即通过焦点损失增加模型对难分样本的关注度, 同时通过梯度协调机制减小难分样本和离群点对模型整体准确率的影响, 从而达到减小类别不平衡对诊断结果的影响。实验结果表明, 所提机制的分割与提取与现有相关方案相比, Jaccard 系数提高了 3%~10%, 达到 82.3%。

关键词: 皮肤病变; 深度残差网络; 特征金字塔; 损失函数
中图分类号: TP391 **文献标识码**: A **文章编号**: 1004-1699(2025)02-0263-09

皮肤病是普遍存在的疾病之一, 分布年龄广。恶性皮肤病变, 如黑色素瘤, 由于恶化速度较快, 致死率极高, 因此是世界恶性程度较大的癌症之一。研究表明, 如果黑色素瘤能在早期被筛选发现, 该类肿瘤可以被完全治愈。因此, 高效及时筛查恶性皮肤病如黑色素瘤具有重要的医疗价值。由于皮肤病存在种类繁多、病变外观的类间相似度高、类内差异化大等特点, 使得黑色素瘤的识别分类非常困难。近年来, 黑色素瘤的确诊和死亡病例在持续增长, 但是通过早期筛查诊断可以极大降低该病的致死率, 因此寻找一种高效可靠、易于使用且准确灵敏的方法来诊断黑色素瘤, 对世界各地的病人进行早期诊

项目来源: 国家自然科学基金项目(61971235); 中国博士后科学基金(面上一等资助)项目(2018M630590); 江苏省"333 高层次人才培养工程"项目; 江苏省博士后科研资助计划项目(2021K501C); 南京邮电大学'1311'人才计划和赛尔网络下一代互联网技术创新项目(NGII20190702)
收稿日期: 2022-01-07 **修改日期**: 2022-06-24

断具有重要意义。

在过去的十年中,神经网络展现出计算高效、灵敏度高和结果准确等特点,因此近年来被广泛应用于分类和回归任务,尤其是在机器视觉方面的医学临床实践中。研究人员通过导入大量的皮肤镜图像数据集和优化的估计参数,对黑色素瘤的特征进行提取训练,其结果提高了皮肤病分类筛查尤其是黑色素瘤分类的特异性、准确性和灵敏度。与皮肤科医生诊断相比,神经网络的误诊率更低,即与专业皮肤科医生的视觉筛查分类相比,计算机视觉的可靠性更高。神经网络可以作为一种出色的诊断工具,具有更高的灵敏度和特异性,可以辅助病人进行早期筛查,减少了非必要的就诊需求;同时对筛查出的有病变风险的病人给予及时准确的医疗建议,辅助病人尽快治疗,使得恶性皮肤病的治疗更加及时准确。

当然,识别皮肤病,尤其是黑色素瘤,主要分为基于人工设计筛查的方法、基于传统机器学习算法的方法和基于卷积神经网络的方法。

①基于人工设计筛查的方法

文献[1-2]将传统的阈值设定、ABCDE 法等分割方法应用于检测黑色素瘤。此类方法虽然增加了黑色素瘤的提取维度,结果优于人工检测,但是对于复杂外观的特征处理能力低,无法准确地分割病变区域。

为解决上述问题,文献[3]阐述了一种基于手动设计的特征分类方法。作者将形状、颜色等手工设计的特征与已优化的特征进行融合以区分病变区域;文献[4]利用皮肤图像中提取的形状等特征设计了一系列特征选择算法来提取特征,并训练支持向量机(Support Vector Machines,SVM)分类器进行分类;文献[5]提出一种新的分类器训练方法。作者用一组特征信息(颜色、纹理等)描述每个候选病变区域,最终训练一个 AdaBoost 分类器;Situ 等[6]从皮肤镜图像提取 16^2 的提取颜色、尺寸等特征,并利用特征包模型(Bag of Feature,BoF)将所得特征聚合为最终表述。上述方法一定程度上提高了系统处理复杂外观的特征提取能力,但是所用特征单一,且着重于全局特征,忽略了局部特征。局部特征有利于利用细节区分黑色素瘤病变区域与非病变区域,对于相似区域的划分作用重大,因此该类方法的特征提取准确性较低。

②基于传统机器学习算法的方法

针对处理复杂外观问题,文献[7]研究了一种毛发检测分割方法。此方法基于病变的特定特性共分为两部分,一部分是利用二维 Gabor 滤波器进行毛发检测;一部分是利用基于色度的集合模型确定病变区域,其中色度特性用于协助控制停止评估。该方法利用色度从背景中提取细节特征,能够有效处理源于毛发的复杂外观,但对于其他因素(如血管)复杂外观处理能力一般,且不能有效区分黑色素瘤与其他相似的皮肤病,因此毛发检测的方法准确率较低。

Khan 等[8]提出一种基于概率分布分割的病变检测与分类方法。该文利用受控熵进行特征选择,融合均匀分布与正态分布两种分割方法确定分割区域,最终利用多类支持向量机对病变进行分类;在文献[9]中,Filali 等提出一种基于对比的病变位置分割方法,先以像素为单位划分区域,再根据相似性计算区域间的连通性,同时得到对比度。为提高对比度的精度,作者对特征进行加权处理,且用一种层次推理技术来控制尺度变化。以上方法基于邻域分量分析改进了特征提取的方法,利用最佳纹路特征进行病变分割,有效提高了病变分割与分类准确性,但是存在计算成本昂贵以及由于样本类别不平衡导致的提取能力差、准确率极低与误诊率极高的问题。

