

基于深度学习的痰湿体质高血压舌象识别

朱倩倩¹, 王澜¹, 姜楠¹, 董昌武²

1. 安徽中医药大学中医学院, 安徽 合肥 230012; 2. 安徽中医药大学第二附属医院名医堂, 安徽 合肥 230061

【摘要】目的:使用深度学习语义分割模型及残差神经网络模型对痰湿体质人群是否患有高血压病进行舌象的客观化识别分析,推动舌象研究的现代化进程,为中医临床决策提供更为客观、科学的依据。**方法:**首先使用LabelMe图像标签标注工具对547例受试者的舌象舌体区域进行划分标注,再使用U-Net分割算法进行舌体分割,将舌体从复杂的背景中单独分离出以便后续的分析。后续研究引入了ResNet-34、ResNet-50以及YOLOv5 3种深度学习模型,将痰湿体质高血压患者舌象和痰湿体质亚健康者舌象进行分类,构建分类模型,绘制混淆矩阵并计算F1值、准确率等对分类模型效果进行评价。**结果:**实验结果显示,3种模型在本次分类任务中均有较好的表现。ResNet-34模型F1值为91.46%,准确率为92.87%,精准率为90.48%,召回率为92.89%。ResNet-50模型总体上更优,F1值为92.08%,准确率为93.05%,精准率为95.26%,召回率为89.11%。YOLOv5模型总体准确率为85.6%,在痰湿体质高血压患者和痰湿体质亚健康者两个类别上,分别取得了85.3%和85.7%的准确率。**结论:**ResNet-34、ResNet-50、YOLOv5在本次分类任务中均表现优异,以ResNet-50最佳。证明了深度学习模型可以较好地完成舌象的分类识别任务,体现了深度学习技术在中医舌诊自动化分类中的巨大潜力,也为中医诊断的现代化、客观化提供了有力的技术支撑。

【关键词】舌象;深度学习;高血压;痰湿体质;神经网络

【中图分类号】R318

【文献标志码】A

【文章编号】1005-202X(2025)04-0534-08

Deep learning-based tongue image recognition for hypertension with phlegm-dampness constitution

ZHU Qianqian¹, WANG Lan¹, JIANG Nan¹, DONG Changwu²

1. College of Chinese Medicine, Anhui University of Chinese Medicine, Hefei 230012, China; 2. Eminent Physician Clinic, The Second Affiliated Hospital of Anhui University of Chinese Medicine, Hefei 230061, China

Abstract Objective To objectively identify whether people with phlegm-dampness constitution suffer from hypertension or not using deep learning semantic segmentation model and residual neural network, so as to promote the modernization of tongue manifestation research, and provide a more objective and scientific basis for clinical decision-making in traditional Chinese medicine (TCM). **Methods** The tongue regions of 547 subjects were outlined and labeled using the Label Me image labeling tool, followed by tongue body segmentation using the U-Net segmentation algorithm which separated the tongue body from the complex background. In the subsequent study, 3 deep learning models, namely ResNet-34, ResNet-50 and YOLOv5, were used to classify the tongue manifestations of hypertensive patients and the sub-health both with phlegm-dampness, and to construct the corresponding classification models whose performances were objectively evaluated by drawing confusion matrix and calculating F1 value and accuracy. **Results** The experimental results showed that all 3 models performed well in the classification task. ResNet-34 vs ResNet-50 had F1 values of 91.46% vs 92.08%, accuracies of 92.87% vs 93.05%, precisions of 90.48% vs 95.26%, and recall rates of 92.89% vs 89.11%. YOLOv5 had an overall accuracy of 85.6%, achieving 85.3% and 85.7% accuracies in the specific classifications for hypertensive patients with phlegm-dampness and the sub-health with phlegm-dampness. **Conclusion** All 3 models (ResNet-34, ResNet-50 and YOLOv5) performed well in the classification task, with ResNet-50 being the best. It proves that the deep learning model can better accomplish the

【收稿日期】2024-11-20

【基金项目】国家自然科学基金(82174277);安徽省中医药科技攻关专项(202303a07020008);安徽省重点研究与开发计划项目(2022h11020018);安徽省临床医学研究转化专项(202304295107020107);新安医学与中医药现代化研究所“揭榜挂帅”项目(2024CXMMTCM002);董昌武安徽省名医工作室项目(皖中医药发展秘[2024]19号)

【作者简介】朱倩倩,硕士研究生,研究方向:四诊客观化及心脑血管疾病中医证候,E-mail: qq1127666@163.com

【通信作者】董昌武,教授,博士生导师,研究方向:四诊客观化及心脑血管疾病中医证候,E-mail: dcw1018@aliyun.com

classification and recognition of tongue manifestations, which reflects the great potential of deep learning in the automated classification for TCM tongue diagnosis, and also provides a strong technical support for the modernization and objectivity of TCM diagnosis.

