

基于YOLOv26n双模型协同的舌象智能识别方法

赵子政¹, 闫沛媛², 刘静超^{1*}, 闻丽芬²

¹西京学院计算机学院, 陕西 西安

²西京学院医学院, 陕西 西安

收稿日期: 2026年6月5日; 录用日期: 2026年7月4日; 发布日期: 2026年7月14日

摘要

针对中医舌象图像识别中存在背景干扰明显、舌象类别视觉差异细微、多种舌象特征可能同时出现以及形态异常特征难以稳定识别等问题, 提出一种基于YOLOv26n双模型协同的多标签舌象智能识别方法。首先利用YOLOv26n目标检测模型对输入图像中的舌体区域进行定位, 获得舌体边界框并裁剪舌体感兴趣区域, 以减少面部、嘴唇和背景区域对后续识别的影响; 随后将舌体ROI输入改进后的YOLOv26n多标签分类模型, 对正常舌、胖大舌、裂纹舌、薄白苔、黄厚苔、白厚苔、剥落苔、红点舌、淡白舌、淡红舌、黄厚腻苔、黑苔、地图舌和齿痕舌等舌象特征进行识别。考虑到一幅舌象图像可能同时包含舌色、舌苔和舌体形态等多种特征, 本文将舌象分类任务建模为多标签识别问题, 采用Sigmoid输出和加权二元交叉熵损失函数进行训练, 并使用mAP、Micro-F1、Macro-F1和Hamming Loss等指标进行评价。针对齿痕舌等边缘形态异常特征, 进一步引入几何形态学辅助分析方法, 通过红色通道提取、阈值分割、轮廓提取、凸包分析和凸缺陷检测计算舌体边缘凹陷得分, 并将几何得分作为辅助特征与深度学习模型输出进行融合。实验表明, 该方法能够实现舌体检测、ROI提取、多标签舌象识别和边缘形态辅助分析, 并可在树莓派5边缘端完成部署验证。所提出方法能够降低非舌体区域干扰, 提高多种舌象特征共存场景下的识别能力, 并为齿痕类形态异常提供可解释的辅助判断依据, 可为中医舌象智能分析提供一种可行的深度学习实现方案。

关键词

YOLOv26n, 舌象识别, 多标签分类, 几何形态学, 边缘部署

An Intelligent Tongue Image Recognition Method Based on YOLOv26n Dual-Model Collaboration

Zizheng Zhao¹, Peiyuan Yan², Jingchao Liu^{1*}, Lifan Wen²

*通讯作者。

文章引用: 赵子政, 闫沛媛, 刘静超, 闻丽芬. 基于YOLOv26n双模型协同的舌象智能识别方法[J]. 计算机科学与应用, 2026, 16(7): 9-21. DOI: 10.12677/csa.2026.167236

Abstract

To address the problems of obvious background interference, subtle visual differences among tongue image categories, coexistence of multiple tongue features, and unstable recognition of morphological abnormalities in traditional Chinese medicine tongue image analysis, this paper proposes a multi-label intelligent tongue image recognition method based on YOLOv26n dual-model collaboration. First, a YOLOv26n object detection model is used to locate the tongue region in the input image. The tongue bounding box is obtained and the tongue region of interest is cropped, thereby reducing the influence of irrelevant regions such as the face, lips, and background on subsequent recognition. Then, the cropped tongue ROI is fed into an improved YOLOv26n multi-label classification model to identify tongue features, including normal tongue, enlarged tongue, fissured tongue, thin white coating, yellow thick coating, white thick coating, peeled coating, red spot tongue, pale tongue, light red tongue, yellow thick greasy coating, black coating, geographic tongue, and scalloped tongue. Considering that a single tongue image may simultaneously contain multiple features related to tongue color, tongue coating, and tongue morphology, the tongue classification task is formulated as a multi-label recognition problem. The model is trained using Sigmoid outputs and a weighted binary cross-entropy loss function, and evaluated using mAP, Micro-F1, Macro-F1, and Hamming Loss. For edge-related morphological abnormalities such as scalloped tongue, a geometric morphology-assisted analysis method is further introduced. This method calculates the tongue edge depression score through red channel extraction, threshold segmentation, contour extraction, convex hull analysis, and convexity defect detection. The geometric score is then fused with the output of the deep learning model as an auxiliary feature. Experimental results show that the proposed method can achieve tongue detection, ROI extraction, multi-label tongue feature recognition, and edge morphology-assisted analysis. It can also be deployed and verified on a Raspberry Pi 5 edge platform. The proposed method can reduce interference from non-tongue regions, improve the recognition ability in scenarios where multiple tongue features coexist, and provide interpretable auxiliary evidence for scalloped tongue-like morphological abnormalities. It offers a feasible deep learning-based solution for intelligent tongue image analysis in traditional Chinese medicine.