③基于卷积神经网络的方法

近年来,卷积神经网络[10-11](Convolutional Neural Network,CNN)被广泛应用于医疗图像领域,并取得了突破。Codella 等[12]将 CNN、稀疏编码、支持向量机融合用于色素性皮肤病的识别;Kawahara 等[13]利用基于 AlexNet 的全卷积神经网络进行黑色素瘤的特征提取。以上应用 CNN 的高辨别力提升了色素性皮肤病识别系统的性能,但是其网络存在局限性,即特征易丢失、特征提取能力差、准确率低,难以得到较优的结果。

针对灵敏度与准确性依赖样本的问题,一些学者提出了集成学习的方法。例如,文献[14]将 CNN 网络应用于来自包含国际皮肤成像协会(International Skin Imaging Collaboration,ISIC)的多个数据库、18 位临床医生的大量样本,平均准确率为 72.1±0.9%;文献[15]用 CNN 网络集成识别 ISIC2018 数据库的 6 类数据,得到的最高准确率为 76%;文献[16]也提出一种新的 CNN 网络集合,作者结合整个图像信息与数据库中分割后被裁剪过的图像信息,经过 900 次训练与 379 次测试得到两类,最终得到曲线下面积(Area Under Curve,AUC)为 0.843,精度为 64.9%。以上集成方法结合不同分类单元的特性,所得结果优于其中每个单独分类器的平均值。然而,这类集成模型需要大量资源,训练开销极大。

针对网络提取能力差与准确率低的问题,一些

学者采用增加网络深度的方法。文献[17]提出一种 5 层 CNN 架构来区分病变区域;文献[18]提出利用一种单一的深度网络进行黑色素瘤的分类;结合卷积神经网络与反卷积神经网络进行语义分割,同时应用了全卷积网络,在 ISIC2015 样本集上得到了 92.03%的交并比(Intersection over Union,IOU),但是准确率不足 70%;文献[19]对包含光照与噪声效应的临床图片进行预处理,利用 CNN 进行分类,应用于包含 70 张黑色素瘤与 100 张黑色素痣图像的数据库中,得到了 81%的准确率。虽然该类方法通过增加网络深度提升了模型的性能,但是随着网络深度的增加,网络的准确度出现饱和甚至下降的现象,网络开始退化,同时提取的特征极易丢失。

基于上述方法存在的问题,Yu 等[20]提出一种基于深度残差网络的多阶段方案,用于黑色素瘤的自动识别,取得了 78.3%的准确率;Codella 等[16]设计了一种集成的深度学习方法,融合深度残差网络、U-Net、Caffe-Net 进行特征提取,取得了 70.2%的准确率;Hagerty 等[21]利用残差网络(Residual Network,ResNet)中 ResNet-50 架构从图像中提取深度特征,在单一数据集上取得了 91.3%的准确率,但多数据集的平均准确率仅有 76.4%。以上方法融合图像中提取的特征信息,通过连接特征向量与分类器来收集数据,有效减轻了网络退化问题,黑色素瘤识别取得了较优的结果。本文基于深度残差网络构建网络模型,采用深度残差网络可以有效解决上述深度网络存在的问题,如网络退化,特征提取能力差。

在上述研究的基础上,本文将深度残差网络与特征金字塔结合,提出一种基于深度残差金字塔的病变分割与特征提取机制。该基于深度残差金字塔的病变分割与特征提取机制将单一尺度拓展为多尺度,即瓶颈层划分为多尺度编码网络,通过提取多尺度特征,输出分割与提取结果。通过改进该结构,可提升神经网络的分割能力和特征提取能力,在提高网络深度和预测性能的同时可降低网络的训练开销,解决了主流 CNN 的退化问题。同时为了解决样本不平衡问题,在焦点损失[22]和梯度协调机制[23]的基础上提出了基于焦点损失的损失梯度监督机制,在解决样本不平衡问题的同时平衡了离群点、易分样本和难分样本的权重,使模型适度关注难分样本,可解决焦点损失过度关注难分样本导致整体准确率降低的问题。实验结果表明,本文所提的基于深度残差金字塔的病变分割与特征提取机制降低了样本不平衡问题对模型的影响,提高了整体分割与提取的准确率,高于主流方案,达到 82.3%。

本文其余部分组织如下:第一节介绍了本文的系统模型;第二节给出了本文所提的深度残差金字塔的病变分割与特征提取机制的具体介绍;第三节是仿真结果与分析;最后第四节总结全文。

1　网络模型

在本节中,本文构建了一个基于深度残差金字塔的皮肤病智能提取模型,该模型由四个部分组成:用户、提取云、模型训练与数据提供。本文模型的四个组成部分如图 1 所示,其中,用户通过采集设备收集病变图片,将数据上传到提取云并从提取云下载提取结果,获取提取服务;提取云通过模型训练获取用于提取的训练模型,并为模型训练提供提取结果数据下载;模型训练通过样本源提供的训练样本,构建残差金字塔多尺度编码网络训练模型;样本源提供用于训练的脱敏数据,同时从模型训练服务器下载提取云上传的用户采集的数据,用于核验分析模型训练的提取结果,调整模型训练的参数并标注样本以扩充样本集。