Keywords: tongue manifestation; deep learning; hypertension; phlegm-dampness constitution; neural network

前言

高血压作为一种普遍且复杂的慢性疾病,其特征是在静息状态下动脉血压异常增高。疾病初期,患者可能会出现头晕、心悸、睡眠障碍及情绪波动等症状,随着病程的发展,疾病将逐渐侵蚀心、脑、肾及视网膜等关键器官,导致功能损害或器质性病变。高血压是心脑血管疾病致死的核心因素之一,全球范围内,高血压的高发病率已成为一个不容忽视的公共卫生问题^[1-2]。痰湿质高血压是指在中医体质分类中,属于痰湿体质的人群所患的高血压病。这类患者往往因脾胃功能受损,导致脾虚不运,进而痰浊内生,形成痰湿体质,并引发高血压^[3]。痰湿体质的人群往往存在代谢异常和气血运行不畅的问题,这些因素可能增加患高血压的风险。及时发现高血压,尽早进行干预和治疗,可以有效预防高血压引起的严重并发症。舌诊作为中医诊断的重要手段之一,在辨别痰湿体质患者是否患有高血压方面具有重要意义。《辨舌指南》^[4]指出:“辨舌质,可决五脏之虚实;视舌苔,可察六淫之浅深”,这一观点强调了舌象在反映人体脏腑功能、气血状况及疾病性质方面的重要作用。舌象作为人体内部环境的外在表现,其形态、颜色、湿润度及舌苔的变化,能直接或间接地揭示人体健康状况的微妙变化。在高血压等疾病的诊疗过程中,中医舌诊发挥着不可替代的作用,其独特的价值与深远的意义,值得深入探究与传承。

近年来,深度学习在医学领域的应用日益广泛,卷积神经网络(Convolutional Neural Networks, CNN)作为深度学习中最具代表性的网络结构之一,擅长处理图像数据^[5]。CNN在图像特征提取方面展现出的卓越性能,可显著提高后续诊断的准确性、高效性与客观性,从而有力推动了中医舌诊领域向现代化、智能化方向的转型与发展。因此,CNN在中医舌诊领域具有广阔的应用前景和重要的应用价值^[6-8]。ResNet(Residual Network)是CNN的一种扩展和改进,它在CNN的基础上增加了残差模块,这些模块通过引入跳跃连接(或残差连接),允许网络中的信号直接传播,这种方法使其具有出色图像识别能力。ResNet-34和ResNet-50都拥有卓越的特征提取能力,它们通过多层的卷积和残差模块捕捉从简单到复杂的图像特征,可以应对不同的图像分类需求。YOLOv5(You Only Look Once version 5)虽然主要用

于目标检测,但其强大的特征提取能力和高效的性能使其在处理一些特殊的图像分类任务时也能发挥较好的效果。因此本研究首先利用深度学习的语义分割模型U-Net从舌象数据中分离出舌体和背景,再使用残差神经网络 ResNet-34、ResNet-50 模型和YOLOv5 目标检测算法将舌象数据按照痰湿体质是否患有高血压将其分类^[9-10]。本研究旨在实现无创、高效、智能的痰湿质高血压辅助诊断,为痰湿体质高血压舌象的现代化分析提供新的思路和方法。

1 资料与方法

1.1 研究对象

本研究数据来源于2023年11月~2024年6月在安徽中医药大学第一附属医院心内科及安徽中医药大学第二附属医院名医堂招募的受试者,经筛选后纳入痰湿体质的受试者共547人,其中痰湿体质患高血压者315名,痰湿体质亚健康者232名。本研究通过安徽中医药大学第二附属医院医学伦理委员会批准(伦理批准号为:2022-zjks-25)。

1.2 数据采集与纳排标准

使用安徽中医药大学与合肥云诊信息科技有限公司联合研发的“中医舌诊AI智慧诊疗系统”采集舌象。该系统将舌面、舌下、面部照片结合问诊信息,可在10 s内快速辨识体质信息,本研究即使用此系统辨识痰湿体质。所有受试者舌象图片的采集时间均设定在早上8:00~11:00,以确保受试者在生理状态下尽量一致。光源选择上午时段充足、柔和的自然光线,保证光线均匀照射在舌面。避免背对阳光入口处,如窗户、门,避免在易反射的有色物体旁拍摄,如有色窗帘等。要求受试者拍摄时处于空腹状态,消除饮食因素对数据采集的干扰。受试者在拍摄前至少有5 min的安静休息。拍摄时受试者被要求采取正坐位,张口伸舌,舌体应尽量平展,舌尖自然向下,以充分暴露舌面的各个区域,确保采集到的舌象信息完整且清晰。同一个受试者采集两次舌象图片,选取质量较高的一张(选取标准:图片清晰明了,舌体边缘无模糊或变形现象。舌面无明暗不均或阴影现象。舌色接近肉眼所见的实际颜色,无偏色或失真现象。完整展示舌体的全貌,包括舌尖、舌边、舌根等部位。图片中无其他干扰因素,如口唇部或其他物体的遮挡,以及唾液、食物残渣等污染物的干

扰)。舌象数据纳入标准:①符合痰湿体质及高血压病诊断标准的患者;②年龄大于18周岁且小于85周岁;③一般资料较为完整;④患者知情同意。排除标准:①患有恶性、消耗性疾病;②年龄小于18周岁或大于85周岁;③资料信息缺失;④因精神、心理、社会或其他任何原因无法配合本研究过程。