Keywords

YOLOv26n, Tongue Image Recognition, Multi-Label Classification, Geometric Morphology, Edge Deployment

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

舌诊是中医诊疗体系中的重要内容之一。医生通常通过观察舌体颜色、舌苔厚薄、舌面纹理以及舌体形态等特征，对人体健康状态进行辅助判断。然而，传统舌诊主要依赖医生经验，结果容易受到主观认知、光照环境、观察角度和拍摄距离等因素影响。随着计算机视觉和深度学习技术的发展，利用图像识别方法对舌象进行自动分析成为中医智能化研究的重要方向。

近年来,卷积神经网络、目标检测算法和图像分类模型在医学图像分析、目标识别和智能诊断辅助场景中得到了广泛应用。YOLO 系列算法由于具有端到端检测、推理速度快和部署灵活等特点,被广泛应用于实时目标检测任务[1]。在舌象识别任务中,深度学习模型能够自动提取颜色、纹理和形态等高层特征,相比传统人工特征方法具有更强的表达能力[2]。

但是,现有舌象识别方法仍存在问题。第一,实际采集的舌象图像通常包含面部、嘴唇、牙齿和背景区域,如果直接对整幅图像进行分类,模型容易受到无关区域干扰。第二,舌象类别之间存在较强相似性,例如正常舌与淡红舌、黄厚苔与黄厚腻苔等类别在颜色和纹理上差异较小,容易造成分类混淆。第三,齿痕舌等形态异常特征主要表现为舌体边缘凹陷,单纯依赖分类模型难以给出可解释的形态依据。第四,在实际应用中,舌象识别方法还需要具备一定轻量化部署能力,以适应基层健康检测和便携式终端场景。

针对上述问题,本文提出一种基于 YOLOv26n 双模型协同的多标签舌象智能识别方法。该方法首先利用 YOLOv26n 检测模型定位舌体区域,裁剪舌体 ROI,再将 ROI 输入改进后的 YOLOv26n 多标签分类模型,对舌色、舌苔和舌体形态等多类舌象特征进行联合识别。考虑到舌象图像中多种特征可能同时存在,本文将分类任务由传统单标签分类调整为多标签识别任务,采用 Sigmoid 输出和加权二元交叉熵损失函数进行训练,并通过 mAP、Micro-F1、Macro-F1 和 Hamming Loss 等指标进行评价。同时,针对齿痕舌等边缘形态异常,引入几何形态学辅助分析方法,对舌体轮廓凹陷程度进行量化计算,并将几何形态特征与深度学习特征进行融合,以提高齿痕类舌象识别的稳定性和可解释性。

2. 相关技术基础

2.1. YOLO 目标检测算法

YOLO 是一类典型的一阶段目标检测算法,其核心思想是将目标定位和类别预测统一到一个端到端网络中完成。与两阶段检测方法相比,YOLO 算法具有较快的推理速度,适合实时检测和嵌入式部署场景[3]。在本文舌象识别任务中,YOLOv26n 被用于舌体检测和舌象分类两个子任务。其中,检测模型负责从原始图像中定位舌体区域,分类模型负责对裁剪后的舌体 ROI 进行舌象特征识别。

舌象图像识别不同于一般自然图像分类。舌体目标通常位于面部区域内部,周围可能存在嘴唇、牙齿、皮肤和背景等干扰区域。因此,在分类之前先进行目标检测和 ROI 裁剪,有助于降低无关区域对分类模型的影响。

2.2. 舌象图像识别的多标签任务特点

舌象识别任务通常涉及舌色、舌苔和舌体形态等多类视觉特征。舌色类特征主要包括正常舌、淡白舌、淡红舌、红点舌等;舌苔类特征包括薄白苔、黄厚苔、白厚苔、黄厚腻苔、剥落苔等;形态类特征包括胖大舌、裂纹舌、地图舌和齿痕舌等。

与普通单标签图像分类任务不同,舌象图像往往具有多标签属性。一幅舌象图像可能同时包含多个舌象特征,例如同一张图像中可能同时出现“淡白舌”“白厚苔”和“齿痕舌”。这些特征分别来自舌色、舌苔和形态三个维度,彼此之间并非完全互斥。因此,如果将舌象识别简单处理为单标签分类任务,模型只能输出一个最高概率类别,容易忽略同一图像中同时存在的其他舌象特征。

基于上述特点,本文将舌象分类任务建模为多标签识别问题。每张舌象 ROI 图像对应一个多维二值标签向量,每一维表示对应舌象特征是否存在。该方式能够同时识别舌色、舌苔和舌体形态等多种特征,更符合中医舌诊中多特征综合观察的实际过程。

2.3. 几何形态学分析方法

图像形态学方法常用于轮廓提取、区域分割、边缘分析和形状描述等任务[4]。在舌象形态分析中，齿痕舌通常表现为舌体左右边缘存在局部凹陷。与颜色和舌苔特征不同，齿痕舌具有明显的边缘几何形态特征，因此可以通过轮廓分析、凸包计算和凸缺陷检测进行辅助判断。