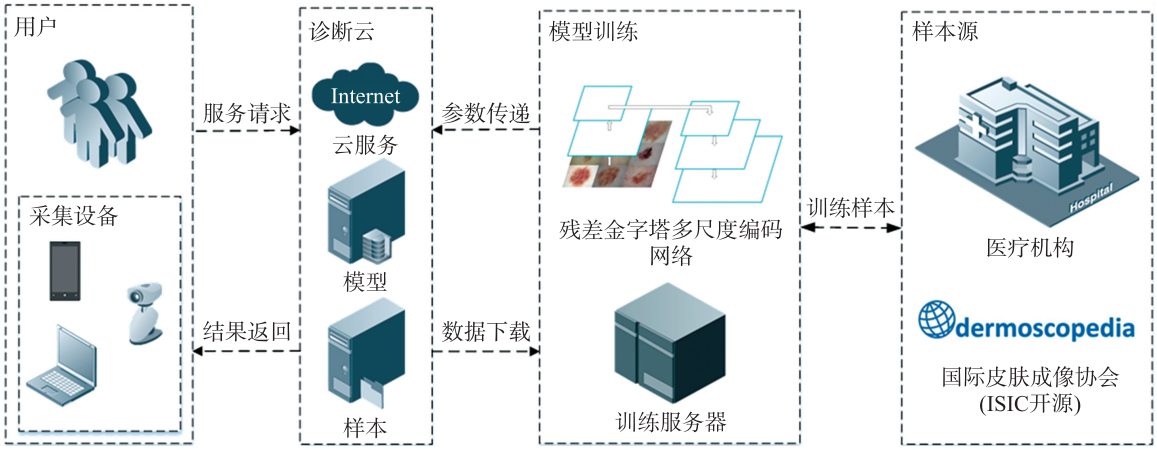


图 1　皮肤病智能提取系统模型

用户:本文考虑的用户为皮肤病患者与皮肤科医生。本文为不同患者提供不同水平的服务,主要分为业余采集用户与专业皮肤科采集用户。业余用户主要通过移动设备(如手机)拍摄病变图片后获得初诊结果,受移动设备拍摄的光照、角度与清晰度等因素影响,服务提取结果的精确度会因此受到影响,作为初步筛查提取参考;专业皮肤科采集用户通过医院皮肤科专业采集设备(如皮肤镜)获取服务,通过皮肤科专业的皮肤镜采集并上传从而辅助医生进行医疗提取,可获得精确的提取结果。

①提取云

本文的提取云由云服务、模型存储与数据缓存组成,主要提供提取与数据存储传输服务。其中,模型存储将模型训练的模型存储用于提取服务,提取模型根据模型训练结果及时更新以提高提取准确率;数据缓存将用户数据缓存用于模型训练的下载,从而提供给样本源,用于核验提取结果与扩充样本集。

②模型训练

本文的模型训练基于残差金字塔多尺度编码网络,在训练服务器中结合数据提供方提供的训练样本进行训练。训练服务器由数个图形处理单元(Graphics Processing Unit, GPU)与中央处理器(Central Processing Unit, CPU)组成,部署本文基于深度残差金字塔的皮肤病智能提取模型与训练环境进行模型训练,并将参数上传到提取云。

③样本源

在数据提供中,样本由 ISIC 提供的开源数据与医疗机构中专业的皮肤镜采集人员采集并由皮肤科医生进行标注的样本构成;皮肤科医生提取采集人员提供的采集数据并提供提取结果,为样本数据进行标注,作为训练样本提供给模型训练;同时皮肤科医生从模型训练下载提取云的提取结果进行结果分析,提取后标注数据以扩充训练样本;皮肤科医生会诊解决疑难杂症,审查提取云的对比结果并评估模型的优势与不足,并根据研讨结果与提取云对比结果调整本文模型的网络结构与网络参数等,优化模型训练并更新参数。

2 基于深度残差金字塔的病变分割与特征提取机制

在本节中,主要阐述基于深度残差金字塔的皮肤病智能提取模型原理与结构设计。为了提高特征提取与训练结果的准确度,本文将传统的深度残差网络[24]与特征金字塔[25]结合,扩展为深度残差金字塔网络。对比主流的单一尺度网络输出,本文将瓶颈层划分为多尺度编码网络,通过提取多尺度特征输出提取结果。此外,为了解决样本不均衡问题,本文设计了一种新颖的损失梯度监督机制,通过焦点损失增加模型对难分样本的关注,同时通过梯度协调机制减少对特别难分样本与离群点的关注,从而减小类别不平衡对提取结果的影响。

2.1 深度残差金字塔多尺度编码网络

本文设计的深度残差金字塔多尺度编码网络基于 CNN。本文模型与 SVM、传统神经网络(Neural Network, NN)相比,通过输入的原始图像提取特征,因此具有更高的准确率。除了输入层与输出层外,由于应用场景的特定性与研究问题的复杂性,本文的 CNN 进一步考虑设置一个或多个隐藏层,该隐藏层的网络拓扑结构在层的类型与数量上有所不同,因此,应用于色素性皮肤病变分割与特征提取的模型分为五层,即输入层、卷积层、激活层、池化层与瓶颈层,详述如下:

①输入层

在输入层,本文模型加载预处理后的样本集,即 RGB 三通道皮肤镜图像。通过色彩恒常性算法、数据增强算法与归一化将图像预处理后,减去训练集图像像素均值并自动缩减为 224 大小以适应网络。

②卷积层

卷积层是本文 CNN 的主要模块,该层作用是从色素性皮肤病样本集中提取不同的输入特征。在反向传播时,卷积核在图像维度上平移。低层卷积层只能提取低级特征,高层卷积可以通过迭代提取低层卷积中的复杂特征。卷积层的色素性皮肤病单元组织为特征图,每个单元通过一定大小的卷积核作用于上一层特征图的局部区域。与主流 CNN 不同,本文 CNN 基于 ResNet。在 ResNet 的基础上,使用 3 个 3×3 卷积核来代替 7×7 卷积核,使用 2 个 3×3 卷积核来代替 5×5 卷积核,即使得 5×5 卷积的全连接网络在 5×5 区域滑动。这样做旨在保证具有相同感知野的条件下提升网络的深度,从而在一定程度上提升神经网络的效果。先采用一个 3×3 的卷积滤波器卷积,然后再采用一个全连接层连接这个 3×3 卷积输出,从而使得全连接层作为 3×3 卷积层。通过该方法即可采用两个 3×3 卷积级联(叠加)代替一个 5×5 卷积。本文模型包含 5 个卷积块,第 1、2 块内有 2 个卷积层,其余 3 块内有 3 个卷积层,所有卷积都使用线性整流函数(Rectified Linear Unit, ReLU)激活函数。这 5 个卷积块的卷积核总数分别为 64、128、256、512、512,每块内的卷积核数目相同。卷积层全部为 3×3 的小卷积核,在增强网络容

量与模型深度的同时减少卷积参数的数量,有效降低计算复杂度。

③激活层

在卷积层后,作用为使训练网络适应非线性。激活层在神经网络中使用非线性函数,如 ReLU 与双曲正切函数(tanh)。对于本文模型而言,ReLU 可以解决梯度消失问题,使网络训练更快且准确性不会受到较大影响。

④池化层

在激活层后,作用为降维,以减少训练参数的数量与计算负担,从而减少训练时间,有效避免模型过拟合,同时降低计算的复杂度。池化有最大池化、最小池化与平均池化等。将输入划分为多个矩阵,对于每个子区域,最大池化过程中保留输出的最大值。在池化层,本文模型中每个卷积块后面都连接一个最大池化层,大小为 2×2。

⑤瓶颈层

处于本文模型最后位置,基于 ResNet[24] 设计瓶颈层。主流 CNN 利用高层特征解决在提取多尺度特征时存在严重不足的问题。但该方法存在的缺陷是,较小特征在下采样过程中极易丢失,这是因为较小的特征本身具有较少的像素信息。为了解决这一缺陷,本文采用深度残差金字塔多尺度编码结构进行多尺度特征的增强。如图 2 所示,瓶颈层由多个瓶颈层 1 和瓶颈层 2 组成,由卷积块、批标准化和线性整流组成。瓶颈层带来的好处是大幅减少计算量,提高网络的训练速度。引入该结构可以对每一种尺度的图像进行特征提取,从而得到多尺度的特征表示,虽然特征金字塔多尺度编码增大了计算量和内存占用量,但残差的瓶颈层大幅减少了计算量和内存占用,同时提高了网络的特征提取能力,从而提高了网络的预测性能和精度。

