

本文引用: 李秋华, 史国峰, 李玥博, 任路. 基于卷积神经网络的“舌边白涎”舌象识别研究[J]. 湖南中医药大学学报, 2024, 44(7): 1254-1260.

基于卷积神经网络的“舌边白涎”舌象识别研究

李秋华^{1,2}, 史国峰², 李玥博¹, 任路^{1*}

1. 辽宁中医药大学, 辽宁 沈阳 110847; 2. 辽宁中医药大学附属第二医院, 辽宁 沈阳 110034

[摘要] **目的** 通过机器学习分析“舌边白涎”舌象特性, 对舌象进行局部特征识别研究, 探讨卷积神经网络算法在舌象识别应用中的性能。**方法** 使用 Python 进行图像预处理, 搭建用于舌象识别的视觉几何组 16 层 (visual geometry group 16, VGG16) 卷积神经网络模型, 分析其对“舌边白涎”舌象鉴别分析的效果, 并结合热力图分析“舌边白涎”典型舌象表现。**结果** 基于 PyTorch 框架, 进行卷积神经网络的舌象鉴别研究, VGG16 及残差网络 50 层 (residual network 50, ResNet50) 模型验证准确率均较高, 达到 80% 以上, 且 ResNet50 模型优于 VGG16 模型, 可为舌象识别提供一定参考。基于加权梯度类激活映射 (gradient-weighted class activation mapping, Grad-CAM) 技术, 通过舌苔舌色差异分布的网络可视化, 有助于直观进行模型评估分析。**结论** 基于卷积神经网络模型对舌象数据库进行分析, 实现“舌边白涎”舌象识别, 有助于临床诊疗的客观化辅助分析, 为舌诊智能化发展提供一定借鉴。

[关键词] 卷积神经网络; 视觉几何组; Python; 人工智能; 舌边白涎

[中图分类号] R241

[文献标志码] A

[文章编号] doi:10.3969/j.issn.1674-070X.2024.07.016

Tongue image recognition of "white saliva on the tongue side" based on convolutional neural network

LI Qiuhua^{1,2}, SHI Guofeng², LI Yuebo¹, REN Lu^{1*}

1. Liaoning University of Chinese Medicine, Shenyang, Liaoning 110847, China; 2. The Second Hospital of Liaoning University of Chinese Medicine, Shenyang, Liaoning 110034, China

[Abstract] **Objective** To study the local feature recognition of tongue image through machine-learning analysis on the characteristics of "white saliva on the tongue side", and to explore the performance of convolutional neural network algorithms in identifying tongue images application. **Methods** Python was used for image preprocessing, and visual geometry group 16 (VGG16) convolutional neural network model was built for tongue image recognition. The effect of the model on tongue image recognition of "white saliva on the tongue side" was identified and analyzed, and the typical tongue image performance of "white saliva on the tongue side" was analyzed combined with heat map. **Results** Based on PyTorch framework, tongue image identification research of convolutional neural network was carried out. The verification accuracy of VGG16 and residual network 50 (ResNet50) models were high, reaching over 80%, and the ResNet50 model outperformed the VGG16 model, providing a certain reference for tongue image recognition. Based on gradient-weighted class activation mapping (Grad-CAM) technology, the network visualization of the difference distribution of tongue coating and tongue color was helpful for intuitive model evaluation and analysis. **Conclusion** The analysis of tongue image database based on convolutional neural network model can realize the tongue image recognition of "white saliva on

[收稿日期] 2023-12-05

[基金项目] 中国博士后科学基金项目 (2021MD703842); 国家中医药管理局中医药国际合作专项项目 (0610-2140NF020629); 辽宁省自然科学基金计划重点项目 (20180540043)。

[通信作者] * 任路, 女, 博士, 二级教授, 博士研究生导师, E-mail: lnzyxkc@sina.com。

the tongue side", which is helpful for objectified auxiliary analysis in clinical diagnosis and treatment, and provides some references for the intelligent development of tongue diagnosis.