1.3 研究方法

实验前期收集并筛选了数据集,使用 Github 开源工具 LabelMe 对收集的舌象图像进行舌体标签的标注,通过旋转、镜像、缩放、高斯噪声等数据增强方式对数据集进行进一步扩充。部署开源深度学习框架,基于深度学习 Pytorch 框架,使用 U-Net 网络分割舌体,使用 ResNet34、ResNet50、YOLOv5 网络进行舌象二分类。

1.4 舌象数据预处理

1.4.1 标注与分割 使用 LabelMe^[11]图像标签标注工具对舌体区域进行划分标注,为后续的图像处理及深度学习模型提供精确的标签数据,见图1。再采用 U-Net^[12]分割算法进行舌体分割,从复杂的背景中精确地提取出目标舌体。这种处理可以有效去除背景的无关信息,突出目标区域,确保每个像素都属于目标本身。U-Net 是一种用于医学图像分割的卷积神经网络,它通过一种精巧的上采样方式有效地使用数据,以便在只有少量标记图像的情况下也能训练得到高性能的模型。U-Net 的核心特点是其对称的“U”形结构,这使得它在捕捉图像的上下文信息和精细的局部特征方面表现出色。此外,该模型采用了多尺度特征融合的策略,通过在编码器和解码器部分之间的跳跃连接来传递特征信息,这不仅增强了网络对图像细节的识别能力,还改善了分割的准确性。在中医诊断中,舌象是重要的健康指标之一。利用 U-Net 网络进行舌象分割,可以自动识别和提取舌体区域,从而辅助中医专家进行更精确的诊断^[13-15]。



图1 舌象标注示意图

Figure 1 Schematic diagram of tongue image labeling

1.4.2 数据增强处理 数据增强是深度学习中的一种常用技术,如缩放、翻转、随机旋转、噪声等,旨在通过对训练数据进行随机变换,生成更丰富的样本,从而提高模型的泛化能力。本次处理首先将 U-Net 网络分割好的 547 张舌象图片的分辨率统一为 224×224 像素,再使用缩放、旋转、翻转、像素平移、添加噪声的图像处理方式将图像数据集扩增,最终得到 2 735 张舌象,其中痰湿体质患有高血压患者图像 1 575 张,痰湿体质无高血压者图像 1 160 张。这样的处理使得模型可以适应图像的对称性变化、专注于图像的不同部分以及在不同角度下识别物体的能力。这些技术共同提高了模型对图像中目标位置、大小、角度等多种变化的鲁棒性,减少过拟合,增强对新样本的适应性^[16]。

本研究得到 U-Net 分割模型评价指标:平均交并比 (Mean Intersection over Union, MIOU) 为 98.37%; 平均类别精度 (mean Per-class Accuracy, mPA) 为 99.26%; 平均精准率 (mean Precision, mPrecision) 为 99.1%; 平均召回率 (mean Recall, mRecall) 为 99.26%。由上述指标可见本研究基于 U-Net 的舌象自动分割模型的分割结果相对稳定,分割效果较为理想,可为下一步分析做好准备。分割效果示例见图2。



图2 舌象分割示意图

Figure 2 Schematic diagram of tongue image segmentation

2 基于 PyTorch 框架的 CNN 舌象识别

2.1 实验环境

模型运行环境:显卡为 NVIDIA GeForce RTX 4090D;操作系统为 Windows 11;编程语言选用 Python(版本:3.10);深度学习框架为 PyTorch。

2.2 模型训练

本文主要基于 ResNet-34、ResNet-50 两种残差网络及 YOLOv5 算法进行训练,构建舌象分类模型。相对于传统的 CNN,ResNet^[17]在小样本学习领域具有一定的优势:ResNet 通过引入残差连接模块有效缓解了梯度消失问题、处理图像数据时能够显著减

少模型参数的数量、可以有效地学习小样本数据集中的深层特征、能通过迁移已在大规模数据集上训练得到的特征表示,来实现良好的泛化能力^[18]。残差模块结构分为 BasicBlock 和 BottleNeck, 见图 3。本文中 ResNet-34 使用 BasicBlock 这一较相对较浅的残差模块, 每个 BasicBlock 通常包含两个卷积层, 每个卷积层后面跟着一个批量归一化层 (Batch Normalization) 和 ReLU 激活函数。这种块设计用于处理较低维度的特征图, 并且在两个卷积层中, 通常使用相同数量的输出通道。ResNet-50 使用 BottleNeck 残差模块, 构建更深网络模型。每个 BottleNeck 包含 3 个卷积层: 一个 1×1 卷积用于降维, 一个 3×3 卷积用于处理降维后的特征, 另一个 1×1 卷积用于扩展维度, 有助于减少计算复杂性, 同时保持网络深度和性能。这两种类型的模块都在它们各自的残差单元中使用跳跃连接 (或残差连接), 这些连接可以将输入直接添加到卷积层输出之后, 有助于解决更深网络中的梯度消失问题。通过使用这些模块, 残差网络能够在增加深度的同时保持训练的稳定性 and 效率。

