

基于深度学习的舌象特征研究*

崔涛¹ 何佳俊¹ 何华² 李睿¹ 赵亮³ 垢德双¹

(¹ 天津市医疗器械审评查验中心 天津 300191 ² 北京神农智能技术有限公司 北京 102208

³ 慧医谷中医药科技(天津)股份有限公司 天津 300392)

〔摘要〕 目的/意义 应用深度学习技术达到舌象分析自动化的目的,为中医舌象标准化提供参考依据,进一步推进中医诊疗技术现代化进程。方法/过程 构建以区域关联性为基础、引入标签松弛技术的语义分割损失函数,显式约束舌象分割模型学习局部区域各像素关联性的同时,对错误标签具有一定的容错能力;从舌象特征中内含的颜色相关性潜在先验出发,在模型构建阶段将舌象特征仅解耦为两类下游多标签分类任务,在加速模型拟合的同时有效降低模型复杂度。结果/结论 在自建数据集上验证算法有效性,舌象分割 MIoU 指标为 96.57%,舌象分析宏 F1 值、平均准确率分别为 88.58%、82.59%。

〔关键词〕 迁移学习;舌象特征;深度学习;舌象分割;中医舌诊

〔中图分类号〕 R-058 〔文献标识码〕 A 〔DOI〕 10.3969/j.issn.1673-6036.2024.07.014

Study on Tongue Image Features Based on Deep Learning

CUI Tao¹, HE Jiajun¹, HE Hua², LI Rui¹, ZHAO Liang³, KOU Deshuang¹

¹Tianjin Medical Device Evaluation and Inspection Center, Tianjin 300191, China; ²Beijing Shennong Intelligent Technology Co. Ltd., Beijing 102208, China; ³Medvalley TCM Technology (Tianjin) Co. Ltd., Tianjin 300392, China

〔Abstract〕 **Purpose/Significance** To apply deep learning technology to achieve the purpose of tongue image analysis automation, so as to provide references for the standardization of tongue image of traditional Chinese medicine (TCM), and further promote the modernization of TCM diagnosis and treatment technology. **Method/Process** It develops a new semantic segmentation loss function with region-based correlation and label relaxation to enhance the capability of tongue image segmentation model to learn pixel relationships and handle mislabeled data. Additionally, leveraging inherent color-related priors in tongue image features, the model is simplified by decomposing it into two multi-label classification tasks, thereby accelerating model training and reducing its complexity. **Result/Conclusion** The proposed algorithm is proven effective on a self-constructed dataset, attaining a high 96.57% MIoU in tongue segmentation, and demonstrating strong performance with a macro F1-score of 88.58% and average accuracy of 82.59%.

〔Keywords〕 transfer learning; tongue image feature; deep learning; tongue segmentation; tongue diagnosis of traditional Chinese medicine (TCM)

〔修回日期〕 2024-04-02

〔作者简介〕 崔涛,工程师,发表论文 4 篇;通信作者:垢德双,高级工程师。

〔基金项目〕 国家重点研发计划(项目编号:2023YFC3606200)。

1 引言

近年来许多学者关注舌诊自动化研究,如舌象预处理、采集、舌体分割及舌象特征定量化等。在舌象分割研究中,Zhang L 等^[1]提出基于自适应阈

值的舌象分割方法，Huang Z P 等^[2]和张希颖^[3]分别使用分水岭、区域生长等以区域为基础的分割方法。这些基于传统机器学习的方法随着深度学习的发展逐渐被舍弃，基于深度学习的舌象语义分割方法成为研究热点。Ruan Q 等^[4]构建基于 UNet 的改进舌象分割模型和边缘分割网络优化舌体边缘分割。Cai Y 等^[5]在原有分割损失函数基础上增加辅助损失函数提高语义分割性能。Wang L 等^[6]使用级联的卷积神经网络分割舌象。深度语义分割算法依赖大量人工标注的数据集，数据集质量直接影响训练效果，然而在舌象语义分割数据集的制作过程中，经常出现舌象边界难以界定导致边界标签错误的情况，因此模型需对边界有一定的容错能力，但是鲜有方法关注舌象边界标签错误问题。

在舌象分析的研究中，Wang X 等^[7]将 ResNet50 作为模型的主干网络，定位和识别裂纹与齿痕区域，并使用梯度加权激活映射可视化裂纹、齿痕区域。

刘梦等^[8]将 Faster R - CNN 和迁移学习应用于齿痕舌、裂纹舌的局部特征识别中。杨晶东等^[9]基于迁移学习的全连接神经网络小样本舌象分类方法，对正常舌、裂纹舌、厚苔、芒刺舌、齿痕舌和剥落苔等舌象特征分类。尽管与传统舌象分类方法相比，这些方法具有更高的准确度，但研究重点集中在单一舌象特征量化或者分类方面，未重视不同舌象之间的特征关联性和舌象特征的潜在先验。