凸包可以理解包围目标轮廓的最小凸多边形，而凸缺陷表示原始轮廓相对于凸包向内凹陷的区域。通过统计舌体左右边缘的有效凸缺陷深度，可以对边缘凹陷程度进行量化分析。本文将该方法作为深度学习分类结果的辅助验证手段，而不是替代分类模型。

3. 基于 YOLOv26n 的舌象识别方法

3.1. 方法总体框架

本文提出的多标签舌象智能识别方法主要包括图像输入、舌体检测、ROI 裁剪、多标签舌象识别、几何形态学辅助分析、特征融合和结果输出七个步骤。首先，系统采集原始舌象图像，并利用 YOLOv26n 检测模型定位舌体区域，得到舌体边界框；随后根据检测框裁剪舌体 ROI，以减少面部、嘴唇、牙齿和背景区域对后续识别的干扰。然后，将裁剪后的舌体 ROI 输入改进后的 YOLOv26n 多标签分类模型，通过 Sigmoid 输出同时判断舌色、舌苔和舌体形态等多种舌象特征。针对齿痕舌等边缘形态异常，本文进一步设计几何形态学辅助分析模块，通过阈值分割、轮廓提取、凸包分析和凸缺陷检测计算边缘凹陷得分。最后，将深度学习模型输出的多标签概率与几何形态特征进行融合，生成最终舌象识别结果和辅助分析报告。该框架既能够提高多特征共存场景下的识别能力，又能够为齿痕类形态异常提供一定的可解释依据。总体流程如图 1 所示。

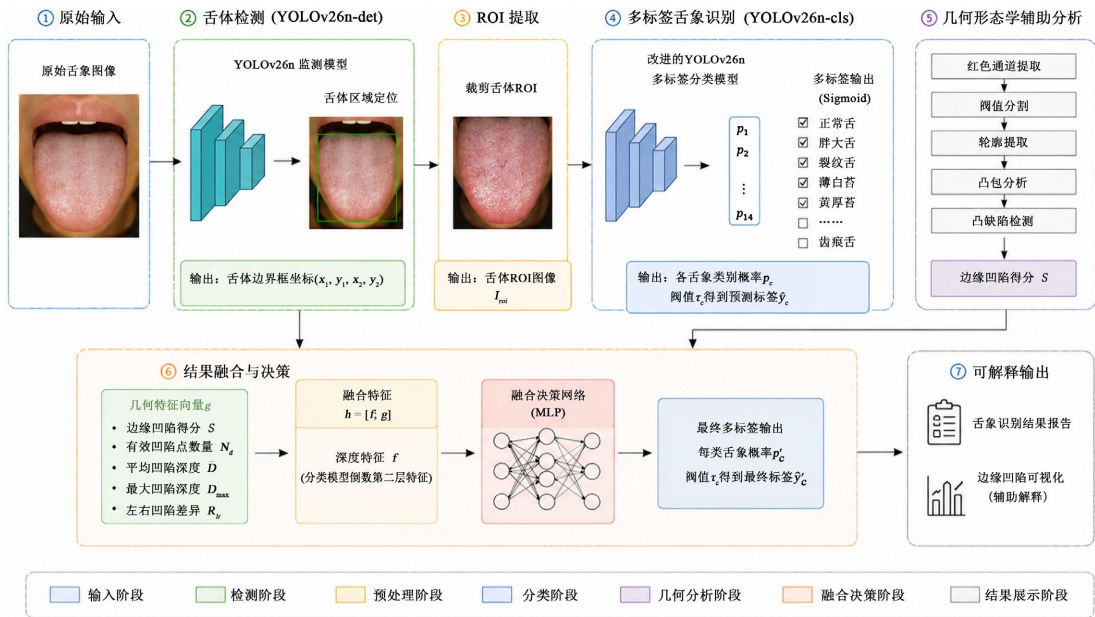


Figure 1. Tongue image recognition workflow based on YOLOv26n dual-model collaboration

图 1. 基于 YOLOv26n 双模型协同的舌象识别流程图

3.2. 舌体区域检测模型

在舌体检测阶段，系统将原始图像输入 YOLOv26n 检测模型，模型输出舌体区域的边界框坐标。设

检测框坐标为:

$$B = (x_1, y_1, x_2, y_2) \quad (1)$$

其中, x_1 和 y_1 表示检测框左上角坐标, x_2 和 y_2 表示检测框右下角坐标。根据检测框坐标, 从原始图像中裁剪舌体 ROI:

$$I_{roi} = I[y_1 : y_2, x_1 : x_2] \quad (2)$$

其中, I 表示原始图像, I_{roi} 表示裁剪后的舌体区域图像。

这种检测先行的方式具有较强实用意义。实际舌象采集过程中, 图像中可能存在皮肤、嘴唇、背景和阴影等非舌体区域。如果不进行舌体定位, 分类模型可能学习到无关信息, 降低泛化能力。通过舌体检测和 ROI 裁剪, 可以使后续分类任务更加聚焦。