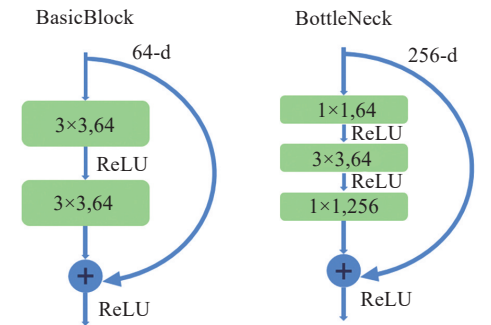


图3 BasicBlock 和 BottleNeck 残差模块
Figure 3 BasicBlock and BottleNeck residual modules

YOLOv5 是当下最先进的目标检测算法之一, 因其高速检测能力、轻量级模型结构、快速部署以及出色的检测精度而广受欢迎^[19]。YOLOv5 的结构是一个典型的 one-stage 目标检测算法框架, 其网络架构分为: 输入端 (Input)、主干网络 (Backbone)、颈部模块 (Neck)、头部模块 (Head) 及输出端 (Output), 如图 4 所示。这种设计将目标检测问题直接转化为回归问题, 省略了传统的检测框生成步骤, 从而显著提高了处理速度。由于其简洁的网络结构, 整个模型可以实现端到端的训练, 无需在各阶段之间调整参数以优化性能, 使得训练过程更加高效和直接。在此次的舌象分类任务中适当调整代码框架, 训练、测试及预测的代码入口位于新建的 classify 文件夹, 这些文件的架构与目标检测任务保持一致。

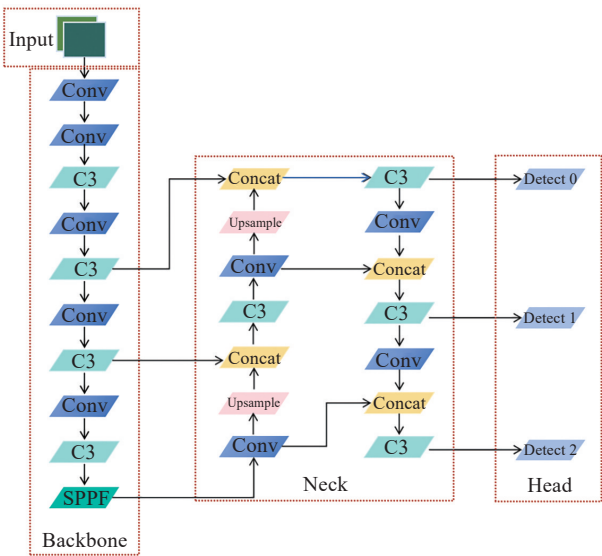


图4 YOLOv5 模型结构
Figure 4 Structure of YOLOv5 model

2.3 模型参数设计

将分割好的舌体数据集按照 8:2 随机划分为训练集和测试集。为适应训练需求与配置环境, 模型参数设计为: 训练批次设置为 16, 训练轮次设置为 36 轮, 学习率设置为 0.000 1, 模型优化器使用 SGD。

2.4 模型评价指标

本研究采用混淆矩阵 (Confusion Matrix) 并计算 F1 值 (F1 Score)、准确率 (Accuracy)、精准率 (Precision)、召回率 (Recall) 等展现模型性能。混淆矩阵以矩阵的形式展示实际类别与模型预测类别之间的关系, 帮助理解某种类别被识别的具体情况, 其中, TP (True Positive) 表示实际为正类且被预测为正类的样本数量; FN (False Negative) 表示实际为正类但被预测为负类的样本数量; FP (False Positive) 表示实际为负类但被预测为正类的样本数量; TN (True Negative) 表示实际为负类且被预测为负类的样本数量^[19]。模型评价指标计算公式如下:

F1 Score = $\frac{2TP}{2TP + FP + FN}$ (1)

Accuracy = $\frac{TP + TN}{P + N}$ (2)

Precision = $\frac{TP}{TP + FP}$ (3)

Recall = $\frac{TP}{TP + FN}$ (4)

3 结果

3.1 ResNet-34 分类结果

ResNet-34 模型混淆矩阵见图 5, F1 值为 91.46%, 准确率为 92.87%, 精准率为 90.48%, 召回率为 92.89%。训练集和验证集的损失函数图如图 6 所示, 准确率图如图 7 所示。

