

基于机器学习的胃肠道疾病舌诊模型构建

张景慧¹,王娟²,赵玉洁³,段森¹,刘毅然⁴,林敏娟¹,譙旭⁴,李真¹,左秀丽¹

(1.山东大学齐鲁医院消化内科,山东 济南 250012; 2.青岛市卫生健康委员会医院发展中心,山东 青岛 266001; 3.山东中医药大学附属医院消化内科,山东 济南 250011; 4.山东大学控制科学与工程学院,山东 济南 250061)

摘要:目的 构建基于机器学习的胃肠道疾病舌诊模型,以寻求更加方便、经济的方式实现对常见胃肠道疾病的非侵入性诊断。方法 前瞻性收集接受电子内镜检查的948名受试者的舌象图片,经过质量筛选,最终获得符合应用标准的3140张图片构成本研究使用的舌象数据集。对原始舌象数据进行预处理、特征提取与模式识别,在传统机器学习方法的基础之上,提出一种从特征融合和决策融合两个方面实现信息融合的方法,以此构建以舌象特征为输入的胃肠道疾病舌诊模型。结果 本研究构建的基于舌象的信息融合诊断模型的曲线下面积(area under the curve, AUC)为0.808,高于单一手工特征(AUC=0.769)和深度特征(AUC=0.779)模型;使用BSFCM混合采样方法进行样本增强提高了该模型对幽门螺杆菌(*Helicobacter pylori*, *H. pylori*)感染(AUC=0.816)、胆汁反流(AUC=0.829)、反流性食管炎(AUC=0.800)、胃糜烂(AUC=0.833)和十二指肠糜烂(AUC=0.818)的分类性能。结论 本研究构建的基于机器学习的智能舌诊模型对多种胃肠道疾病具有较高的区分度,或为胃肠道疾病的诊断与筛查提供一种新的、有价值的思路与方法。

关键词:人工智能;舌象;胃肠道疾病;机器学习;舌诊模型

中图分类号:R241.25;R57

文献标志码:A

Construction of a machine learning-based tongue diagnosis model for gastrointestinal diseases

ZHANG Jinghui¹, WANG Juan², ZHAO Yujie³, DUAN Miao¹, LIU Yiran⁴,

LIN Minjuan¹, QIAO Xu⁴, LI Zhen¹, ZUO Xiuli¹

(1. Department of Gastroenterology, Qilu Hospital of Shandong University, Jinan 250012, Shandong, China;

2. Hospital Development Center of Qingdao Municipal Health Commission, Qingdao 266001, Shandong, China;

3. Department of Gastroenterology, Affiliated Hospital of Shandong University of Traditional Chinese Medicine, Jinan 250011, Shandong, China; 4. School of Control Science and Engineering, Shandong University, Jinan 250061, Shandong, China)

Abstract: Objective To construct a machine learning (ML)-based tongue diagnostic model for the diagnosis of gastrointestinal diseases so as to realize the non-invasive auxiliary diagnosis of common gastrointestinal diseases in a more convenient and faster way. **Methods** Tongue images of 948 subjects who underwent electronic endoscopy were prospectively collected. After quality screening, 3,140 images that met the application criteria were finally obtained to constitute the tongue image data set, which underwent preprocessing, feature extraction and pattern recognition. On the basis of traditional machine learning methods, a method to realize information fusion in terms of feature fusion and decision fusion was proposed, and a tongue diagnosis model of gastrointestinal diseases was constructed. **Results** The area under the curve (AUC) of the model was 0.808, which was higher than that of the single handcrafted feature (AUC=0.769) and deep feature (AUC=0.779) models. Sample enhancement using the BSFCM hybrid sampling method improved the model's performance for *Helicobacter pylori* (*H. pylori*) infection (AUC=0.816), bile reflux (AUC=0.829), reflux esophagitis (AUC=0.800), gastric erosion (AUC=0.833) and duodenal erosion (AUC=0.818).

Conclusion The intelligent tongue diagnostic model based on ML constructed in this study shows a high degree of differentiation for a variety of gastrointestinal diseases, and may provide a new and valuable idea and method for the diagnosis and screening of gastrointestinal diseases.

Key words: Artificial intelligence; Tongue image; Gastrointestinal diseases; Machine learning; Tongue diagnosis model

一项涵盖了 204 个国家和地区的研究数据表明,全球消化系统疾病在 2019 年的年龄标准化发病率为 95 582 例/10 万人年,占流行性疾病病例的 1/3 以上,造成了巨大的社会和经济负担^[1]。胆汁反流、胃糜烂、十二指肠糜烂等常见胃肠道疾病的诊断目前主要依靠电子内镜和病理组织学检查,由于成本费用高、具有侵袭性以及需要专业内镜医生等,使其应用受到极大限制。因此,开发一种方便、经济、无创的胃肠道疾病诊断方法势在必行。

祖国医学是几千年来经验应用和保存的医学和文化遗产^[2]。舌为心之苗,脾之外候,舌苔由胃气所生,脏腑通过经脉与舌相联,故脏腑病变可在舌质和舌苔上反映出来。因此,中医舌诊也构成了中医望诊的核心内容。舌的颜色、形状、质地及齿痕等变化可以反映出多种疾病的病理生理信息^[3-4]。舌作为消化道的起始部分,其与脾胃的关系极为密切。有学者发现唾液量和脾胃区舌苔量与反流性食管炎患者的严重程度之间存在关系,表明舌可作为评估和预测反流性食管炎的标志^[5]。另一项研究表明舌红或紫、苔厚是消化道症状患者感染幽门螺杆菌 (*Helicobacter pylori*, *H. pylori*) 的潜在指标,提示舌的特征可以作为伴有消化道症状患者感染 *H. pylori* 的可行预测指标^[6]。然而,中医舌诊具有较强的主观性,医生的经验和知识储备极大地限制了中医舌诊在临床中的应用,实现舌诊标准化、客观化和智能化的问题亟待解决。