3.3. ROI 引导的多标签舌象识别模型

在获得舌体 ROI 后, 系统将裁剪后的舌体区域输入 YOLOv26n 分类模型。考虑到舌象图像中可能同时存在多个舌象特征, 本文对原有分类模型进行多标签化改造。具体而言, 保留 YOLOv26n 的主干特征提取结构, 将原有用于单标签分类的 Softmax 输出层替换为 Sigmoid 输出层, 使模型能够对每一个舌象类别独立输出存在概率。本文设置的舌象类别如表 1 所示。

Table 1. Tongue image category setting

表 1. 舌象类别设置

类别编号	类别名称	类别类型
0	正常舌	正常类
1	胖大舌	形态类
2	裂纹舌	形态类
3	薄白苔	舌苔类
4	黄厚苔	舌苔类
5	白厚苔	舌苔类
6	剥落苔	舌苔类
7	红点舌	舌色类
8	淡白舌	舌色类
9	淡红舌	舌色类
10	黄厚腻苔	舌苔类
11	黑苔	舌色/舌苔类
12	地图舌	形态类
13	齿痕舌	形态类

设舌象类别数为 C , 第 i 张舌象图像的真实标签表示为:

$$y_i = [y_{i1}, y_{i2}, \dots, y_{iC}], y_{ic} \in \{0, 1\} \quad (3)$$

其中, $y_{ic} = 1$ 表示第 i 张图像存在第 c 类舌象特征, $y_{ic} = 0$ 表示不存在该特征。模型输出 logits 表示为:

$$z_i = [z_{i1}, z_{i2}, \dots, z_{iC}] \quad (4)$$

经过 Sigmoid 函数后得到每一类舌象特征的预测概率:

$$p_{ic} = \sigma(z_{ic}) = 1 + \frac{1}{e^{-z_{ic}}} \quad (5)$$

3.4. 类别不平衡处理与损失函数设计

由于不同舌象类别在数据集中出现频率不同, 部分舌象类别样本数量较少, 容易导致模型训练过程中偏向高频类别。为缓解类别不平衡问题, 本文在多标签分类训练中采用加权二元交叉熵损失函数。

对于第 i 个样本和第 c 个类别, 其损失定义为:

$$L_{ic} = -[w_c y_{ic} \log(p_{ic}) + (1 - y_{ic}) \log(1 - p_{ic})] \quad (6)$$

整体损失函数为:

$$L = -\frac{1}{NC} \sum_{i=1}^N \sum_{c=1}^C [w_c y_{ic} \log(p_{ic}) + (1 - y_{ic}) \log(1 - p_{ic})] \quad (7)$$

其中, N 表示样本数量, C 表示类别数量, p_{ic} 表示模型对第 c 类的预测概率, y_{ic} 表示真实标签, w_c 表示第 c 类正样本权重。本文根据训练集中每一类正负样本比例设置权重:

$$w_c = \frac{N_c^-}{N_c^+} \quad (8)$$

其中, N_c^+ 表示第 c 类正样本数量, N_c^- 表示第 c 类负样本数量。对于样本较少的类别, 较大的类别权重能够提高模型对该类特征的关注程度, 从而缓解类别不平衡造成的少数类识别能力不足问题。

同时, 本文在训练阶段对少数类样本进行适度数据增强, 包括随机亮度变化、颜色扰动、随机旋转、尺度缩放和水平翻转等, 以提高模型对不同采集条件下舌象图像的适应能力。

3.5. 多标签舌象识别评价指标

为适应舌象图像中多种特征共存的特点, 本文采用 mAP、Micro-F1、Macro-F1 和 Hamming Loss 作为多标签分类模型的主要评价指标。

mAP 用于衡量模型在所有类别上的综合识别能力, 其定义为:

$$\text{mAP} = \frac{1}{C} \sum_{c=1}^C AP_c \quad (9)$$

其中, AP_c 表示第 c 类舌象特征的平均精度, C 表示类别数量。

Micro-F1 从整体层面统计所有类别的真正例、假正例和假负例, 适合评价模型整体识别效果:

$$\text{Micro-F1} = \frac{2TP}{2TP + FP + FN} \quad (10)$$

Macro-F1 先分别计算每个类别的 F1 值, 再对所有类别求平均, 更能反映模型对少数类别的识别能力:

$$\text{Macro-F1} = \frac{1}{C} \sum_{c=1}^C F1_c \quad (11)$$

Hamming Loss 用于衡量多标签预测结果与真实标签之间的不一致程度:

$$\text{Hamming Loss} = \frac{1}{NC} \sum_{i=1}^N \sum_{c=1}^C I(\hat{y}_{ic} \neq y_{ic}) \quad (12)$$