[**Keywords**] convolutional neural network; visual geometry group; Python; artificial intelligence; white saliva on the tongue side

望舌主要包括望舌质的神、色、形、态及舌苔的颜色、厚薄与润燥等。但相关望舌苔的研究中,未单独论及“舌边白涎”现象,这类舌象主要是被归属于“舌苔黏腻”范畴。多数研究者认为,舌边白涎的形成与情志刺激及情志疾病相关^[1]。“舌边白涎”由朱良春先生于20世纪80年代提出,并首见于其弟子俞淦琪^[2]的研究文献中。“舌边白涎”是指患者首次就诊时,舌象可见舌两侧边缘各现一条由白色涎沫聚凝而成的、断续或连贯的线索状泡沫带^[2-4]。

人工智能机器学习算法在图像裁剪、去噪、分类、调节、分析方面,识别效果较好,可以推进舌诊的客观化、智能化发展^[5-6]。本研究将图像处理与卷积神经网络技术相融合,选取“舌边白涎”舌象为切入点,将图像处理和卷积神经网络技术应用到中医舌象识别,实现临床诊疗的客观化辅助分析。应用Python进行图像预处理,通过PyTorch平台,基于视觉几何组16层(visual geometry group 16, VGG16)卷积神经网络模型^[7-8],搭建舌象识别卷积神经网络模型,对舌象特征进行精准识别和多维度定量分析,对比卷积神经网络模型舌象识别的效果。同时,通过加权梯度类激活映射(gradient-weighted class activation mapping, Grad-CAM)方法^[9],可视化解析卷积神经网络模型,为舌诊的智能化发展提供一定的借鉴。

1 材料与方法

1.1 数据来源

本文舌象采集自2018年7月至2021年7月就诊于辽宁中医药大学附属第二医院的患者。采集舌象保存建档,存储为JPG格式,共采集医案(包含舌象)合计405例,由3名副主任中医师以上职称人员对舌象进行判定记录,同时结合Grad-CAM方法,通过热力图进行筛选,最终结合主客观双重判断,构建“舌边白涎”的舌象组(观察组)与非“舌边白涎”的舌象组(对照组)的舌象数据集。其中,观察组共计178例,对照组共计227例。舌象数据采集过程的基本要求:患者在舌象拍摄采集前1h内,不可食用有可能导致舌苔染色的食物、饮料或者药物,因舌体收缩、颤动、卷曲等影响舌象照片清晰的情况不予纳

入;拍摄舌象时,需要将患者上下嘴唇及舌体整体均摄入,女性患者需要在拍摄前清除口红等,在拍摄舌象图片时,避免光线过强或过暗。

1.2 研究方法

1.2.1 舌象采集分割方法 为保证良好的特征识别,保留患者舌体的有效信息,分割有效的舌象细节,分类识别舌体,主要围绕舌质、舌苔展开,聚焦于舌质的神色和形态以及舌苔的质地和颜色,将机器学习应用于图像识别领域,进而提高识别准确度、泛化能力。

1.2.2 图像预处理 标注舌部轮廓,去除干扰因素,应用Python的cv2库,将原始图像缩放成224×224×3像素,裁剪去噪处理、分割,提取舌象形状、纹理和颜色特征,经过量化、对比、色彩调节,进行模糊、调亮、调暗、水平或垂直翻转、移动、旋转(30°、60°、90°、180°、270°)、0值填充处理,经过预处理,扩充图像,建构训练集及测试集舌象。

1.2.3 模型训练 训练基于VGG16、残差网络50层(residual network 50, ResNet50)舌象特征分类模型,卷积层主要用于提取图像特征。数学表达式为:

$$H_i=f(W_i+H_{i-1}+b_i) \quad (1)$$

公式中: H_i 代表第*i*层的特征图; f 指激活函数; W_i 表示第*i*层的卷积核的权重值; H_{i-1} 表示第*i-1*层输入的特征图; b_i 表示第*i*层的偏移量。

池化层即汇聚层,用于进行特征选择,降低特征数量,从而减少参数数量。全连接层则是对提取的特征进行非线性组合,产生输出结果。分类输出时,使用Softmax函数进行分类,函数表达式为:

$$\text{Softmax}(z_i)=\frac{e^{z_i}}{\sum_{c=1}^C e^{z_c}} \quad (2)$$

公式中: z_i 为第*i*个节点输出值, e 为指数函数, e^{z_i} 为指数函数的 z_i 次幂, e^{z_c} 为指数函数的 z_c 次幂, C 为输出节点个数,即分类的类别个数,通过Softmax函数可将多分类的输出值转换为范围在[0,1]和为1的概率分布。从输入层输入图像,进入卷积层,通过过滤器进行卷积运算,运算结果是从数据中提取的特征。多次卷积、下采样,行全连接,获取特征^[10],整合后得到输出层,详见图1。