近年来,人工智能技术与医学领域深度融合^[7]。基于人工智能和计算机图像处理等技术已在多个医学领域中的图像识别方面取得突破性进展^[8-12]。通过人工智能探索舌与疾病之间的关系已成为当前研究领域的热点^[13-16]。在提取患者舌象特征的基础上,使用人工智能技术构建疾病诊断模型,已实现非酒精性脂肪肝的诊断和糖尿病的风险预测^[9, 17-18]。在胃肠道疾病领域,基于舌象的人工智能诊断模型可以很好地区分胃癌和非胃癌受试者^[12]。依托人工智能技术实现舌诊标准化、客观化和智能化,对胃肠道疾病无创分类诊断,减少患者非必要有创检查的痛苦与费用,将具有较高的临床应用价值。

本研究开发了一种基于舌象图像的诊断模型,

使用机器学习技术提取并处理舌象特征,实现了智能化诊断胃肠道疾病的目的,为常见胃肠道疾病的诊断与筛查提供了一种方便、经济、无创的方法。

1 资料与方法

1.1 研究对象

前瞻性选取 2021 年 3 至 5 月在山东大学齐鲁医院消化内科接受电子内镜检查的 948 名受试者。

入选标准:①年龄在 18~80 岁之间。②接受电子内镜检查,必要时进行病理组织学检查。③入选患病组的受试者,其临床诊断中至少包含以下疾病类型中的一种:*H. pylori* 感染、胆汁反流、反流性食管炎、胃糜烂或十二指肠糜烂。其中,*H. pylori* 感染是指快速尿素酶试验、¹³C-尿素呼吸试验或活检标本中的 *H. pylori* 生物体鉴定阳性;胆汁反流定义为内镜检查期间胃内存在胆汁;糜烂性反流性食管炎是指食管远端黏膜炎症引起的糜烂;胃糜烂和十二指肠糜烂指相应部位的黏膜损伤或溃疡。④入选健康组的受试者,其内镜检查结果显示无明显异常。

排除标准:①患有严重的基础疾病(例如:冠状动脉粥样硬化性心脏病、肝病、肾病或精神疾病等)。②怀孕或哺乳期妇女。③1 个月内使用过糖皮质激素或抗生素。④近期服用铋剂或其他染色药物。⑤无法或不愿提供知情同意书。

本研究经山东大学齐鲁医院伦理委员会批准 (KLYY-202008-056),所有受试者均已签署知情同意书,并在临床试验中提供个人照片和基本信息。

1.2 方法

1.2.1 分组

根据受试者的临床诊断结果,将其分为健康组、*H. pylori* 感染组、胆汁反流组、反流性食管炎组、胃糜烂组和十二指肠糜烂组。

1.2.2 舌象图片收集

使用普通智能手机拍摄受试者的舌象图片,图像尺寸为 1 000×1 333 像素。在图片采集的过程中,没有严格控制周围环境。研究人员指导受试者平坐于座位上,目光平视前方,嘱其做吞咽动作并将舌的前 1/2 至 2/3 部分伸出,保持舌体放松、舌尖向下弯曲、舌面平整的状态,在不同的光照、亮度和拍摄角

度下为每个受试者拍摄3~4张完整的舌象图片。

1.2.3 舌象图片质控与预处理

为确保图像质量一致并减少冗余信息干扰,具有以下特征的舌象图片将被剔除:①图片模糊(雾状、曝光不足、曝光过度或失焦等)。②舌部异常染色。③部分舌象区域缺失(例如:未捕捉到舌尖等)。④异常舌形。对于最终纳入研究的舌象图片,使用直方图均衡化以增加图像对比度和增强图像细节,同时使用高斯滤波去除图像噪点。