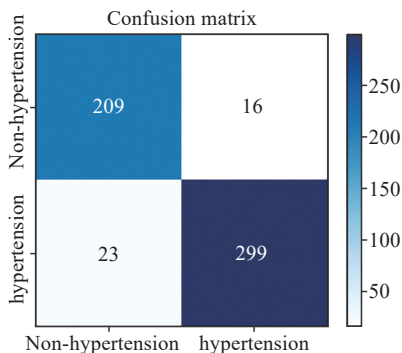


图5 ResNet-34模型混淆矩阵

Figure 5 Confusion matrix of ResNet-34 model

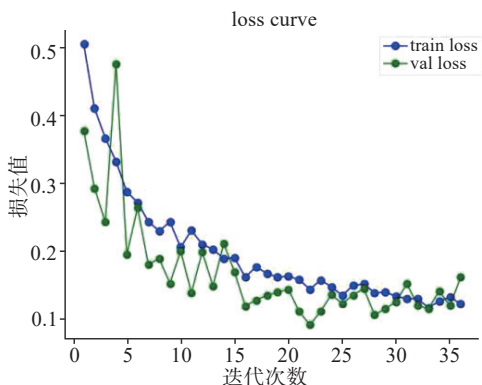


图6 ResNet-34模型损失函数图

Figure 6 Loss function of ResNet-34 model

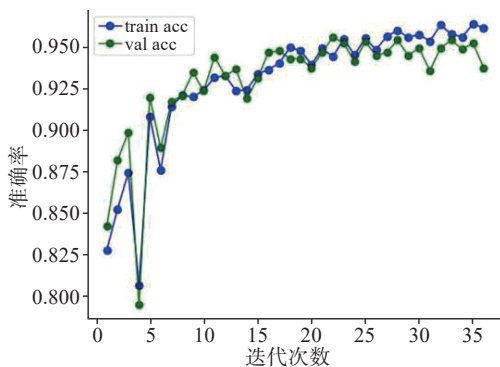


图7 ResNet-34模型准确率图

Figure 7 Accuracy of ResNet-34 model

3.2 ResNet-50分类结果

ResNet-50模型混淆矩阵见图8,F1值为92.08%,准确率为93.05%,精准率为95.26%,召回率为89.11%。训练集和验证集的损失函数图如图9所示,准确率图如图10所示。

YOLOv5模型展现出稳定的分类性能,整体准确率达到85.6%,其中痰湿体质高血压患者和痰湿体质亚健康者分别取得85.3%和85.7%的准确率,显示出两类样本分类效果的一致性。

由以上指标可以看出,ResNet-34、ResNet-50、YOLOv5在本次分类任务中均表现良好。这些模型的性能指标共同表明模型具有高效的学习能力和良好的泛化性,能够在保证准确性的同时捕捉大多数的正类样本。由图6和图9可见在训练过程中损失逐渐减小,反映了模型在逐步学习和优化。在图7和图10中训练集和验证集准确率在训练总体变化趋势较为接近,这表明模型在本次任务中泛化能力强,避免了过拟合。

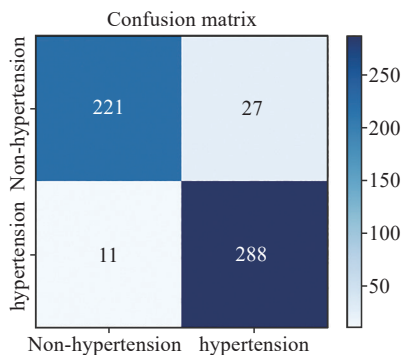


图8 ResNet-50模型混淆矩阵

Figure 8 Confusion matrix of ResNet-50 model

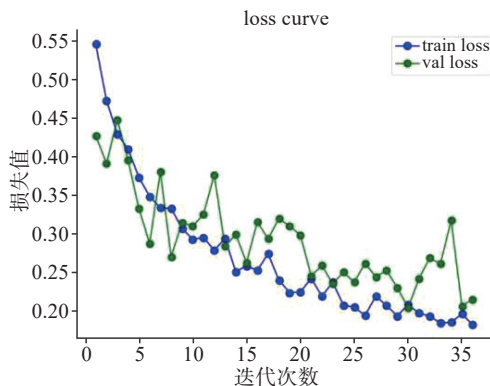


图9 ResNet-50模型损失函数图

Figure 9 Loss function of ResNet-50 model

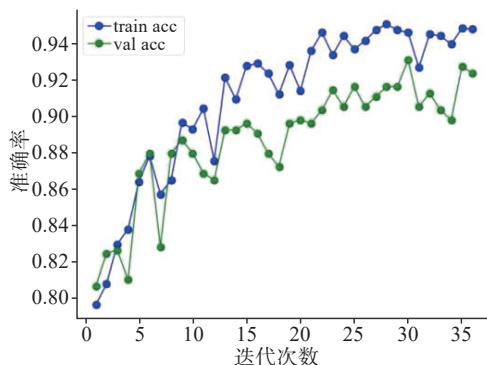


图10 ResNet-50模型准确率图

Figure 10 Accuracy of ResNet-50 model

ResNet-50 在此次分类任务表现最佳,其较高的准确率(93.05%)和精确率(95.26%)表明该模型在分类任务中对正样本的识别能力强,尽管召回率稍低为 89.11%,说明模型在识别负样本时可能存在一些漏检,但整体性能依然优于其他两个模型。ResNet-34 的性能也很优秀,其召回率为 92.89%,高于 ResNet-50 的召回率,说明在某些情况下,它能够更好地识别正类样本。这种均衡性使得 ResNet-34 在实际应用中可能更具灵活性,特别是在对误分类影响较小的任务中。F1 值是综合考虑精确率与召回率的重要指标,ResNet-50 的 F1 值为 92.08%,高于 ResNet-34,这显示了在精确率和召回率之间的整体平衡性能。对于需要权衡这两个指标的任务,ResNet-50 可能是更合适的选择。YOLOv5 的准确率明显低于 ResNet 系列模型,仅为 85.6%,这可能表明在特定的图像分类任务中,YOLOv5 表现不如 ResNet 系列模型,尽管 YOLOv5 在目标检测方面有其独特的优势。