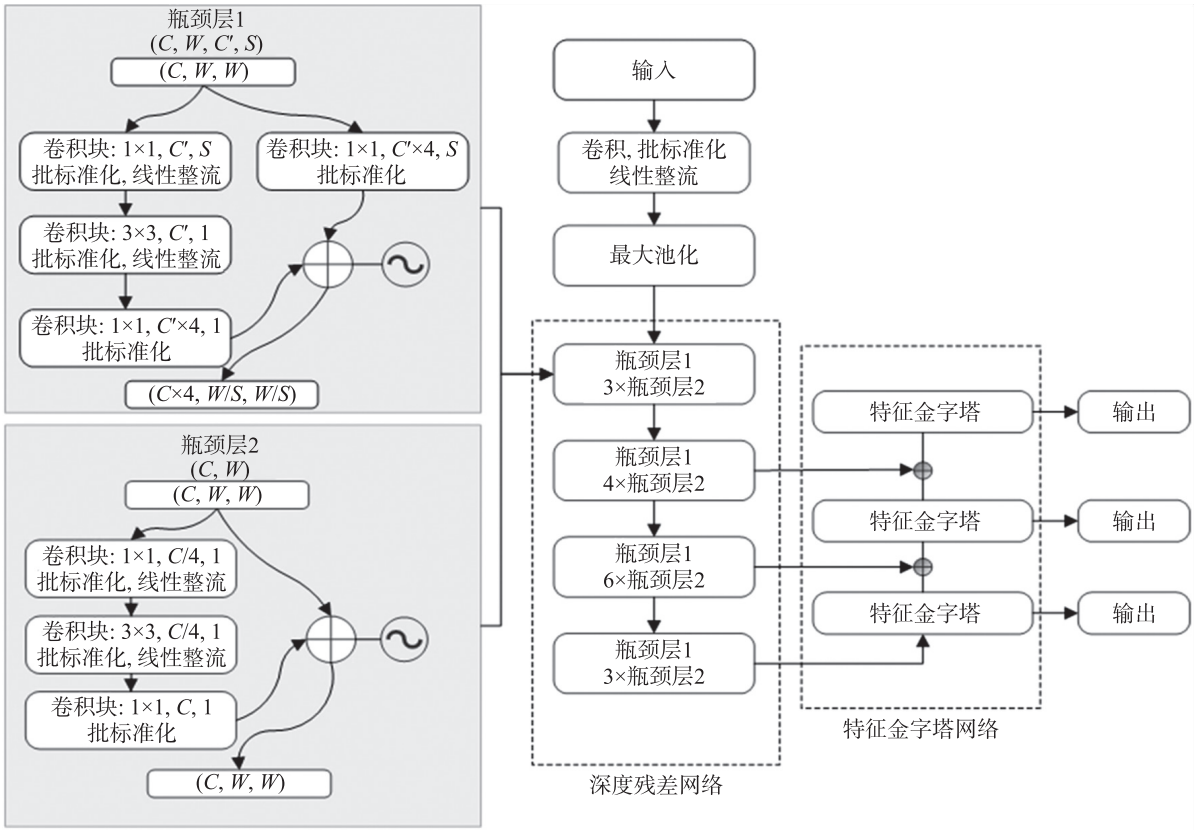


图 2　深度残差金字塔多尺度编码网络

2.2　基于焦点损失的损失梯度监督机制

本文引入损失旨在解决色素性皮肤病类别间存在的极大不平衡问题。采用的损失函数从二元分类的交叉熵(Cross Entropy,CE)损失引入。

CE 损失定义为:

$$\mathrm{CE}(p,y)=\begin{cases}-\log p & y=1\\ -\log(1-p) & \text{otherwise}\end{cases} \tag{1}$$

在式(1)中,对二分类任务而言,$y\in\{\pm1\}$。$p\in[0,1]$,表示模型对带有标签 $y=1$ 的概率估计。为了简化,用 p_t 代替 p,即为式(2):

$$p_t=\begin{cases}p & y=1\\ 1-p & \text{otherwise}\end{cases} \tag{2}$$

则(1)可重写为:

$$\mathrm{CE}(p,y)=\mathrm{CE}(p_t)=-\log(p_t) \tag{3}$$

解决类别不平衡的常用方法是为 $y=1$ 引入加权因子 $\alpha\in[0,1]$，为 $y=-1$ 引入加权因子 $1-\alpha$。为了符号上方便，类似于定义 p_t 的方式定义 α_t，即将 α 平衡后的 CE 损失写为：

$$\mathrm{CE}(p_t)=-\alpha_t\log(p_t) \tag{4}$$

这种损失是对 CE 的简单扩展，即本文梯度协调机制的实验基准。

在训练过程中遇到的类别失衡影响了交叉熵损失，容易导致分类的负值构成了损失的大部分并影响梯度。尽管上文加权因子 α 平衡了正负样本所占权值，但其并未区分简单与复杂样本。为解决该问题，在交叉熵损失中添加一个调制系数 $(1-p_t)^{\gamma}$，且参数 $\gamma\geqslant 0$。焦点损失[22]（Focal Loss，FL）定义为：

$$\mathrm{FL}(p_t)=-(1-p_t)^{\gamma}\log(p_t) \tag{5}$$

当样本分类错误且 p_t 较小时，调制系数接近 1，并且损失不受影响。当 $p_t\to 1$ 时，因子变为 0，并且对分类准确的样本的损失进行了加权。参数 γ 对简单样本的权重进行平滑调整。当 $\gamma=0$ 时，FL 等效于 CE，并且随着 γ 的增加，调制系数的作用也会增加。调制系数减少了简单样本的损耗影响，并扩展了样本接收低损耗的范围。使用 α 损失平衡焦点损失为：

$$\mathrm{FL}(p_t)=-\alpha_t(1-p_t)^{\gamma}\log(p_t) \tag{6}$$

上述机制存在的缺陷为：如果让模型过多关注难分样本（如样本中的离群点），已经收敛的模型会因为这些离群点造成被判别错误的问题，因此对于易分样本和难分样本都不应过度关注；同时参数的取值全从实验得出且各参数要联合实验，因为其取值会相互影响，这不仅降低了模型的准确率，还增加了训练的计算量。

因为易分样本和特别难分样本数量都要比一般样本多一些，因此本文衰减单位区间数量多的那类样本（即物理密度），则改进的损失梯度监督机制[21]为：

$$\mathrm{GD}(g)=\frac{1}{l_e}\sum_{i=1}^{N}\delta_e(g_k,g) \tag{7}$$

式中：$\delta_e(g_k,g)$ 表示样本梯度模长分布在 $(g-e/2,g+e/2)$ 范围内的样本个数，$l_e(g)$ 表示区间 $(g-e/2,g+e/2)$ 的长度。

综上所述，本文平衡后采用的损失梯度监督机制为：

$$\mathrm{Loss}=\lambda\sum_{i=1}^{N}\frac{\mathrm{CE}(p_i,p_i^{*})}{\mathrm{GD}(g_i)}+(\lambda-1)\alpha_t(1-p_t)^{\gamma}\log(p_t) \tag{8}$$