1.2.4 舌象自动分割模块

对数据集中的500张舌象图片进行人工标注,使用Labelme软件(版本:v5.2.1)标注出舌象轮廓,见图1。

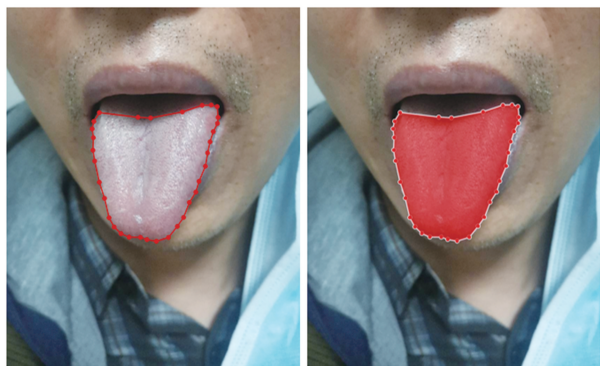


图1 舌象区域手工标注

Figure 1 Manual labeling of the tongue region

将标注好的舌象图片按照4:1的比例随机拆分

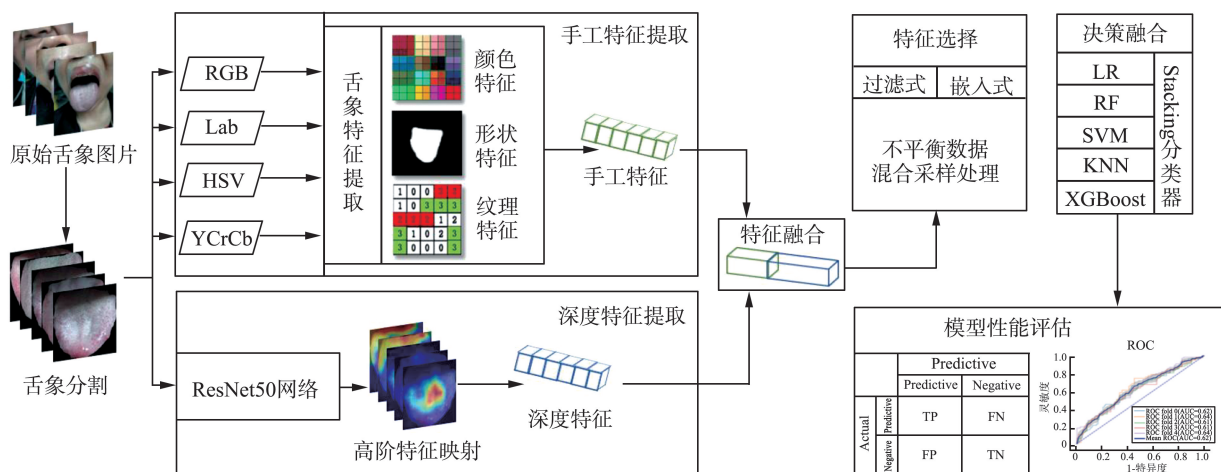


图2 胃肠道疾病分类模块的构建策略

Figure 2 Strategy for constructing a classification module for gastrointestinal diseases

1.2.5.1 手工特征提取

将舌象图片分别映射到RGB、HSV、Lab、YCrCb颜色空间,并对每个颜色通道分别提取舌象图片的以下特征:①颜色特征:基于数学统计方法,通过计算矩描述舌象图像色彩的分布。用一阶、二阶和三阶颜色矩分别代表颜色成分的平均强度、颜

为训练集和测试集。使用U-Net分割网络^[19]准确识别舌的位置并进行单独分割,以消除无关背景对其正确分类的影响。对于训练中使用的参数,设定学习率 $\alpha=0.0001$,批次大小 $b=16$,迭代次数 $\text{epoch}=100$,应用Adam优化器和交叉熵损失函数。U-Net模型对舌象分割的性能使用训练集中所有图像的交并比(intersection and union, IoU)评价。IoU是指预测区域和实际区域的交集和重合的比率,IoU越接近1预测结果越好。给定一个真实标签图像G和一个预测图像P,IoU度量定义如(1-1)所示:

$$IoU = \frac{|GR \cap PR|}{|GR| + |PR| - |GR \cap PR|} \quad (1-1)$$

其中,GR和PR分别表示图像G和P中的阳性像素集合, $|GR|$ 和 $|PR|$ 代表GR和PR中的像素数目。

用500张手工分割的舌象图片作为真实标签来训练U-Net模型,并对其余未分割图像进行自动分割。分割后的图像是标准的,根据舌的轮廓进行裁剪,只保留单独的舌体区域作为感兴趣区域(region of interest, ROI),图像尺寸被缩放为 448×448 像素。基于U-Net的自动舌象分割模型在测试集中IoU为98.43%,舌象可被正确分割。

1.2.5 胃肠道疾病分类模块

本研究针对胃肠道疾病分类模块的整体构建策略,见图2。包括手工特征提取、深度特征提取、特征融合、特征选择、不平衡样本处理和决策融合。

色方差和偏度。此外,使用颜色直方图提取舌象图片的颜色特征,从不同角度表示图像,并计算出图像中色块的频率分布。②纹理特征:纹理特征表示了图像中的纹理,反映了物体表面的形态、细节和结构等信息,具有较强的稳定性和不变性。使用灰度共生矩阵(grey level co-occurrence matrix, GLCM)^[20]、局

部二进制模式(local binary pattern, LBP)^[21]和小波分解(wavelet decomposition, WD)^[22]三种技术提取所需的舌象纹理特征。③形状特征:使用圆形和像素面积比描述边界特征,利用 Hu M-K^[23]推导出的7个矩不变量描述区域特征。以上各手工特征,见

表1 手工特征描述
Table 1 Description of manual features

特征类型	特征名称	具体特征
颜色特征	颜色矩	一阶、二阶、三阶颜色矩
	颜色直方图	颜色直方图的均值、方差、能量、对比度、熵
	GLCM	GLCM 4 方向的能量、相关性、对比度、熵、逆差矩
纹理特征	LBP	LBP 的均值、方差、能量、对比度、熵
	WD	WD 的水平、垂直和对角三个方向上的均值、方差、能量、对比度、熵
形状特征	圆形度, Hu 不变矩	圆形度, 7 个 Hu 不变矩

1.2.5.2 深度特征提取

深度特征提取流程,见图3。本研究将卷积神经(convolutional neural networks, CNN)网络模型最后一个卷积层的输出作为图像的高阶深度特征,使用全局平均池化将每个通道的特征图压缩成一个数值,得到一个固定大小的特征向量。基于迁移学习的策略,使用预训练模型提取深度特

征来描述舌象图片的深层信息。预处理操作后的舌象图片作为卷积神经网络的输入,使用在 ImageNet 数据集上预训练好的模型权重作为网络参数,加载至 ResNet50 网络中,最后提取经池化层操作后的特征向量。利用 ResNet50 模型作为深度特征提取器,在每张舌象图片中各提取 2 048 维高阶深度特征向量。