其中, $I()$ 为指示函数。Hamming Loss 越低, 说明模型多标签预测错误越少。

3.6. 几何形态学辅助验证方法

针对齿痕舌等边缘形态异常特征，本文设计了几何形态学辅助验证方法[5]。该方法主要用于对舌体边缘凹陷程度进行量化分析。流程如图 2 所示。具体步骤如下：

首先，提取舌体 ROI 图像的红色通道。由于舌体区域在红色通道上通常具有较明显灰度响应，该步骤有助于增强舌体与背景之间的差异。

其次，对红色通道图像进行高斯滤波，降低局部噪声和纹理干扰。随后使用 Otsu 阈值分割方法获得二值图像，并通过形态学闭运算填补舌体区域内部小孔洞，使舌体轮廓更加连续。

然后，提取二值图像中的最大外轮廓，作为舌体主轮廓。对该轮廓进行多边形近似，并计算其凸包和凸缺陷。对于每一个凸缺陷点，系统根据缺陷深度、位置和跨度进行筛选，仅保留位于舌体左右边缘、深度较大且跨度适中的有效凹陷点。

最后，定义边缘凹陷得分 S ：

$$S = \frac{D}{L} \quad (13)$$

其中， D 表示有效凸缺陷深度总和， L 表示舌体轮廓周长。当 S 大于设定阈值时，认为舌体边缘存在较明显凹陷特征。该得分作为齿痕类形态异常的辅助判断依据。

几何形态学辅助验证流程图



Figure 2. Flowchart of geometric morphology-assisted verification
图 2. 几何形态学辅助验证流程图

3.7. 深度学习与几何形态特征融合策略

分类模型和几何形态学模块之间并不是简单的替代关系。分类模型能够从舌体 ROI 中提取颜色、纹

理和语义特征，适合识别舌色和舌苔等整体视觉特征；几何形态学模块则重点关注舌体边缘凹陷情况，能够为齿痕舌等形态异常提供可解释的几何依据。为避免仅依赖人工规则判断，本文设计了基于概率校正的融合策略。

设多标签分类模型输出的齿痕舌概率为 p_{sc} ，几何形态学模块输出的边缘凹陷得分为 S 。首先将分类模型输出概率转换为 logit 形式：

$$\text{logit}(p_{sc}) = \log \frac{p_{sc}}{1 - p_{sc}} \quad (14)$$

然后利用几何凹陷得分对齿痕舌概率进行校正：

$$p'_{sc} = \sigma(\text{logit}(p_{sc}) + \beta S) \quad (15)$$

其中， p'_{sc} 表示融合后的齿痕舌概率， β 表示几何得分的融合系数。本文在验证集上搜索 β 的最优取值，使齿痕舌类别的 F1 值达到最大。通过该方式，当舌体边缘凹陷明显时，系统会适当提高齿痕舌的输出概率；当几何凹陷不明显时，则主要保留分类模型的判断结果。

进一步地，为提高融合策略的自适应能力，本文还将几何形态学模块输出的多个特征组成几何特征向量：

$$g = [S, N_d, \bar{D}, D_{\max}, R_{lr}] \quad (16)$$

其中， S 表示边缘凹陷得分， N_d 表示有效凹陷点数量， $\bar{D}$ 表示平均凹陷深度， $D_{\max}$ 表示最大凹陷深度， R_{lr} 表示左右边缘凹陷分布差异。将该几何特征向量与分类模型倒数第二层输出的深度特征 f 进行拼接：

$$h = [f; g] \quad (17)$$

并输入轻量级全连接网络进行最终多标签预测：

$$p = \sigma(MLP(h)) \quad (18)$$

该方法能够使模型同时利用深度学习特征和几何形态特征，相比简单规则判断具有更强的自适应能力。

4. 实验设计与结果分析

4.1. 数据集与实验设置

本文实验数据集为自建中医舌象图像数据集，包含正常舌、胖大舌、裂纹舌、薄白苔、黄厚苔、白厚苔、剥落苔、红点舌、淡白舌、淡红舌、黄厚腻苔、黑苔、地图舌和齿痕舌等多类舌象特征。数据主要通过摄像头和移动设备在自然室内光照及辅助补光条件下采集，采集过程中尽量保持舌体位于图像中心区域，并要求被采集者自然伸舌，以减少姿态变化对后续识别的影响。

为符合学术伦理和隐私保护要求，数据采集过程中不记录姓名、身份证号、联系方式等个人身份信息。对于包含面部局部区域的原始图像，仅用于舌体检测模型训练；在分类模型训练和论文结果展示中，主要使用裁剪后的舌体 ROI 图像，以降低无关身份信息泄露风险。