式中：λ 为调制因子，取值范围为 $[0,1]$，取值基于样本和实验调整。

3 实验

3.1 样本集

本文样本集为红绿蓝颜色系统（Red Green Blue，RGB）三通道皮肤镜图像，样本集分为训练集和测试集，由 ISIC 开源的皮肤病样本集与医疗机构中专业的皮肤镜采集人员采集并由皮肤科医生进行标注的样本构成。本文样本为使用各种类型的皮肤镜采集的不包括粘膜与指甲的皮肤图像，每个病变图像包括一个原发病变，不包括较小的继发性病变或其他色素区。病变样本的分布符合现实情况，即良性病变多于恶性病变，但恶性病变种类较多。

3.2 实验

本文训练集构成为 12609 张皮肤镜病变图像，主要包含的病变类型为黑色素瘤（MEL）、黑色素细胞痣（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）和血管病变（VASC），训练集选自 ISIC2018 和医疗机构脱敏采样图片，如图 3 所示，图片不带有患者任何个人信息，从而保护患者隐私。训练集中人工标注样本有 220 张；由于样本人工标注不统一，准确度低，无法提供统一准确的标签用于训练，因此本文选取 2 600 张图片进行蒙版处理，作为标签训练的样本集，未经处理的 9 789 张图片作为无标签训练样本集。为了避免采集到的训练集样本存在毛发、气泡、光照等因素对实验结果的影响，本文采用去噪、色彩恒常性等算法对图片进行预处理，同时对样本集平移、翻转、拉伸、旋转扩充训练集至 20 000 张。

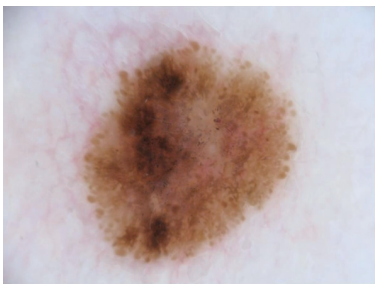


图 3 训练集图片示例

基于标签训练的样本集主要是人工像素定界输出 PNG 格式的二进制蒙版图像，作为准确的病变类型边界结果。指示每个输入病变图像中原发皮肤病变的位置，如图 4 所示。蒙版图像与其对应的病变图像具有完全相同的尺寸。为了无损压缩图片，蒙版图像被编码为单通道（灰度图）8 位 PNG，其中像素分为 0 与 255。像素为 0 代表该区域为图像的背景或原发病变位置以外的区域；像素为 255 代表该

区域为图像原发病变位置内的区域。由于原发皮肤病变是单个连续区域,因此蒙版图像仅包含单个连续病变位置区域。

图 4　标准病变边界

为避免偶然性对实验结果的影响,采用 k 折交叉验证,即运行 k 次迭代完成 k 折交叉验证,每次迭代时每个样本只在训练集中或者测试集中,且在测试集中出现一次。测试数据选取 1 000 张验证,图片尺寸与输入数据一致,如图 5 所示。

图 5　测试图片示例

通过本文模型分割与提取后的输出结果如图 6 所示,其中病变位置用灰色与白色像素标记,特征部位用白色像素标记,非病变位置用黑色像素标记。特征部位所在属于病变位置,是有助于该类疾病诊断分类的核心部位;病变位置由蒙板图像半监督学习得到,特征部位由特征金字塔网络提取像素得到。在提取时,特征金字塔网络在卷积后得到小尺寸特征图,通过上采样还原后得到更大尺寸的特征图。将得到的特征图叠加但不增加通道数,使得在不增加图像维度的情况下增加了每一维的信息量,从而提高了特征提取的准确性。

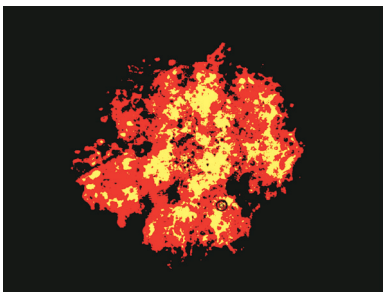


图 6　输出结果

上文阐释了本文的主要研究内容,即构建了一个用于病变分割与特征提取的神经网络模型,提高了分割与提取的准确度。为了进一步微观地评判模型,本文采用 Jaccard 系数作为评判分割与提取准确度的标准。设 T 为真实病变边界分割与特征提取结果,P 为预测边界分割与特征提取结果,则 Jaccard 系数计算为式(9):

$$\text{Jaccard}=\frac{T\times P}{T+P-T\times P} \tag{9}$$

表 1 展示了 CE、FL 与本文损失梯度监督机制相对比的情况。从表中可以看出,对于大样本,CE 与本文所提的梯度监督机制准确率接近,而 FL 准确率较低,这是因为 FL 会过度关注难分样本和离群点,导致整体准确率偏低。在样本充足的情况下 CE 的模型训练效果接近本文,拥有良好的分割与提取能力;而 FL 会过于关注难分样本从而降低模型整体的预测能力,导致已经收敛的模型会因为这些难分离群点造成被判别错误的问题。

表 1　不同损失方案的分割与提取结果对比(Jaccard)

损失	大样本	小样本	平均
CE	0.817	0.529	0.752
FL	0.798	0.815	0.807
本文	0.831	0.812	0.823

对于小样本,FL 和本文所提梯度损失监督机制准确率接近,而 CE 的分割与提取准确率极低,这是因为 CE 的局限性导致其无法处理样本不均衡问题,只能基于样本数量给出平均分割与提取的结果,模型难以充分训练,从而导致平均准确率较低。说明 FL 和本文提出的梯度损失监督机制有效解决了样本不平衡问题。而由于 FL 参数的取值会相互影响,因此其各参数要联合实验从而得出取值,导致 FL 的整体精确度低于本模型。因此本文所提损失方案的平均准确率高于 FL 和 CE。