控制在相同的硬件及数据集条件下,此次的图像分类任务中,3 种模型的训练时间和计算资源消耗也存在区别。ResNet-34 是训练时间最短且计算资源消耗最少的模型,ResNet-50 提供了更高的分类精度,因此相比于 ResNet-34 需要更多的计算资源和更长的训练时间。YOLOv5 在此次分类任务中的计算资

源消耗更高且训练时间最长,因其是目标检测领域的先进模型,结构更加复杂,在仅进行分类时仍然会消耗额外的计算资源。因此在选择模型时,可以考虑到具体的应用场景和需求,以此权衡训练时间和计算资源消耗。如果追求更快的训练速度和更少的计算资源消耗,ResNet-34 可能是一个更好的选择,追求更高的分类准确率则 ResNet-50 可能更合适。而如果总体任务涉及到目标检测,YOLOv5 可能是一个更好的选择。

4 ResNet-50 可视化

使用 Grad-CAM (Gradient-weighted Class Activation Mapping) 可视化 ResNet-50 模型中痰湿体质高血压患者与痰湿体质亚健康者舌象,该方法可生成热力图以展示模型在分类任务时较为关键的特征区域^[20]。如图 11 所示,红色区域表示模型在做出决策时重点关注的图像部分,即该区域对于模型决策的贡献度较大,周围的蓝、绿色区域相对较小。痰湿体质高血压患者舌象的热点区域多集中在裂纹与黄腻苔部位,这与高血压患者常有血行不畅、体内热邪较盛、津液耗伤的情况较为符合。因此模型分类与临床诊断经验接近,体现了显著的辅助诊断潜力和重要的临床应用价值。

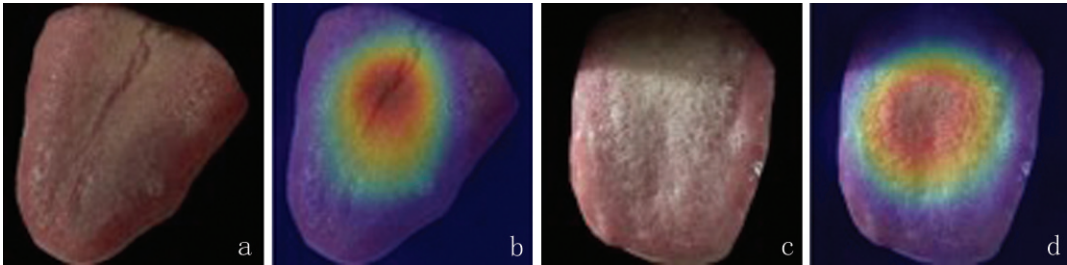


图 11 Grad-CAM 热力图
Figure 11 Grad-CAM thermal maps
a、b:痰湿体质高血压患者舌象及其热力图;c、d:痰湿体质亚健康者舌象及其热力图

5 讨论与展望

痰湿体质作为高血压的易感体质之一,其舌象变化往往与高血压的病理过程相呼应,通过观察舌的形态、色泽、舌苔及其变化,可以初步判断痰湿体质者是否患有高血压,从而实现对病症的早发现、早诊断以及方便随时查看治疗效果^[21-22]。

随着信息技术的飞速进步与计算能力的显著提升,深度学习技术正逐步渗透到中医舌诊领域,特别是在舌象分类这一关键环节,展现出了巨大的潜力。例如张景慧等^[23]通过从特征融合和决策融合两个方面实现信息融合,构建了胃肠疾病的舌诊模型;王朝雨等^[24]通过迁移学习微调改进 DenseNet201 提取特

征向量,选用 4 种特征提取算法决定输入分类器的特征向量,分类结果在脑卒中中医舌象辨证分型中达到 95.74% 的准确率;宁智敏等^[25]对比分析 AlexNet、GoogLeNet、VGG16 和 ResNet50 这 4 种 CNN 模型对虚证舌象识别的效果,认为 ResNet-50 模型对于虚证舌象的识别相较于另外 3 种更灵敏,可以辅助中医医师诊断病人是否为虚证体质。深度学习以其强大的特征学习能力、复杂的非线性映射能力以及大规模数据处理能力,为中医舌诊的客观化、量化及自动化提供了强有力的技术支持^[26]。通过大量舌象数据集的训练与优化,深度学习模型能够精准地捕捉舌象的细微差异,减少了传统舌诊中因医师个人经验

差异等因素导致的判断模糊性与主观性,提升了诊断的准确性^[27]。

使用分类模型对舌体图像进行分类,通过自动识别舌体图像中的异常或特定区域,可以提升医生的工作效率和诊断准确性。分类模型可以快速、准确地将图像归类为健康或异常类别,帮助诊断高血压疾病。这种技术能够处理大量的舌体图像数据,减轻医生的工作负担,减少人为错误,同时为远程医疗和自动化诊断提供支持,有助于实现更加精准的医疗决策^[28-29]。本研究结果验证了深度残差网络 ResNet-34 与 ResNet-50 在图像分类任务中的卓越表现。ResNet-50 在绝大多数指标上超过了 ResNet-34,显示出更好的整体性能,能够提供高准确率和良好的精确率,适合需要高精度的应用场景。ResNet-34 则提供了更均衡的性能,适合对召回率有更高要求的任务。从模型结构上来看,ResNet-50 相较于 ResNet-34 拥有更多的网络层次和参数数量,构建了一个更为深邃的架构,这通常意味着更强的特征提取能力和更高的计算复杂度,在本次任务中,这种复杂度的增加带来了性能上的显著提升。YOLOv5 的表现虽然在总体准确率上略低于 ResNet,但它是为复杂的目标检测任务设计的,因此在单一的分类任务上可能不是最优选择^[30]。