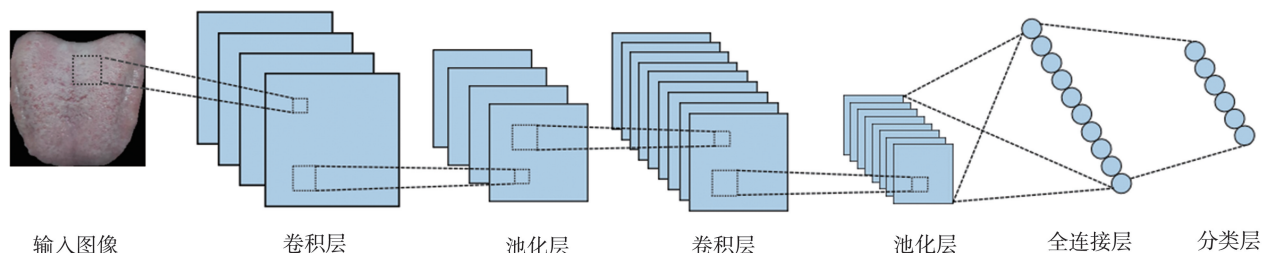


图3 深度特征提取示意图

Figure 3 Schematic diagram of deep feature extraction

1.2.5.3 特征融合

对从舌象图片提取得到的手工特征矩阵和深度特征矩阵,按照(1-2)进行线性融合:

$$F = \alpha F_1 + \beta F_2 \quad (1-2)$$

其中, α 和 β 均表示权重系数。

1.2.5.4 特征选择

基于机器学习算法中的集成学习思想,本研究提出了一种将过滤式和嵌入式特征选择方法相组合的逐级特征选择算法,用于舌象图片特征选择,见图4。该特征选择过程分为以下两步进行:①使用过滤式特征选择方法对融合特征进行筛选,去除大部分冗余特征。②对每个子分类器根据嵌入式特征选择算法 GBDT 进行更精确的特征子集搜索,得到与胃肠道疾病最相关的特征子集。

1.2.5.5 不平衡样本处理

考虑到样本数据的不平衡特性,在特征选择之

后对训练集数据使用一种新的混合重采样方法 BSFCM^[24],该方法将 Borderline-SMOTE 过采样和基于模糊均值聚类下采样相结合,可缓解模型训练的过拟合,解决样本数据的类内和类间不平衡问题。未在测试集中使用上述方法,以保证测试集样本数据分布不受破坏,减少过拟合。

1.2.5.6 决策融合

使用 Stacking^[25]方法进行决策融合,选取支持向量机(support vector machines, SVM)、随机森林(random forest, RF)、K-近邻算法(K-nearest neighbor, KNN)、XGBoost 分类器作为第一层的子学习器,逻辑回归(Logistic regression, LR)作为第二层的元学习器,融合各子学习器的分类结果,最终得到泛化能力较强的强学习器。

1.3 统计学处理

使用 Python(版本:3.7.13)编程语言进行模型

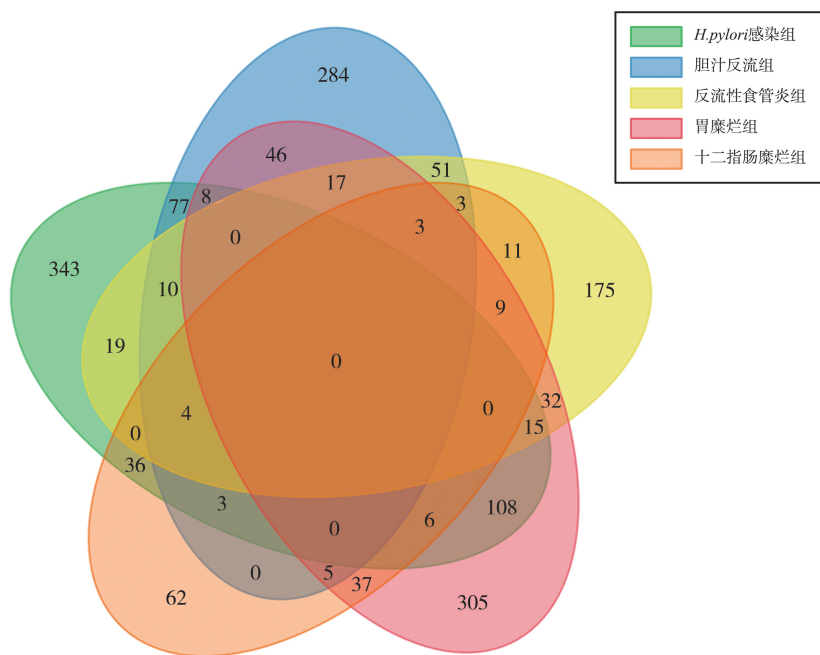


图6 患病组不同疾病类型纳入的舌象图片数目

Figure 6 Number of tongue images included in different disease types in the sick group

经过图像质控,剔除不符合标准的图片 171 张,其余符合标准的 3 140 张图片构成了用于后续实验分析的多标签舌象图片数据集,包括健康组 905 张和患病组 1 262 张。构建 6 个二分类模型以评估分类模型对不同胃肠道疾病的诊断能力,即患病/健康(1 262 张/905 张)、*H. pylori* 感染/健康

(629 张/905 张)、胆汁反流/健康(511 张/905 张)、反流性食管炎/健康(349 张/905 张)、胃糜烂/健康(591 张/901 张)和十二指肠糜烂/健康(179 张/905 张)。

2.2 特征融合法的有效性评估

对特征融合法的有效性进行评估,见表 2。

表 2 特征融合法的有效性评估

Table 2 Evaluation of the validity of feature fusion method

方法	特征	分类器	评价指标				
			AUC	ACC	TPR	PPV	F1 值
特征融合	手工特征(RGB)	Stacking	0.712	0.672	0.803	0.702	0.749
	手工特征(Lab)	Stacking	0.738	0.673	0.746	0.724	0.735
	手工特征(HSV)	Stacking	0.724	0.680	0.819	0.705	0.758
	手工特征(YCrCb)	Stacking	0.736	0.698	0.777	0.740	0.758
	手工特征(多色彩空间融合)	Stacking	0.769	0.714	0.792	0.752	0.771
	深度特征(Resnet50)	Stacking	0.779	0.747	0.858	0.743	0.779
	深度特征(VGG16)	Stacking	0.772	0.730	0.813	0.752	0.761
	深度特征(Resnet50+VGG16)	Stacking	0.783	0.742	0.841	0.747	0.774
	融合特征	Stacking	0.808	0.753	0.850	0.761	0.803