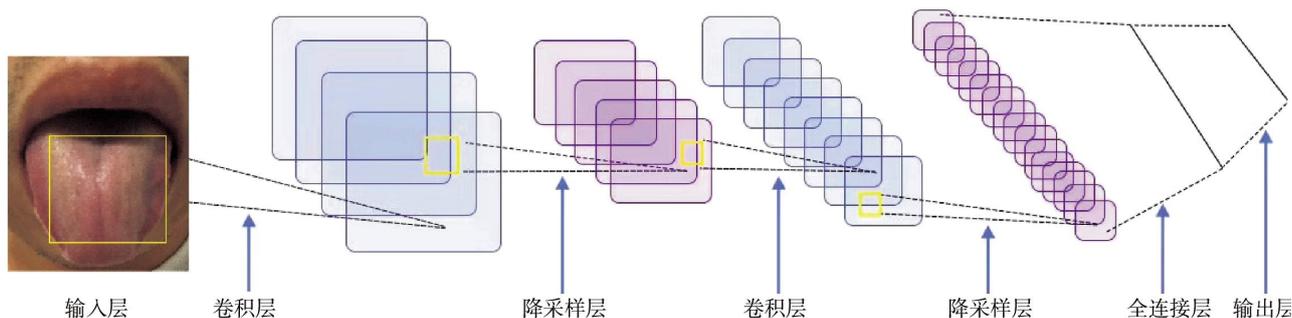


图 1 舌象卷积神经网络结构图

Fig.1 Structure diagram of tongue convolutional neural networks

对模型进行训练,将模型迭代次数作为模型最终训练结束标志,应用验证集进行验证,记录验证准确率。具体模型训练流程见图 2。

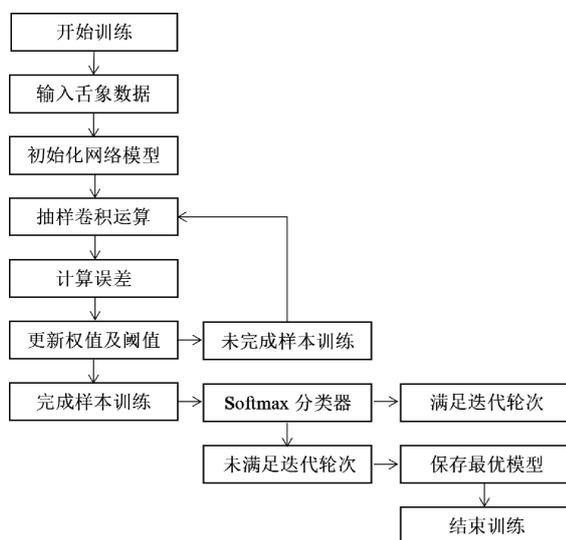


图 2 舌象特征分类模型训练流程图

Fig.2 Training process of tongue image feature classification model

1.2.4 VGG16 模型 VGG 模型是一种深度卷积神经网络^[11]。VGGNet 相当于在传统卷积神经网络模型基础上的拓展,是针对较为复杂的模型结构的设计。VGG 模型通过使用多个卷积层和池化层来提高网络的深度,并通过在每个卷积层后添加 ReLU 激活函数来提高网络的非线性能力。针对过拟合的问题,可以通过增加训练数据集的样本数量,减少模型对训练数据的过度拟合,提高模型的泛化能力,在应用训练图像数据时,可采取随机水平翻转或随机 RGB 色差的方法,来增加数据集,对经过缩放的图片进行随机排序,选择尺寸为 224×224 像素的训练图像。

(1)卷积层的计算公式为:

$$Y[i, j]=\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} X[i+m, j+n] \cdot W[m, n]+b \quad (3)$$