图 7 所示为本文训练样本准确率(accuracy)和损失(loss)随着训练迭代数(epoch)的变化曲线。

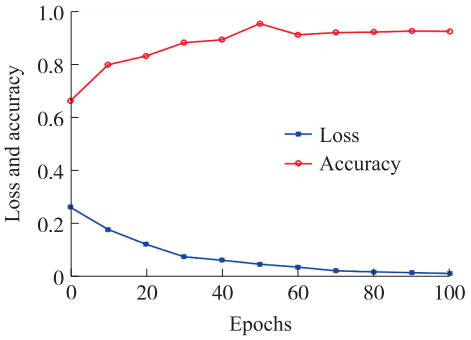


图 7　损失和训练准确率随迭代次数变化结果

随着 epoch 的增加,本文模型的 loss 逐渐降低,并在 epoch>70 时趋于收敛,表明本文的损失模型设计合理;训练集的 accuracy 逐渐提高,并在 epoch>70 时趋于收敛,表明本文的训练模型设计高效。

如表 2 所示,本文的分割与提取模型较主流的分割与提取模型在准确率方面有较大优势,原因在于:SVM 和 NN 由于无法通过输入的原始图像提取特征,因此准确率较低。与本文模型比,其余模型在特征提取、尺度依赖、损失等方面都有一定局限,因此准确率较本文模型低。本文分割与提取的准确率高是因为,本文将瓶颈层划分为多尺度编码网络,通过提取多尺度特征输出提取结果。通过这种网络结构,可以有效避免特征在卷积、池化的过程中丢失,因此特征提取的准确性高。为了避免梯度爆炸、梯度消失、过拟合和欠拟合问题,提高网络的训练效率,本文网络基于深度残差网络。对于样本不均衡问题,本文设计的损失梯度监督机制,通过焦点损失增加模型对难分样本的关注,同时通过梯度协调机制减少对特别难分样本与离群点的关注,从而减小类别不平衡对提取结果的影响,因此与主流方案对比具有较高的准确率。

表 2 不同模型分割与提取结果 Jaccard 系数对比

模型	Jaccard
SVM	0.382
NN	0.425
U-Net[16]	0.729
Inception v3[26]	0.747
VGG16[20]	0.786
ResNet50[27]	0.792
本文	0.823

如图 8 所示,本文同时验证了基于 k 折交叉验证的结果,从图中可以看出,本文提出的模型精确率随 k 值变化波动较小,且均高于其他主流模型,表明本文模型稳定性好。k 折交叉验证避免了偶然性对实验结果造成的影响,验证了预测能力的稳定性,有效避免了偶然误差和数据集随机取值对结果的影响。

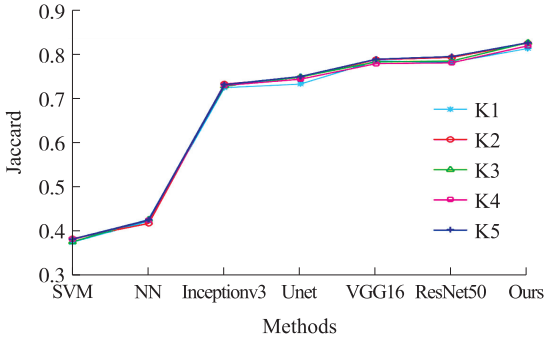


图 8 不同模型 k 折交叉分割与提取结果对比

4 总结

本文提出了一种基于深度残差金字塔的皮肤病变分割与特征提取机制。具体而言,设计了基于深度残差金字塔的多尺度编码网络,以提取多尺度特征从而提高模型的提取能力。此外,设计了基于焦点损失的梯度损失监督机制,可有效解决样本不均衡问题同时克服焦点损失的局限性。最后,与现有方法的对比结果表明,本文所提机制可以显著提高提取准确性并减少样本不均衡问题带来的影响。

参考文献:

[1] Olugbara O O,Taiwo T B,Heukelman D. Segmentation of Melanoma Skin Lesion Using Perceptual Color Difference Saliency with Morphological Analysis[J]. Mathematical Problems in Engineering,2018,2018(1):1524286.

[2] Zhou H,Schaefer G,Sadka A H,et al. Anisotropic Mean Shift Based Fuzzy C-Means Segmentation of Dermoscopy Images[J]. IEEE Journal of Selected Topics in Signal Processing,2009,3(1):26-34.

[3] Abbas Q,Celebi M E,Fondón García I,et al. Lesion Border Detection in Dermoscopy Images Using Dynamic Programming[J]. Skin Research and Technology,2011,17(1):91-100.

[4] Celebi M E,Kingravi H A,Uddin B,et al. A Methodological Approach to the Classification of Dermoscopy Images[J]. Computerized Medical Imaging and Graphics,2007,31(6):362-373.

[5] Capdehourat G,Corez A,Bazzano A,et al. Toward a Combined Tool to Assist Dermatologists in Melanoma Detection from Dermoscopic Images of Pigmented Skin Lesions[J]. Pattern Recognition Letters,2011,32(16):2187-2196.