首先,对多颜色空间校正的结果进行验证。四种颜色空间的特征提取结果表明,使用多颜色空间的手工特征的 AUC 为 0.769,ACC 为 0.714,高于其他四种单一颜色空间的手工特征。其诊断的灵敏度(TPR=0.792)和特异度(TNR=0.658)高于 Lab 颜色空间(TPR=0.764,TNR=0.621)。在单一颜色空间的手工特征中,RGB 颜色空间的诊断性能最差(AUC=0.712,ACC=0.672),对亮度变化不敏感的 YCrCb 颜色空间的性能相对较好(AUC=0.736,

ACC=0.698)。因此,在手工特征提取时,选用多颜色空间校正方法提高了胃肠道疾病诊断模型的性能。

使用 Stacking 集成分类器对手工特征和深度特征进行特征融合,融合特征模型的 AUC 为 0.808,ACC 为 0.753,高于单一深度特征(AUC=0.779,ACC=0.747)和单一手工特征(AUC=0.769,ACC=0.714),其他评价指标也优于单一特征。以上结果表明,融合特征包含更多的舌象信息。因此,选用特

征融合方法提高了诊断模型的性能。

2.3 BSFCM 混合采样法处理不平衡样本对模型分类性能的影响

对于不平衡样本,本研究在训练集上使用 BSFCM 混合采样方法来增强样本,见表 3。

使用 BSFCM 混合采样方法进行数据增强明显提高了融合模型对 *H.pylori* 感染、胆汁反流、反流性食管炎、胃糜烂或十二指肠糜烂的诊断的 AUC 和 ACC 值,表明少数类样本得到了重点考虑,模型的

分类性能得到明显提高。在增强样本后,融合模型的灵敏度 TPR 在五类疾病的分类中均提升 0.1 以上,尤其是对于类别最不平衡的十二指肠糜烂,TPR 从 0.324 提高到了 0.526。五种疾病的混淆矩阵亦可看出,在对不平衡样本使用数据增强后,融合模型可明显减少被误判为阴性的患病阳性样本的数量,即漏诊率降低,这对于模型的疾病筛查应用而言至关重要,见图 7。

表 3 使用或不使用数据增强算法的融合模型的疾病分类性能评估

Table 3 Evaluation of fusion models with and without data enhancement algorithms for disease classification

诊断模型	疾病类型	评价指标					
		AUC	ACC	TPR	TNR	PPV	F1 值
不使用数据增强算法的融合模型	<i>H.pylori</i> 感染	0.806	0.712	0.531	0.843	0.708	0.607
	胆汁反流	0.771	0.768	0.516	0.894	0.71	0.598
	反流性食管炎	0.777	0.769	0.368	0.943	0.737	0.491
	胃糜烂	0.808	0.777	0.636	0.868	0.758	0.692
	十二指肠糜烂	0.804	0.866	0.250	0.973	0.615	0.355
使用数据增强算法的融合模型	<i>H.pylori</i> 感染	0.816	0.746	0.667	0.806	0.721	0.693
	胆汁反流	0.829	0.759	0.610	0.847	0.703	0.653
	反流性食管炎	0.800	0.777	0.539	0.88	0.661	0.594
	胃糜烂	0.833	0.793	0.729	0.835	0.741	0.735
	十二指肠糜烂	0.818	0.871	0.526	0.944	0.667	0.588

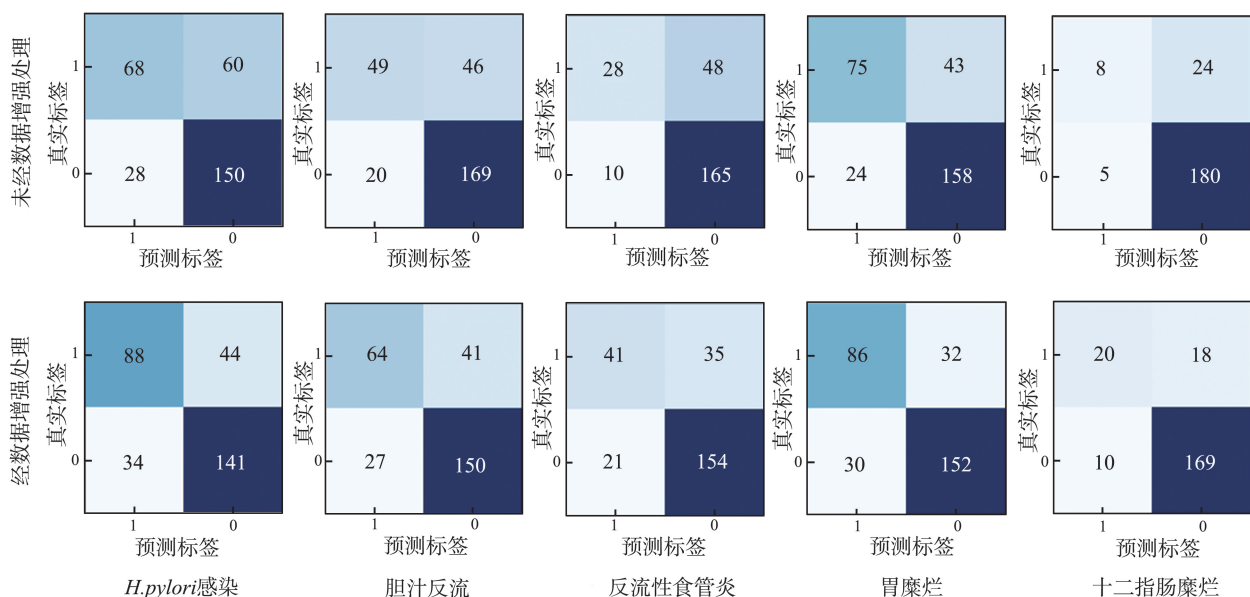


图 7 使用 BSFCM 混合采样法处理前后五种疾病诊断性能的混淆矩阵对比

Figure 7 Confusion matrix comparison of diagnostic performance of five diseases using BSFCM hybrid sampling