其中, X 为输入特征图, Y 为输出特征图, W 为卷积核, b 为偏置项, i 和 j 分别代表卷积核的第 i 行和第 j 列, m, n 为维度。

(2)最大池化层的计算公式为:

$$Y[i, j]=\max_{m, n} X[2^{i+m}, 2^{j+n}] \quad (4)$$

其中, X 为输入特征图, Y 为输出特征图, m 和 n 为维度,取 0 或 1,用于选择局部区域中的最大值。

(3)全连接层的计算公式为:

$$Y=ReLU(W \cdot X+b) \quad (5)$$

其中, $ReLU$ 为线性整流函数, Y 为输出特征图, W 为卷积核,即权重矩阵, X 为输入特征图, b 为偏置项。

VGG 网络的训练过程使用基于随机梯度下降的反向传播算法,用于优化网络参数,实现对图像分类任务的准确预测。

1.2.5 ResNet50 模型 2015 年,微软团队提出 ResNet,在获得分类任务、目标检测及图像分割方面具有积极意义^[12]。ResNet 由残差学习模块构成,在分类问题上表现出较好的性能,很多网络模型均在 ResNet 基础之上设计。在引入残差学习模块后,打破了神经网络深度限制,提高了网络的训练效率和性能。近些年,ResNet50 在 ImageNet 大规模视觉识别中体现出较好表现的主流算法。在给定的图像上选择候选区域,然后对这些区域进行特征提取,最后使用训练分类器进行分类识别^[13-14]。

(1)ResNet50 网络结构。ResNet 中,每个 conv 层后均使用批归一化,进行 Xavier 初始化。而 BN 层的加入,对于抑制过拟合具有一定作用。

(2)残差网络单元图。ResNet 网络关键是结构中的残差单元,单元中包含跨层连接,曲线可以将输入直接跨层传递,进行同等映射,其后与经过卷积操作的结果相加。

假设输入图像: x ,输出: $H(x)$,中间经卷积后,输出为 $F(x)$ 非线性函数,最终输出: $H(x)=F(x)+x$,对输出进行非线性变换,残差即“差” $F(x)$,网络转化为求残差函数 $F(x)=H(x)-x$,应用 Focal Loss 损失函数,调整难、易分类样本的权重,使模型更关注难分类样本,残差网络单元示意图详见图 3。

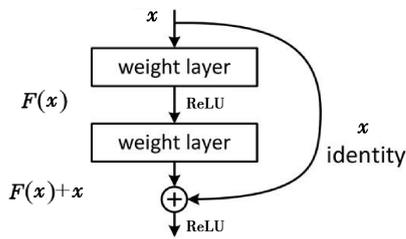


图 3 残差网络单元示意图

Fig.3 Schematic diagram of residual network units

1.2.6 Grad-CAM 本研究通过 Grad-CAM 方法,实现可视化解析卷积神经网络模型,可让模型最终卷积层中的目标形成比较粗略的定位,将其对模型贡献最高的部位以热力图形式展示^[15]。

1.2.7 环境构建及模型参数 (1)环境构建。显卡: NVIDIA GeForce GTX 1080Ti,CPU: Intel Core i7-9700K,32 GB 内存。在 Linux 系统中,基于 PyTorch-gpu 1.11.0,CUDA Toolkit 10.2 深度学习相关软件包,使用 Python 3.9 语言。(2)训练舌象分类模型设计。训练批次大小设为 64,训练样本数量为 64,每

一个批次处理完,更新一次参数。训练轮次 epochs 200 轮,学习率为 0.000 1,应用 PyTorch 深度学习框架中的优化算法 Adam 模型优化器。

2 结果与分析

基于 VGG16、ResNet50 卷积神经网络模型的舌象特征分类模型,训练、保存舌象特征分类模型实验结果,模型训练时长:VGG16 0.2 h,ResNet50 0.6 h。相关数据集的训练准确率和验证准确率情况如下:VGG16 训练准确率为 99.6%,验证准确率为 77.2%;ResNet50 训练准确率为 99.2%,验证准确率为 77.6%。详见图 4—5。

基于图 4—5 可以看出,训练准确率明显高于验证准确率,且验证集准确率曲线波动起伏较大,模型泛化能力差,出现过拟合。进一步通过图像预处理算法进行扩充,再次训练。

再次训练后,VGG16 训练准确率为 99.8%,验证准确率为 81.4%;ResNet50 训练准确率为 99.3%,验证准确率为 82.8%。验证集准确率曲线的波动幅度较小,表明模型识别效果较好,能较好地识别“舌边白涎”舌象的特征,且 ResNet50 模型优于 VGG 模型。详见图 6—7。