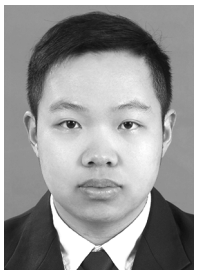
[6] Situ N,Yuan X,Chen J,et al. Malignant Melanoma Detection by Bag-of-Features Classification[C]//2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society(EMBS),Vancouver,Canada,IEEE,2008:3110-3113.

[7] Pathan S,Prabhu K G,Siddalingaswamy P C. Hair Detection and Lesion Segmentation in Dermoscopic Images Using Domain Knowledge[J]. Medical & Biological Engineering & Computing,2018,56(11):2051-2065.

[8] Khan M A,Akram T,Sharif M,et al. An Implementation of Normal Distribution-Based Segmentation and Entropy-Controlled Features Selection for Skin Lesion Detection and Classification[J]. BMC Cancer,2018,18(1):638.

[9] Filali I,Belkadi M. Multi-Scale Contrast Based Skin Lesion Segmentation in Digital Images[J]. Optik,2019,

185:794-811.
[10] 肖秦,张雨昕,王颖瀛,等. 基于卷积神经网络的酵母出芽特征识别方法[J]. 电子器件,2023,46(4):1096-1103.
[11] 吴明孝,杨威,孙武魁,等. 基于结构优化卷积深度网络的变压器溶解气体分析[J]. 电子器件,2024,47(3):788-795.
[12] Codella N,Cai J,Abedini M,et al. Deep Learning,Sparse Coding, and SVM for Melanoma Recognition in Dermoscopy Images [C]//2015 6th International Workshop on Machine Learning in Medical Imaging(MLMI),Munich,Germany,Springer Verlag,2015:118-126.
[13] Kawahara J,Bentaieb A,Hamarneh G. Deep Features to Classify Skin Lesions[C]//2016 IEEE International Symposium on Biomedical Imaging(ISBI),Prague,Czech Republic,IEEE,2016:1397-1400.
[14] 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.
[15] Li Y X, Shen L L. Skin Lesion Analysis Towards Melanoma Detection Using Deep Learning Network[J]. Sensors,2018,18(2):556.
[16] Codella N C F, Nguyen Q B, Pankanti S, et al. Deep Learning Ensembles for Melanoma Recognition in Dermoscopy Images[J]. IBM Journal of Research and Development,2017,61(4-5):5.
[17] Demyanov S, Chakravorty R, Abedini M, et al. Classification of Dermoscopy Patterns Using Deep Convolutional Neural Networks[C]//2016 IEEE International Symposium on Biomedical Imaging(ISBI),Prague,Czech Republic,IEEE,2016:364-368.
[18] Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation[C]//2015 18th International Conference on Medical Image Computing and Computer-Assisted Intervention(MICCAI), Munich, Germany,Springer Verlag,2015:234-241.
[19] Nasr-Esfahani E, Samavi S, Karimi N, et al. Melanoma Detection by Analysis of Clinical Images Using Convolutional Neural Network [C]//2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society(EMBS),Orlando,USA,IEEE,2016:1373-1376.
[20] Yu L,Chen H,Dou Q,et al. Automated Melanoma Recognition in Dermoscopy Images via Very Deep Residual Networks[J]. IEEE Transactions on Medical Imaging,2016,36(4):994-1004.
[21] 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]//2018 IEEE 15th International Symposium on Biomedical Imaging(ISBI), Washington, DC,USA,IEEE,2018:168-172.
[22] Lin T Y,Goyal P,Girshick R,et al. Focal Loss for Dense Object Detection [C]//2017 IEEE International Conference on Computer Vision(ICCV), Venice, Italy, IEEE,2017:2999-3007.
[23] Li B,Liu Y,Wang X. Gradient Harmonized Single-Stage Detector[C]//2019 33rd AAAI Conference on Artificial Intelligence,Honolulu,HI,USA,AAAI,2019:8577-8584.
[24] He K,Zhang X,Ren S,et al. Deep Residual Learning for Image Recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition(CVPR), Nevada, NV,USA,IEEE,2016:770-778.
[25] Lin T Y,Dollár P,Girshick R,et al. Feature Pyramid Networks for Object Detection[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition(CVPR), Honolulu,HI,USA,IEEE,2017:936-944.
[26] Yoo T K,Choi J Y,Kim H K et al. Adopting Low-Shot Deep Learning for the Detection of Conjunctival Melanoma Using Ocular Surface Images[J]. Computer Methods and Programs in Biomedicine,2021,205:106086.
[27] Akkoca Gazioğlu B S,Kamaşak M E. Effects of Objects and Image Quality on Melanoma Classification Using Deep Neural Networks[J]. Biomedical Signal Processing and Control,2021,67:102530.



董春序(1997—),男,南京邮电大学在读硕士研究生。主要研究方向为卷积神经网络与边缘智能,tcdcx1@163.com;

李　雪(1986—),女,南京医科大学附属妇产医院(南京市妇幼保健院)副主任医师。主要研究方向为色素性皮肤病、基于AI的皮肤病诊断,lixuedm@163.com;

陈思光(1984—),男,通讯作者,南京邮电大学物联网学院院长助理、教授、硕士生导师。主要研究方向为边缘智能与智慧物联网等,sgchen@njupt.edu.cn。