2.4 决策融合法的有效性评估

使用集成学习策略来实现决策融合,子分类器的决策融合分别通过 Voting 集成法和 Stacking 集成法实现,结果见表 4。

且与单一分类算法相比,以上两种集成模型具有更好的性能,AUC 分别为 0.791 和 0.808,ACC 分

别为 0.726 和 0.753,灵敏度分别为 0.799 和 0.850,特异度分别为 0.674 和 0.683。其中,Stacking 模型的性能相对更好,其 AUC、ACC、灵敏度、特异度和 F1 值均比 Voting 模型高。以上结果表明了本研究中使用集成学习策略的决策融合法在模型分类中的有效性,此方法可有效改善舌象分类模型的性能。

表4 决策融合法有效性评估
Table 4 Evaluation of the validity of decision fusion method

方法	特征	分类器	评价指标					
			AUC	ACC	TPR	TNR	PPV	F1 值
决策融合	融合特征	LR	0.774	0.717	0.858	0.616	0.714	0.779
	融合特征	SVM	0.798	0.725	0.845	0.639	0.704	0.768
	融合特征	KNN	0.781	0.718	0.811	0.651	0.742	0.790
	融合特征	RF	0.766	0.700	0.790	0.635	0.716	0.754
	融合特征	XGBoost	0.783	0.732	0.825	0.665	0.720	0.779
	融合特征	Voting	0.791	0.726	0.799	0.674	0.762	0.780
	融合特征	Stacking	0.808	0.753	0.850	0.683	0.761	0.803

3 讨论

消化系统疾病具有高发病率和患病率的显著特点^[1]。全世界约有一半人口感染 *H. pylori*, *H. pylori* 感染会导致慢性胃炎,进而发展为严重的胃十二指肠病变,包括消化性溃疡、胃癌和胃黏膜相关淋巴瘤淋巴癌^[26]。最近的一项临床横断面研究表明,胆汁反流是发生胃癌前病变和胃癌的独立危险因素^[27]。因此,提高对早期消化道疾病的检出率至关重要。目前,多数胃肠道疾病的诊断和筛查依赖电子内镜及病理组织学检查,给患者个人及社会带来沉重负担。因此,迫切需要发展非侵入性、方便且有效的筛查和诊断方法,以减少患者痛苦以及社会和经济负担。

人工智能技术的发展日新月异,已被应用于医学成像^[28]、癌症靶点识别和药物发现^[29]等领域,为医学的创新发展持续供能。本研究构建了一个基于机器学习的胃肠道疾病舌诊模型。首先,使用智能手机收集了健康、*H. pylori* 感染、胆汁反流、反流性食管炎、胃糜烂和十二指肠糜烂受试者的 3 140 张舌象图片,通过 U-Net 分割模型对舌象进行自动分割,得到标准化舌象。然后,分别提取舌象的手工特征和深度特征,并基于信息融合的思想,进行特征融合与决策融合,以构建诊断模型。在原始舌象图片数据集中,模型的总体 AUC 为 0.808,疾病诊断准确率为 0.753,灵敏度为 0.850,特异度为 0.683。在高精确率(PPV=0.761)情况下,该模型对疾病分类具有较高的灵敏度,表明该模型具有漏诊率低的特征,即可检测出绝大部分 *H. pylori* 感染(AUC=0.816)、胆汁反流(AUC=0.829)、反流性食管炎(AUC=0.800)、胃糜烂(AUC=0.833)和十二指肠糜烂(AUC=0.818)患者。本文提供了一种无创且有效的常见胃肠道疾病诊断与筛查的新方法。

根据中医的观点,舌象与人体的脏腑经络密切

相关。舌诊则是以观察病人舌象变化以诊察疾病的一种方法,其为中医辨证诊疗提供较为重要的客观依据。已有多项研究探讨了舌象特征与包括胃肠道疾病在内的多种疾病之间的关系,如冠心病^[30-31]、糖尿病^[32]、脂肪肝^[33]及多种肿瘤如肺癌^[34]、胃癌^[12]等。然而,中医舌诊具有极强的主观性和经验性,将人工智能技术与中医舌诊相结合,通过预处理、舌象检测、分割、特征提取和舌象分析等技术,或可使舌诊断的标准化和客观化得到提升^[35]。本研究将人工智能机器学习运用于探讨舌象在常见胃肠道疾病诊断中的价值,为舌象作为 *H. pylori* 感染、胆汁反流、反流性食管炎、胃糜烂和十二指肠糜烂的诊断标志物提供了有力的依据。

在疾病筛查的实际应用中,正常样本数量往往远多于患病样本数量,即存在样本数据不平衡问题。若不对样本进行不平衡问题的处理,使得漏诊率偏高,将极大削弱诊断模型的实际应用价值。考虑不同疾病类别分组的舌象图片数据的不平衡性,本研究在特征选择之后对训练集数据使用了一种新的混合重采样方法 BSFCM^[24]。该方法将过采样和欠采样的优势相结合,综合解决了样本类内和类间不平衡问题,使得模型对五种疾病的漏诊率显著降低。