与 VGG16 模型的结果相比,ResNet50 模型曲线收敛较好,说明模型可以很好地学习到“舌边白涎”舌象的特征。详见图 8—9。

通过 Grad-CAM 可视化分析,将“舌边白涎”舌象进行热力图展示,在模型最终卷积使用梯度加权类激活映射分析,捕获高级别的图像语义及空间信

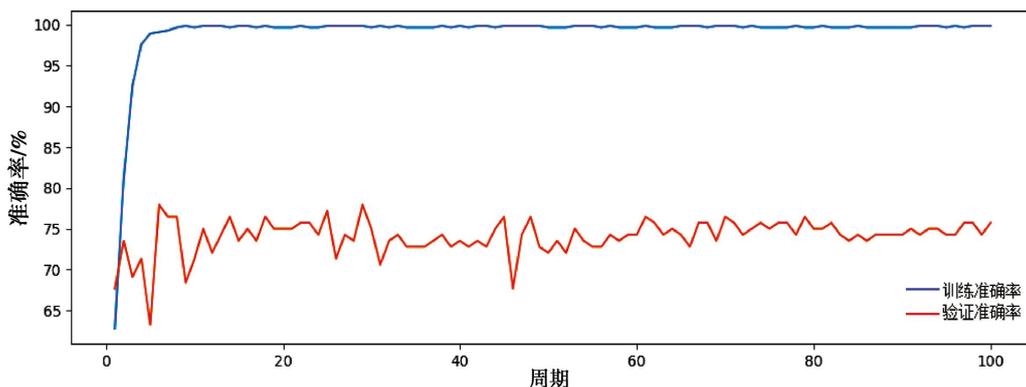


图 4 VGG16 模型的训练准确率和验证准确率

Fig.4 Training accuracy and validation accuracy of VGG16 model

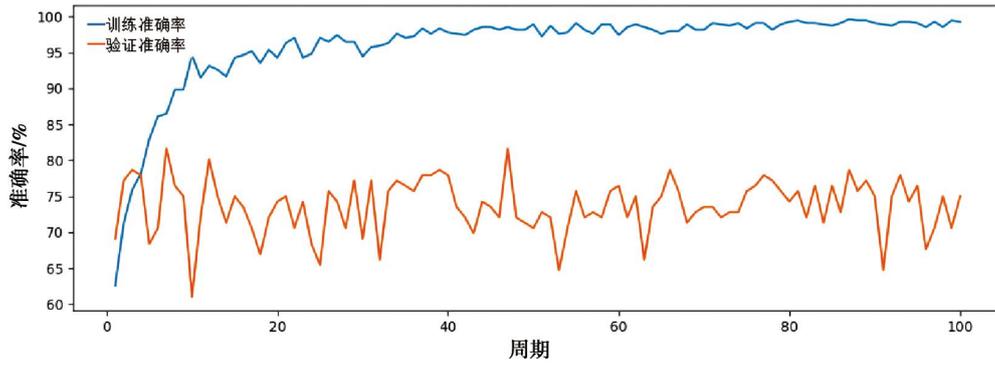


图 5 ResNet50 模型的训练准确率和验证准确率

Fig.5 Training accuracy and validation accuracy of ResNet50 model

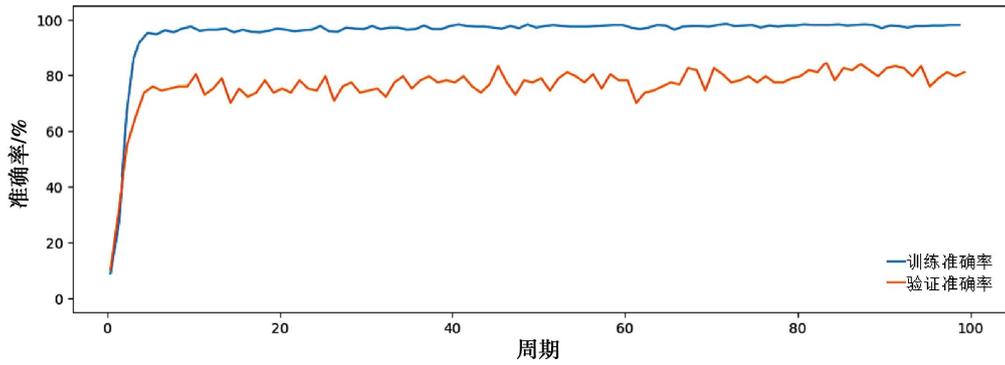


图 6 VGG16 扩充模型的训练准确率和验证准确率

Fig.6 Training accuracy and validation accuracy of the VGG16 extended model

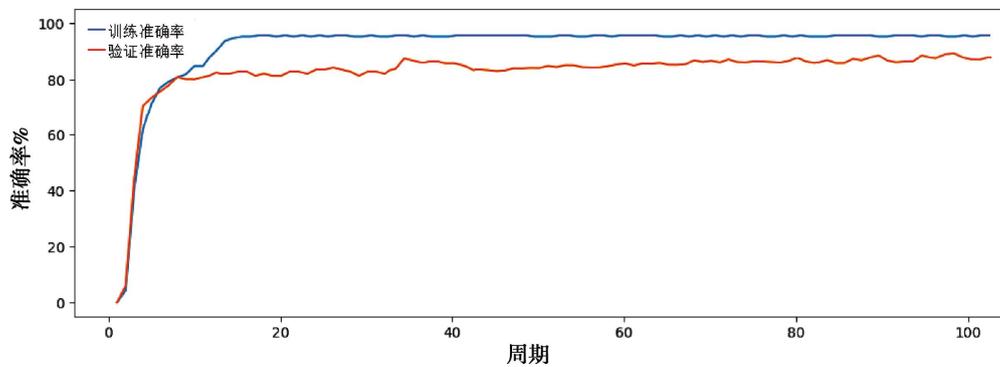


图 7 ResNet50 扩充模型的训练准确率和验证准确率

Fig.7 Training accuracy and validation accuracy of ResNet50 extended model

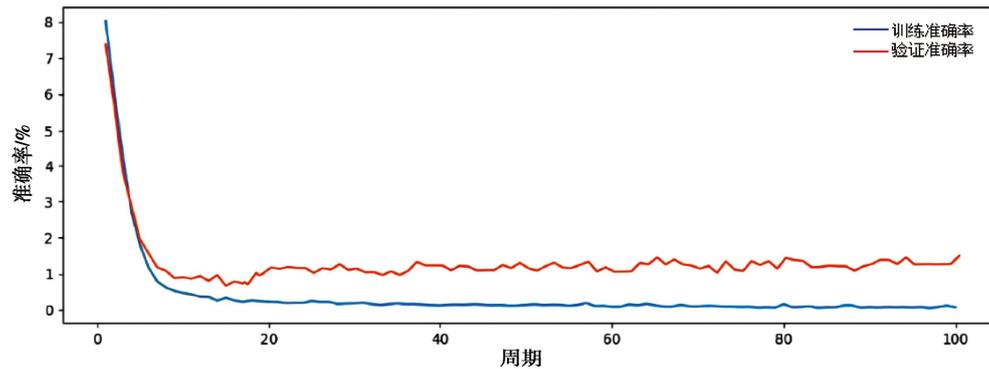