本研究通过深度学习模型初步探索了深度学习技术在中医舌诊自动化分类中的应用潜力,并在分类任务中取得了较好的结果。但受限于样本数量、模型设计细节及数据多样性等因素,当前的研究成果尚属初步阶段。为了进一步验证并深化这一领域的探索,未来研究可通过扩大样本量、优化模型结构、多模态数据融合、在更大规模的临床环境中训练和验证深度学习模型,提高舌象辨识的效率与准确性。深度学习技术的引入,为中医舌诊的现代化转型提供了强有力的技术支持。它不仅促进了舌诊的客观化、定量化发展,还为实现中医诊疗的智能化、精准化奠定了坚实基础。未来,随着技术的不断进步与应用的深入拓展,深度学习在中医舌诊领域的应用前景将更加广阔,有望为中医学的传承与创新发展注入新的活力。

【参考文献】

- [1] 赵连友,李妍,牛晓琳.对我国高血压防控策略的思考[J].中华高血压杂志,2024,32(1):2-5.
Zhao LY, Li Y, Niu XL. Reflections on Hypertension Prevention and Control Strategies in China[J]. Chinese Journal of Hypertension, 2024, 32(1): 2-5.
- [2] 马丽媛,王增武,樊静,等.《中国心血管健康与疾病报告2022》要点解读[J].中国全科医学,2023,26(32):3975-3994.
Ma LY, Wang ZW, Fan J, et al. Interpretation of report on cardiovascular health and diseases in China 2022[J]. Chinese

- Journal of General Practice, 2023, 26(32): 3975-3994.
- [3] 袁亚樱,韩景波.高血压痰湿壅盛型中医诊治研究进展[J].国医论坛,2024,39(4):71-74.
Yuan YY, Han JB. Research advance in hypertension due to abundant phlegm-dampness with TCM diagnosis and treatment[J]. Forum on Traditional Chinese Medicine, 2024, 39(4): 71-74.
- [4] 曹炳章.辨舌指南[M].南京:江苏人民出版社,1962.
Cao BZ. A Guide to Tongue Diagnosis[M]. Nanjing: Jiangsu People's Publishing House, 1962.
- [5] Haug CJ, Drazen JM. Artificial intelligence and machine learning in clinical medicine, 2023[J]. N Engl J Med, 2023, 388(13): 1201-1208.
- [6] 吴欣,徐红,林卓胜,等.深度学习在舌象分类中的研究综述[J].计算机科学与探索,2023,17(2):303-323.
Wu X, Xu H, Lin ZS, et al. Review of deep learning in classification of tongue image[J]. Journal of Frontiers of Computer Science and Technology, 2023, 17(2): 303-323.
- [7] 刘梦,王曦廷,周璐,等.基于深度学习与迁移学习的中医舌象提取识别研究[J].中医杂志,2019,60(10):835-840.
Liu M, Wang XT, Zhou L, et al. Study on extraction and recognition of traditional Chinese medicine tongue manifestation: based on deep learning and migration learning[J]. Journal of Traditional Chinese Medicine, 2019, 60(10): 835-840.
- [8] Li JW, Zhang ZD, Zhu XL, et al. Automatic classification framework of tongue feature based on convolutional neural networks[J]. Micromachines (Basel), 2022, 13(4): 501.
- [9] 周涛,霍兵强,陆惠玲,等.残差神经网络及其在医学图像处理中的应用研究[J].电子学报,2020,48(7):1436-1447.
Zhou T, Huo BQ, Lu HL, et al. Research on residual neural network and its application on medical image processing[J]. Acta Electronica Sinica, 2020, 48(7): 1436-1447.
- [10] Redmon J, Divvala S, Girshick R, et al. You only look once: unified, real-time object detection[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2016: 779-788.
- [11] Russell BC, Torralba A, Murphy KP, et al. LabelMe: a database and web-based tool for image annotation[J]. Int J Comput Vis, 2008, 77(1): 157-173.
- [12] Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation[C]//Medical Image Computing and Computer-Assisted Intervention-MICCAI 2015. Cham: Springer International Publishing, 2015: 234-241.
- [13] Siddique N, Paheding S, Elkin CP, et al. U-Net and its variants for medical image segmentation: a review of theory and applications[J]. IEEE Access, 2021, 9: 82031-82057.
- [14] 徐光宪,冯春,马飞.基于UNet的医学图像分割综述[J].计算机科学与探索,2023,17(8):1776-1792.
Xu GX, Feng C, Ma F. Review of medical image segmentation based on UNet[J]. Journal of Frontiers of Computer Science and Technology, 2023, 17(8): 1776-1792.
- [15] 江智泉,周作建,鲍剑洋,等.改进U-Net网络的中医舌图像分割方法研究[J].软件导刊,2021,20(11):186-190.
Jiang ZQ, Zhou ZJ, Bao JY, et al. Research on the segmentation method of TCM tongue images based on improved U-NET network[J]. Software Guide, 2021, 20(11): 186-190.
- [16] Shorten C, Khoshgoftaar TM. A survey on image data augmentation for deep learning[J]. J Big Data, 2019, 6(1): 60.
- [17] He KM, Zhang XY, Ren SQ, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2016: 770-778.
- [18] 王晓兵,刘琳,王俊卿,等.基于卷积神经网络ResNet50残差网络的岩石图像岩性识别研究[J].岩土工程技术,2024,38(3):294-302.
Wang XB, Liu L, Wang JQ, et al. Lithology recognition of rock image based on convolutional neural ResNet50 residual network[J]. Geotechnical Engineering Technique, 2024, 38(3): 294-302.
- [19] 刘博文.基于图像分割与检测的舌体特征识别研究[D].成都:电子科技大学,2023.
Liu BW. Research on tongue feature recognition based on image segmentation and detection[D]. Chengdu: University of Electronic Science and Technology of China, 2023.
- [20] Selvaraju RR, Cogswell M, Das A, et al. Grad-CAM: visual