数据标注采用多标签标注方式。每张舌象图像可以同时标注多个舌象特征，例如“淡白舌”“白厚苔”和“齿痕舌”可以共同出现在同一张图像中。标注过程按照舌色、舌苔和舌体形态三个维度进行。首先由标注人员根据预先制定的类别标准完成初始标注；随后对存在争议的样本进行复核，并通过讨论确定最终标签。为保证训练、验证和测试结果的可靠性，数据集按照 70%、15% 和 15% 的比例划分为训练集、验证集和测试集，划分结果见表 2。

Table 2. Dataset partition
表 2. 数据集划分情况

数据集	图像数量	占比
训练集	1680	70%
验证集	360	15%
测试集	360	15%
合计	2400	100%

实验中，检测模型用于舌体区域定位，分类模型用于 ROI 特征识别。各类别正样本数量统计见表 3。训练过程中采用数据增强策略，包括图像缩放、随机翻转、亮度调整和颜色扰动等，以提高模型对不同采集条件的适应能力。

Table 3. Statistics of positive samples for each category in the tongue image dataset
表 3. 舌象数据集各类别正样本数量统计

类别编号	类别名称	类别类型	正样本数量	占比/%
0	正常舌	正常类	360	15.00
1	胖大舌	形态类	238	9.92
2	裂纹舌	形态类	286	11.92
3	薄白苔	舌苔类	432	18.00
4	黄厚苔	舌苔类	258	10.75
5	白厚苔	舌苔类	376	15.67
6	剥落苔	舌苔类	164	6.83
7	红点舌	舌色类	214	8.92
8	淡白舌	舌色类	302	12.58
9	淡红舌	舌色类	518	21.58
10	黄厚腻苔	舌苔类	226	9.42
11	黑苔	舌色类	96	4.00
12	地图舌	形态类	128	5.33
13	齿痕舌	形态类	344	14.33

4.2. 舌体检测模型结果分析

舌体检测模型用于从原始图像中定位舌体区域。本文采用 Precision、Recall、mAP@0.5 和 mAP@0.5:0.95 作为检测性能评价指标[6]。其中，Precision 表示预测为舌体的目标中实际为舌体的比例，Recall 表示实际舌体目标被正确检测出的比例，mAP 用于综合评价不同置信度和 IoU 阈值下的检测性能，具体结果见表 4。

Table 4. Performance of tongue detection model
表 4. 舌体检测模型性能

模型	Precision/%	Recall/%	mAP@0.5/%	mAP@0.5:0.95/%
YOLOv26n-det	94.6	92.8	96.1	71.4

从检测结果看，YOLOv26n 检测模型在舌体定位任务中取得了较好的检测效果。其中，mAP@0.5 达到 96.1%，说明模型能够较准确地定位舌体区域；mAP@0.5:0.95 为 71.4%，表明在更严格的 IoU 评价条件下仍具有一定的定位稳定性。相较于直接整图分类，舌体检测能够减少非舌体区域干扰，使分类模型更加关注舌象特征。

4.3. 舌象分类模型结果分析

舌象分类模型用于对舌体 ROI 中的舌色、舌苔和形态特征进行多标签识别。为验证多标签建模、类别权重和几何特征融合的有效性，本文设置单标签基线、多标签基线、加权 BCE 多标签模型以及几何融合模型进行对比，结果如表 5 所示。

Table 5. Performance comparison of multi-label tongue image recognition models
表 5. 多标签舌象识别模型性能对比

模型	损失函数	不平衡处理	mAP/%	Micro-F1/%	Macro-F1/%	Hamming Loss
YOLOv26n 单标签基线	CE Loss	无	83.6	80.2	76.8	0.086
YOLOv26n 多标签基线	BCE Loss	无	86.9	82.7	79.5	0.071
YOLOv26n 多标签模型	加权 BCE Loss	类别权重	88.4	84.1	82.3	0.062
YOLOv26n + 几何融合	加权 BCE Loss	类别权重 + 几何特征	89.7	85.6	83.9	0.054

从表中可以看出，相较于单标签分类基线，多标签模型能够更好地适应舌象图像中多种特征共存的情况。采用 BCE Loss 后，模型可以对每个舌象类别独立输出预测概率，避免了 Softmax 单标签分类中类别互斥假设带来的限制。在此基础上，引入类别权重后，模型对少数类舌象特征的识别能力得到提升，Macro-F1 相比未加权模型有所改善。进一步融合几何形态特征后，模型对齿痕舌等边缘形态异常类别的识别更加稳定，说明几何特征能够为深度学习模型提供有效补充。

4.4. 几何形态辅助分析结果

为验证几何形态学辅助分析方法的有效性，本文对部分具有边缘凹陷特征的舌象图像进行轮廓分析。系统通过凸包与凸缺陷检测得到舌体边缘凹陷点，并计算边缘凹陷得分[7]，分析示例见表 6。