舌象的手工特征提取是以中医理论为基础的,有助于区分舌象特征,其工作机制包括大量经过临床实践检验的先验知识理论。传统的卷积神经网络通过增加网络层数来提升网络性能,但随着网络层数的不断增加,会出现梯度消失和退化现象,使网络模型在训练过程中难以收敛。针对上述问题,He 等^[36]提出了深度残差网络 ResNet,与传统的 CNN 算法相比,ResNet 通过引入残差连接机制,可增加网络深度的同时提高网络性能。本研究在深度特征提取过程中,选用 ResNet50 算法作为一个特征提取主干网络,在每张舌象图片中各提取 2 048 维高阶深度特征向量,该算法提高了分类器的准确性,AUC 值高于手工特征。

为避免单一特征的不足,本研究对提取的手工特征和深度特征进行了融合,AUC值为0.808,高于单一手工特征(AUC=0.769)和深度特征(AUC=0.779)模型。传统的基于特征的机器学习方法在表现数据的自然特征方面具有局限性,原因在于这些方法需要大量的专业知识来设计特征提取器,将获得的数据转化为特征向量,随后通过分类模型对其进行检测与分类。经典的机器学习算法擅长分析结构性数据,而深度学习算法是机器学习算法的一个高级子领域,它可以将图像与给定的任务联系起来,并能自动提取和分析图像的高级特征信息^[37]。本研究采用的特征融合方法结合了机器学习和深度学习的优势之处,将手工特征和深度特征相结合,从而准确全面地表征了舌象信息。在构建分类器时,分别使用 Voting 和 Stacking 方法构建决策融合模型。该模型在融合各子学习器分类结果的情况下,最终得到的学习器具有泛化能力较强的优势,能够充分提取并拟合舌象图像特征,实现胃肠道疾病的准确分类。与大部分单一分类算法相比,决策融合模型具有更好的性能。

本研究构建的模型的优点是模型训练速度快、时间复杂度低,基于研究者先验知识的人工特征比深度模型更具可解释性。该模型实现了基于舌象特征的几种常见胃肠道疾病的诊断,兼具无创和有效的优势,且漏诊率较低,具有两种潜在的应用场景。首先,该模型可辅助临床医生进行相关疾病诊断。该模型对于常规临床使用来说足够简单,通过拍摄舌象照片,即可快速给出胃肠道疾病的诊断结果以辅助医生诊断。其次,该模型也适用于社区或个人的胃肠道疾病筛查。舌象特征和胃肠道疾病之间具有紧密联系,舌象特征可能是诊断胃肠道疾病的一个潜在标志物。因此,患者可以利用手机上拍摄的舌象图片实现对胃肠道疾病的初步筛查,且该模型漏诊率低,符合疾病筛查和诊断中的应检尽检原则。

本研究具有一定的局限性。首先,只纳入了中国受试者,将来纳入多种族的人群可提高模型的普适性。其次,本研究依据 *H. pylori* 感染、胆汁反流、反流性食管炎、胃糜烂和十二指肠糜烂数据进行了模型的开发和验证,仍需要丰富胃肠道疾病中常见病、多发病及现有疾病的相近病症数据,以提高该模型的鉴别诊断能力和可推广性。此外,纳入的样本量有待于进一步的扩充,舌象图片采集设备尚需进一步升级,且将来仍需通过临床实践对该模型进行优化。

总之,本研究基于特征融合和决策融合方法构

建了疾病诊断模型,从特征提取和特征选择的角度优化了机器学习建模过程,提高了对常见胃肠道疾病的诊断性能,提供了一种方便、经济、无创的常见胃肠道疾病筛查方法,能够减少患者痛苦以及减轻消化道疾病所带来的社会和经济负担,具有临床应用价值。

参考文献:

- [1] Wang YC, Huang YT, Chase RC, et al. Global burden of digestive diseases: a systematic analysis of the global burden of diseases study, 1990 to 2019[J]. *Gastroenterology*, 2023, 165(3): 773-783.
- [2] Wang WY, Zhou H, Wang YF, et al. Current policies and measures on the development of traditional Chinese medicine in China[J]. *Pharmacol Res*, 2021, 163: 105187. doi:10.1016/j.phrs.2020.105187.
- [3] 王庆盛,高慧,许朝霞,等. 冠心病及其不同合并病患者的舌诊参数特征分析[J]. *中华中医药杂志*, 2022, 37(3): 1316-1320.
WANG Qingsheng, GAO Hui, XU Zhaoxia, et al. Analysis of tongue diagnostic parameters in patients with coronary heart disease and its different complications[J]. *China Journal of Traditional Chinese Medicine and Pharmacy*, 2022, 37(3): 1316-1320.
- [4] 于然,娄彦妮,梁婉娴,等. 基于食管癌高发区人群筛查探索反流性食管炎及 Barrett 食管的舌象转化规律[J]. *中医杂志*, 2021, 62(9): 782-788.
YU Ran, LOU Yanni, LIANG Wanxian, et al. The change characteristics of tongue manifestation of reflux esophagitis and barrett's esophagus based on an esophageal carcinoma screening study in high-risk areas of China[J]. *Journal of Traditional Chinese Medicine*, 2021, 62(9): 782-788.
- [5] Wu TC, Lu CN, Hu WL, et al. Tongue diagnosis indices for gastroesophageal reflux disease: a cross-sectional, case-controlled observational study[J]. *Medicine*, 2020, 99(29): e20471. doi:10.1097/MD.00000000000020471.
- [6] Hou B, Zeng Y, Ling H, et al. Correlation between helicobacter pylori infection and tongue manifestations: a meta-analysis[J]. *Digital Chinese Medicine*, 2018, 1(2): 155-163.
- [7] Haug CJ, Drazen JM. Artificial intelligence and machine learning in clinical medicine, 2023[J]. *N Engl J Med*, 2023, 388(13): 1201-1208.
- [8] Litjens G, Kooi T, Bejnordi BE, et al. A survey on deep learning in medical image analysis[J]. *Med Image Anal*, 2017, 42: 60-88. doi:10.1016/j.media.2017.07.005.
- [9] Fan SY, Chen B, Zhang XR, et al. Machine learning al-