图 8 VGG16 模型 loss 曲线

Fig.8 VGG16 model loss curve

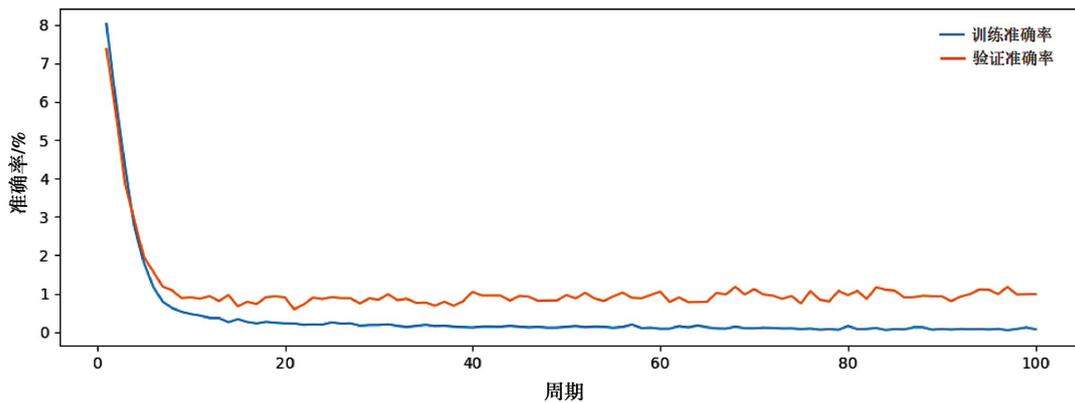


图9 ResNet50 模型 loss 曲线

Fig.9 ResNet50 model loss curve

息,进行典型“舌边白涎”舌象的热力图展示,直观展示出不同苔质的局部区域差异性。应用 Grad-CAM进行深度学习网络可视化分析,有助于模型的评估,辨析舌象的类别差异性分区,提取出“舌边白涎”舌象两侧白涎线这一重要特征,提供给研究者更加直观的信息以辅助诊断。详见图 10。



图 10 观察组部分抽样患者舌象热力图

Fig.10 Heat map of tongue images of partially sampled patients in the observation group

注:蓝色表示舌面的基础情况,即淡红舌颜色的体现;红色表示舌面白涎,即需要识别的“舌边白涎”舌象的典型色彩区域。

3 讨论

PyTorch 是一个拥有自动求导功能的强大的深度神经网络^[16]。近几年来,相关医工学科交叉融合、相关技术与医学的融汇越来越密切^[17]。人工智能技术在临床医学诊疗活动中的应用日益广泛^[18]。目前,中医学与人工智能主要应用成果包括舌诊、脉诊设备以及诊疗系统等,舌象是中医的主要诊察手段,也是研究的热点之一,基于各种算法以及相关理论的舌象研究在持续更新换代。

舌诊属于中医辨证论治的重要依据,关于舌象的脏腑分区与经络的关系,在《灵枢·经脉》中有“厥