Table 6. Examples of geometric morphology-assisted analysis
表 6. 几何形态辅助分析示例

样本编号	分类预测结果	是否包含齿痕舌	边缘凹陷得分	辅助判断
1	齿痕舌	是	0.021	齿痕特征明显
2	淡白舌	是	0.018	疑似齿痕特征
3	裂纹舌	否	0.009	未触发齿痕提示

从分析结果可以看出，几何形态学模块能够为齿痕类边缘异常提供量化依据。当分类模型结果与几何得分一致时，系统判断更加可靠；当分类模型置信度较低而几何得分较高时，几何模块可作为辅助提示，降低漏检风险。

4.5. 边缘端部署性能测试

为验证所提出方法在嵌入式平台上的应用可行性，本文将训练后的模型部署至树莓派 5 平台进行测试，测试结果见表 7。边缘端测试主要关注推理耗时、帧率、模型大小和系统稳定性。

Table 7. Edge-side inference performance
表 7. 边缘端推理性能

平台	模型	输入尺寸	推理耗时/ms	FPS	模型大小/MB
Raspberry Pi 5	YOLOv26n-det	640 × 640	286.4	3.5	5.8
Raspberry Pi 5	YOLOv26n-cls	224 × 224	64.7	15.4	4.2

实验结果表明，YOLOv26n 模型具备一定轻量化特点，可在树莓派 5 端完成本地化推理。其中，检测模型平均推理耗时为 286.4 ms，分类模型平均推理耗时为 64.7 ms，能够满足单次拍照识别和辅助分析场景的基本需求。由于检测模型输入尺寸较大，其推理耗时明显高于分类模型。后续可通过模型剪枝、量化、ONNX Runtime 或 NCNN 等推理引擎进一步提升树莓派端推理速度[8]。

4.6. 实验结果

为进一步验证本文方法在实际应用场景中的可行性，本文基于所设计的中医舌象多模态问诊终端进行了功能展示实验。实验界面如图 3 所示。系统启动后进入实时监测界面，主界面上方显示系统名称，中部为摄像头实时预览区域，下方设置开启实时监测、停止摄像头、抓拍并出报告、上传照片检测和退出系统等功能按钮。该界面能够完成舌象图像实时采集、图片上传检测和静态报告生成等操作，为后续舌象识别实验提供可视化交互环境。

从图 3 可以看出，系统界面结构清晰，左侧大区域用于显示摄像头实时画面，底部功能按钮能够支持用户完成图像采集与检测操作。该界面为本文提出的舌象识别方法提供了实际运行载体，使模型推理结果能够以可视化方式直接反馈给用户。



Figure 3. Real-time monitoring interface of the system
图 3. 系统实时监测界面

在舌象识别实验中，用户通过摄像头采集舌象图像，并点击“抓拍并出报告”按钮进行检测。检测

结果如图 4 所示。左侧区域显示原始舌象图像，右侧区域显示 AI 综合异常分析报告。系统首先利用 YOLOv26n 检测模型定位舌体区域，并在右侧结果图中以绿色矩形框标注舌体位置。随后将舌体 ROI 输入 YOLOv26n 多标签分类模型进行舌象特征识别。该样本中，“淡红舌”类别输出概率为 97.0%，高于设定阈值，因此系统将其作为主要舌象特征显示；其余舌苔类和形态类异常特征未超过对应阈值。同时，几何形态学模块计算得到的边缘凹陷得分低于齿痕舌判断阈值，因此系统在报告中给出“未见明显齿痕”的辅助提示。



Figure 4. Display of tongue image recognition result and auxiliary report

图 4. 舌象识别结果与辅助报告展示

由图 4 可见，本文方法能够较准确地定位舌体区域，检测框基本覆盖舌体主体区域，说明 YOLOv26n 检测模型能够为后续 ROI 分类提供有效输入。在分类结果方面，系统输出“淡红舌”作为核心状态，置信度达到 97.0%，表明分类模型对该样本具有较高的预测置信度。同时，系统在报告中给出“未见明显齿痕”的形态辅助判断，说明几何形态学模块在该样本中未检测到明显的舌体边缘凹陷特征。

此外，系统还显示了单次推理延迟，实验样例中的完整推理延迟约为 1290.8 ms。该耗时包括图像采集、舌体检测、ROI 裁剪、舌象分类、几何形态学辅助分析以及结果显示等多个环节。由于本文方法面向拍照式舌象辅助分析场景，而非连续高帧率视频检测场景，因此该推理速度能够满足单次舌象识别与辅助报告生成的基本需求。后续可通过模型量化、输入尺寸优化、ONNX Runtime 或 NCNN 等端侧推理加速方式进一步降低推理耗时。