- gorithms in classifying TCM tongue features in diabetes mellitus and symptoms of gastric disease[J]. *Eur J Integr Med*, 2021, 43: 101288. doi: 10.1016/j.eujim.2021.101288.
- [10] Bi WL, Hosny A, Schabath MB, et al. Artificial intelligence in cancer imaging: clinical challenges and applications[J]. *CA Cancer J Clin*, 2019, 69(2): 127-157.
- [11] Sun TG, Mao L, Chai ZK, et al. Predicting the proliferation of tongue cancer with artificial intelligence in contrast-enhanced CT[J]. *Front Oncol*, 2022, 12: 841262. doi:10.3389/fonc.2022.841262.
- [12] Yuan L, Yang L, Zhang SC, et al. Development of a tongue image-based machine learning tool for the diagnosis of gastric cancer: a prospective multicentre clinical cohort study [J]. *E Clinical Medicine*, 2023, 57: 101834. doi:10.1016/j.eclinm.2023.101834.
- [13] Li MY, Zhu DJ, Xu W, et al. Application of U-net with global convolution network module in computer-aided tongue diagnosis [J]. *J Healthc Eng*, 2021: 5853128. doi:10.1155/2021/5853128.
- [14] Ma CZ, Zhang P, Du SY, et al. Construction of tongue image-based machine learning model for screening patients with gastric precancerous lesions[J]. *J Pers Med*, 2023, 13(2): 271. doi:10.3390/jpm13020271.
- [15] Wang XZ, Luo SY, Tian GH, et al. Deep learning based tongue prickles detection in traditional Chinese medicine [J]. *Evid Based Complement Alternat Med*, 2022: 5899975. doi:10.1155/2022/5899975.
- [16] 姜楠, 袁莉, 汪莉, 等. 初诊胃癌患者舌象特征的客观化研究[J]. *中华中医药杂志*, 2023, 38(1): 427-433. JIANG Nan, YUAN Li, WANG Li, et al. Objective study on tongue features of newly diagnosed gastric cancer patients [J]. *China Journal of Traditional Chinese Medicine and Pharmacy*, 2023, 38(1): 427-433.
- [17] Jiang T, Guo XJ, Tu LP, et al. Application of computer tongue image analysis technology in the diagnosis of NAFLD[J]. *Comput Biol Med*, 2021, 135:104622. doi: 10.1016/j.combiomed.2021.104622.
- [18] Li J, Chen QG, Hu XJ, et al. Establishment of noninvasive diabetes risk prediction model based on tongue features and machine learning techniques[J]. *Int J Med Inform*, 2021, 149: 104429. doi:10.1016/j.ijmedinf.2021.104429.
- [19] Olaf R, Philipp F, Thomas B. U-Net: convolutional networks for biomedical image segmentation [C]. *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18, 2015: 234-241.*
- [20] Mohanaiah P, Sathyanarayana P, GuruKumar L, et al. Image texture feature extraction using GLCM approach [J]. *In J Sci Res Publ*, 2013, 3(5): 1-5.
- [21] Kaplan K, Kaya Y, Kuncan M, et al. Brain tumor classification using modified local binary patterns (LBP) feature extraction methods[J]. *Med Hypotheses*, 2020, 139: 109696. doi:10.1016/j.mehy.2020.109696.
- [22] Arivazhagan S, Ganesan L, Priyal SP. Texture classification using Gabor wavelets based rotation invariant features[J]. *Pattern Recognit Lett*, 2006, 27(16): 1976-1982.
- [23] Hu MK. Visual pattern recognition by moment invariants [J]. *IRE Trans Inf Theory*, 1962, 8(2): 179-187.
- [24] Liu YR, Qiao X, Gao R. Plankton classification on imbalanced dataset via hybrid resample method with Light-BGM[C]. 2021 6th International Conference on Image, Vision and Computing (ICIVC). July 23-25, 2021, Qingdao, China. IEEE, 2021: 191-195. doi: 10.1109/ICIVC52351.2021.9526988.
- [25] Džeroski S, Ženko B. Is combining classifiers with stacking better than selecting the best one? [J]. *Mach Learn*, 2004, 54(3): 255-273.
- [26] Malfertheiner P, Camargo MC, El-Omar E, et al. *Helicobacter pylori* infection [J]. *Nat Rev Dis Primers*, 2023, 9: 19. doi:10.1038/s41572-023-00431-8.
- [27] Zhang LY, Zhang J, Li D, et al. Bile reflux is an independent risk factor for precancerous gastric lesions and gastric cancer: an observational cross-sectional study[J]. *J Dig Dis*, 2021, 22(5): 282-290.
- [28] Gore JC. Artificial intelligence in medical imaging[J]. *Magn Reson Imaging*, 2020, 68: 1-4. doi: 10.1016/j.mri.2019.12.006.
- [29] You YJ, Lai X, Pan Y, et al. Artificial intelligence in cancer target identification and drug discovery[J]. *Signal Transduct Target Ther*, 2022, 7(1): 156. doi:10.1038/s41392-022-00994-0.
- [30] Lin S, Li ZG, Fu BW, et al. Feasibility of using deep learning to detect coronary artery disease based on facial photo[J]. *Eur Heart J*, 2020, 41(46): 4400-4411.
- [31] Liang B, Li R, Lu J, et al. Tongue diagnostic parameters-based diagnostic signature in coronary artery disease patients with clopidogrel resistance after percutaneous coronary intervention[J]. *Explore*, 2023, 19(4): 528-535.
- [32] Zhang NN, Jiang ZX, Li JX, et al. Multiple color representation and fusion for diabetes mellitus diagnosis based on back tongue images[J]. *Comput Biol Med*, 2023, 155: 106652. doi:10.1016/j.combiomed.2023.106652.