explanations from deep networks via gradient-based localization [C]//2017 IEEE International Conference on Computer Vision (ICCV). Piscataway, NJ, USA: IEEE, 2017: 618-626.

[21] 徐思雨, 周思敏, 林建国, 等. 从痰湿论治临界性高血压合并代谢紊乱[J]. 中医药导报, 2024, 30(9): 102-106.

Xu SY, Zhou SM, Lin JG, et al. Study on the treatment of borderline hypertension with metabolic disorders based on phlegm dampness theory[J]. Guiding Journal of Traditional Chinese Medicine and Pharmacology, 2024, 30(9): 102-106.

[22] 胡润萍, 张振宇, 朱玲逸, 等. 中医体质辨识社区高血压病进展[J]. 中国中医药现代远程教育, 2024, 22(17): 87-90.

Hu RP, Zhang ZY, Zhu LY, et al. Progress on traditional Chinese medicine constitution identification of hypertension in community [J]. Chinese Medicine Modern Distance Education of China, 2024, 22(17): 87-90.

[23] 张景慧, 王娟, 赵玉洁, 等. 基于机器学习的胃肠道疾病舌诊模型构建[J]. 山东大学学报(医学版), 2024, 62(1): 38-47.

Zhang JH, Wang J, Zhao YJ, et al. Construction of a machine learning-based tongue diagnosis model for gastrointestinal diseases [J]. Journal of Shandong University (Health Sciences), 2024, 62 (1): 38-47.

[24] 王朝雨, 黄奎麟, 代国威, 等. 基于卷积神经网络的脑卒中中医辨证分型舌象分类研究[J]. 湖南中医药大学学报, 2023, 43(8): 1460-1467.

Wang ZY, Huang KL, Dai GW, et al. Convolutional neural networks-based TCM pattern differentiation and classification of tongue manifestations in stroke patients[J]. Journal of Hunan University of Chinese Medicine, 2023, 43(8): 1460-1467.

[25] 宁智敏, 董志学, 吴方. 基于卷积神经网络的虚症人群舌象识别研究[J]. 信息技术与信息化, 2023(5): 195-198.

Ning ZM, Dong ZX, Wu F. Research on tongue image recognition in virtual population based on convolutional neural network[J]. Information Technology and Informatization, 2023(5): 195-198.

[26] Xu JT, Jiang T, Liu S. Research status and prospect of tongue image diagnosis analysis based on machine learning[J]. Digital Chin Med, 2024, 7(1): 3-12.

[27] 董易杭, 王建勋, 王晶, 等. 基于深度学习的阳虚质与阴虚质舌象分类研究[J]. 中华中医药学刊, 2024, 42(7): 27-30.

Dong YH, Wang JX, Wang J, et al. Study of tongue image classification for yang deficiency and yin deficiency constitutions based on deep learning[J]. Chinese Archives of Traditional Chinese Medicine, 2024, 42(7): 27-30.

[28] 肖雯迪, 朱龙, 王洋, 等. 中医望诊多模态融合技术的应用与思考[J]. 中医杂志, 2024, 65(17): 1741-1746.

Xiao WD, Zhu L, Wang Y, et al. Application and thinking on multimodal fusion technology of traditional Chinese medicine inspection[J]. Journal of Traditional Chinese Medicine, 2024, 65 (17): 1741-1746.

[29] Zhuang QB, Gan SZ, Zhang LY. Human-computer interaction based health diagnostics using ResNet34 for tongue image classification[J]. Comput Methods Programs Biomed, 2022, 226: 107096.

[30] 梁海翔, 唐艳慧, 王宇庆, 等. 高分辨率遥感图像的目标检测[J]. 液晶与显示, 2024, 39(10): 1350-1360.

Liang HX, Tang YH, Wang YQ, et al. Object detection methods for high resolution remote sensing images[J]. Chinese Journal of Liquid Crystals and Displays, 2024, 39(10): 1350-1360.

(编辑: 薛泽玲)