阴者肝脉也……而脉络于舌本也”的记载,梁玉瑜在《舌鉴辨证·全舌分经图》中指出“舌边分主五脏六腑,舌边左主肝,右主胆”,吴坤安在《伤寒指掌·察舌辨证法》中提出“满舌属胃……两旁属肝胆”,沈月光的《伤寒第一书·舌上三图总论》中也指出“舌之尖属心经……两旁属肝胆”,将舌象脏腑做了明确分区。从脏腑辨证来看,舌面两侧在从属上是肝胆所主。舌诊中少阳经所主部位在舌两侧,两侧见白涎属肝胆疏泄失常,少阳经气运行受阻。朱良春教授提出“舌边白涎”,即舌之两侧边缘各有一条由泡沫聚凝而成的线索状涎沫^[19]。有学者将舌之两侧泡沫状白涎排列线称为“舌纓线”,并且指出这种线常对称分布^[20]。“舌边白涎”是一种病理信号,是肝郁气滞、脾失健运之征,朱良春教授常以豁痰渗湿、调气开郁的方法论治本病。临床上,常见情志不遂之人舌边有白涎,津液输布障碍,循经脉上泛,舌边两侧对应肝胆,故见“舌边白涎”。有学者认为,“舌上白涎”反映机体存在复杂的病理状态^[21],非独小柴胡汤主之,柴胡类方皆主之^[22]。可见“舌边白涎”舌象的分析研究具有一定临床探索价值。有学者称之为“液线”^[23]。王彦晖将其称为“肝郁线”,认为这种舌象与肝胆气滞具有相关性^[24]。也有学者指出,“舌边白涎”属于少阳证,主以小柴胡汤治疗,故命名为“小柴胡线”^[21]。

本研究中的训练基于 VGG16、ResNet50 卷积神经网络模型的舌象特征分类模型,其中 VGG 可以看作对 Alexnet 的改进,整个网络由卷积层和全连接层叠加而成,增大了网络的非线性能力。ResNet 引入残差学习和深度可分离卷积,解决了神经网络中的梯度消失和信息丢失问题,为深度加深打下

基础。本研究发现,ResNet50 模型性能相对较好,具有较为优秀的识别效果。结合热力图可视化分析,对不同舌象图像区域的关注程度差异度进行区别,确定舌边特定区域对辨识的贡献,判断模型的舌象特征,有助于对模型进行评估。