5. 讨论

本文方法的核心优势在于将舌体检测和多标签舌象识别分离处理，形成由粗到细的识别流程。与直接整图分类相比，检测模型能够先定位舌体区域，减少背景、面部和嘴唇区域干扰；多标签分类模型则在舌体 ROI 上进行特征识别，能够同时判断舌色、舌苔和舌体形态等多种舌象特征。

相比传统单标签分类方式，多标签建模更符合中医舌象分析的实际特点。一幅舌象图像可能同时包含多个异常特征，例如舌色偏淡、舌苔较厚以及舌体边缘存在齿痕。如果采用 Softmax 单标签分类，模型只能输出一个最高概率类别，容易造成其他共存特征漏检。本文采用 Sigmoid 输出和加权二元交叉熵损失函数，使各舌象类别能够独立预测，并通过 mAP、Micro-F1、Macro-F1 和 Hamming Loss 等指标对模型进行综合评价。

此外，几何形态学辅助分析提高了齿痕舌识别的可解释性。深度学习模型虽然具有较强的特征表达

能力,但其判断过程相对黑箱。通过凸包和凸缺陷检测计算边缘凹陷得分,可以为齿痕类形态异常提供直观的几何依据。本文进一步将几何得分作为辅助特征参与融合决策,而不是仅依靠固定规则进行判断,从而提高系统对复杂舌象样本的适应能力。

不过,本文方法仍存在一些不足。首先,舌象图像容易受到光照、拍摄设备和白平衡影响,后续需要引入颜色校准和标准化采集方法。其次,多标签标注依赖标注人员对舌象特征的准确判断,部分类别之间边界较模糊,未来需要进一步提高标注一致性。再次,几何形态学方法对舌体分割质量较敏感,当舌体边缘遮挡或轮廓提取不完整时,凹陷得分可能受到影响。最后,树莓派 5 端推理速度仍受硬件算力限制,后续可进一步研究模型剪枝、量化和端侧推理加速部署。

6. 结论

本文提出了一种基于 YOLOv26n 双模型协同的多标签舌象智能识别方法。该方法首先利用 YOLOv26n 检测模型定位舌体区域并裁剪 ROI,再通过改进后的 YOLOv26n 多标签分类模型识别舌色、舌苔和舌体形态等多种舌象特征。针对舌象图像中多种特征可能同时存在的问题,本文将分类任务建模为多标签识别任务,采用 Sigmoid 输出和加权二元交叉熵损失函数进行训练,并使用 mAP、Micro-F1、Macro-F1 和 Hamming Loss 等指标进行评价。针对齿痕舌等边缘形态异常,本文进一步引入几何形态学辅助分析方法,通过凸包和凸缺陷检测计算舌体边缘凹陷得分,并将其作为辅助特征与深度学习模型结果进行融合。实验结果表明,该方法能够实现舌体检测、ROI 提取、多标签舌象识别、形态辅助验证和边缘端部署验证。未来将进一步扩充数据集规模,优化多标签标注体系,提升模型轻量化部署能力,并结合颜色校准和多模态健康数据提高舌象智能识别的准确性和实用性。

致 谢

感谢指导教师在研究思路、实验设计和论文撰写过程中给予的指导与帮助。感谢团队成员在数据整理、模型训练、系统测试和边缘端部署过程中的支持。感谢相关开源社区提供的计算机视觉与深度学习工具,为本文实验实现提供了重要基础。

基金项目

本研究得到陕西省大学生创新创业训练计划项目(项目编号: S202612715083): 守望——阿尔茨海默症早期筛查与家庭支持社区网络。

参考文献

- [1] 邵延华,郭涛,王子明,等. 基于深度学习的 YOLO 目标检测综述[J]. 电子与信息学报, 2022, 44(10): 3697-3708.
- [2] 卢运西,张涛,杨志明,等. 中医舌象分割技术研究进展: 方法、性能与展望[J]. 自动化学报, 2021, 47(9): 2033-2050.
- [3] 宋超,刘鹏,王晓东,等. 基于深度迁移学习的舌象特征分类方法研究[J]. 计算机工程与科学, 2021, 43(8): 1488-1495.
- [4] 卓力,黄建,刘芳,等. 有噪声标注情况下的中医舌色分类方法[J]. 电子与信息学报, 2022, 44(9): 3171-3179.
- [5] 张勇,王忆勤,陈家旭,等. 中医舌诊客观化研究进展[J]. 世界科学技术-中医药现代化, 2020, 22(12): 4356-4364.
- [6] 章毓晋. 图像处理和分析[M]. 北京: 清华大学出版社, 2018.
- [7] 陈积明,张宁,李勇,等. 边缘计算研究进展与展望[J]. 计算机研究与发展, 2019, 56(1): 69-89.
- [8] 王田苗,陶永,李硕,等. 人工智能在医学图像分析中的应用研究综述[J]. 中国图象图形学报, 2021, 26(1): 1-18.