深度学习技术与医学相结合,可以为医学诊疗提供新的技术手段,通过自动分析舌象图像,完成舌诊的客观化研究,减少主观判断诊断的偏差,保障辨识结果的可靠性^[25]。本研究团队下一步将整理大样本的临证医案,包含症状、体征等信息,将文本识别与图像识别联合建模,探索“舌边白涎”舌象与抑郁及焦虑等评分的相关性,以及与相关主要症状的内关联。

参考文献

- [1] 邱义勇,熊永法.情志病研究概述[J].江西中医药,2023,54(2):78-80.
- [2] 俞淦琪,朱良春“舌边白涎”诊法的经验[J].北京中医,1988,7(6):8-9.
- [3] 陈雪莲,吴星星,刘可心,等.舌涎白涎[J].江苏中医药,2021,53(2):72-73.
- [4] 张佳元,王天芳,杨浩婕.“舌边白涎”形成机理的初步探讨[J].现代中医临床,2015,22(6):42-43.
- [5] 余楷杰,袁芳君,马庆宇,等.机器学习驱动中医诊断智能化的发展现状、问题及解决路径[J].中国中医基础医学杂志,2024,30(3):398-406.
- [6] 张晓维,张晓雨,赵晨,等.循证视角下数智中医广义望诊研究及相关装备研制思路[J/OL].中国实验方剂学杂志,1-8[2024-06-07].<https://doi.org/10.13422/j.cnki.syfjx.20240927>.
- [7] 沈筱梅,张新颖,王权泳,等.基于卷积神经网络的涎腺肿瘤超声图像分类研究[J].临床超声医学杂志,2023,25(10):849-855.
- [8] 廖清江,刘婷,张星月,等.基于VGGNet的人脸表情识别研究[J].软件工程,2023,26(11):59-62.
- [9] 张明凯,胡军国,刘江南,等.基于深度学习的可视化图表分类方法研究[J].电子技术应用,2024,50(5):58-65.
- [10] GU J X, WANG Z H, KUEN J, et al. Recent advances in convolutional neural networks[J]. Pattern Recognition, 2018, 77(C): 354-377.
- [11] 何依婷.基于孪生网络的肺部图像识别[J].电视技术,2023,47(7):24-27.
- [12] 朱家群,王东阳,顾玉宛,等.基于多尺度融合CNNs特征和Transformer特征的显著目标检测[J].常州大学学报(自然科学版),2023,35(6):35-44.
- [13] 邵定琴,张乾,岳诗琴,等.基于ResNet50网络的年龄估计[J].信息与电脑(理论版),2021,33(1):21-24.
- [14] 杜元翰,汤铭,查易艺,等.融合注意力机制与全局特征网络的行人重识别算法研究[J].测试技术学报,2023,37(1):37-42.
- [15] 曾子强.基于Grad-CAM联合Mask R-CNN的乳腺钼靶图像分类可信度分析[D].南昌:南昌大学,2023.
- [16] 陈昭明,洪锦魁.PyTorch深度学习应用实战[M].北京:清华大学出版社,2023:23.
- [17] 景城阳,戴泽琦,吴雪,等.中医药领域预测模型研究的可视化分析[J].中国实验方剂学杂志,2024,30(14):209-217.
- [18] 陈美伊.人工智能技术在医疗系统中的应用[J].电子技术,2023,52(9):313-315.
- [19] 李韬芝,贾立群,杨闪闪,等.贾立群对白涎线舌象的认识[J].中医学报,2023,38(1):83-87.
- [20] 张丽宁.原发性干燥综合征的舌象特点及其与疾病活动度的相关性研究[D].北京:北京中医药大学,2022.
- [21] 李敏,何庆勇.基于“舌上白胎”试论小柴胡证[J].中医杂志,2021,62(10):917-920.
- [22] 鲁颖晔.从“舌边白涎”浅论柴胡证[J].浙江中医杂志,2023,58(4):269-270.
- [23] 何勇,陈晓雯.舌诊“液线”分析[J].山东中医杂志,2017,36(4):335-336.
- [24] 李鹏飞,刘培,张辰龙,等.王彦晖治疗痰郁互结型甲状腺功能亢进症经验[J].中医药通报,2022,21(12):14-16.
- [25] 孙忠人,游小晴,韩其琛,等.人工智能在中医药领域的应用进展及现状思考[J].世界科学技术:中医药现代化,2021,23(6):1803-1811.

(本文编辑 周旦)