本文在构建语义分割数据集时发现，舌象数据集制作中存在不可避免的边界标注错误。因此提出一种基于区域关联性和边界标签松弛的损失函数，以显式地指导模型在学习局部邻域像素间相关性的同时具有一定的标注容错能力，有效提升舌象分割精度；同时从舌象特征中内含的颜色相关潜在先验出发，将舌象特征解耦为两个下游多标签分类任务，在加速模型拟合的同时有效降低模型复杂度，具体舌象特征识别架构，见图 1。

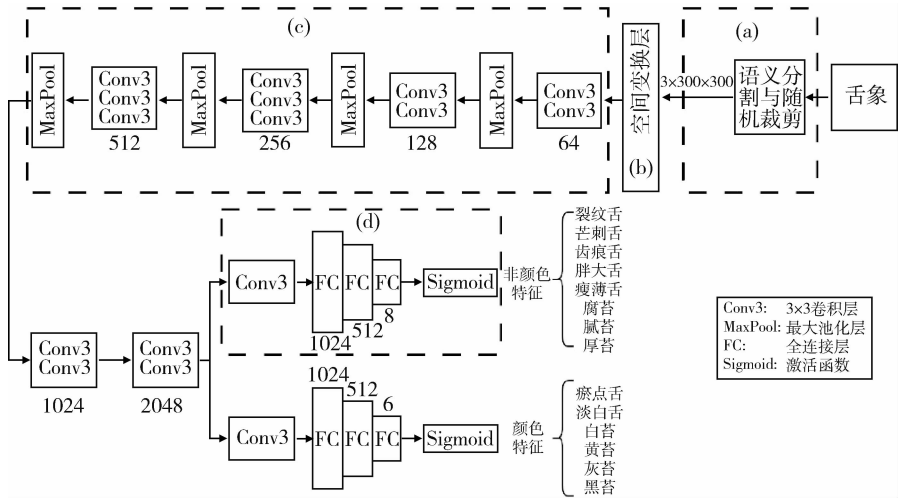


图 1 舌象特征识别整体构架

2 相关工作

现阶段，随着信息技术持续更新及发展，提升舌诊客观性已成为舌诊自动化研究的主要方向^[10]。舌诊自动化流程分为 4 步：数据采集、舌象分割、舌象分析、四诊合参^[11]。

2.1 舌体分割

现阶段舌体分割多在对传统图像处理予以改进的基础上完成。Zhang L 等^[1]首先通过灰度投影定位舌体，然后通过最大类间方差法分割。Huang Z P 等^[2]提出以子块生长与最大类间方差法为基础的舌体自动分割算法。高丽等^[12]提出以改进模糊算子与

形态学为基础的舌体分割方法。Li Q L 等^[13] 基于光谱角度匹配算法将超光谱图像转换为光谱角度立方体后选取一维脉冲波形密切监测边缘信息，得到基于超光谱图像的舌体分割方法。Zhou C 等^[14] 提出使用特征金字塔网络提取特征，构建检测分割多任务学习网络。Huang X 等^[15] 使用深度残差网络、特征金字塔网络分别作为编码器和解码器，融合多尺度特征实现舌象分割。

2.2 舌象分析

舌诊自动化的重点在于自动分析舌图明确舌象特征及病变。对舌象的分析实质上是识别舌体视觉特征。瞿婷婷等^[16] 提出基于改进 Gabor 小波变换的舌苔腐腻识别方法。Wang S 等^[17] 提出基于传统视觉的舌象斑点识别与提取方法。刘斌等^[18] 通过小波与灰度共生矩阵相结合方式对舌苔厚度予以自动提取，从而量化舌苔厚度。杨晶东等^[9] 基于全连接神经网络迁移学习对小样本舌象特征分类。

3 舌体语义分割模型构建

分割舌体是确保分析精准性的前提条件，常规图像处理舌体分割的精准性较低，且易受开放环境影响，因此应重视舌体分割方法的合理选择。语义分割算法输出为像素的边缘概率分布，且损失以像素的负对数损失为基础，不能确保模型学习到像素间的关联，空间一致性不足^[19]。同时舌象语义分割数据集的制作过程中，经常出现舌象边界难以界定

从而出现边界标签错误的情况，因此需要模型对边界有一定的容错能力。

损失函数的设计从两个方面出发，对于非边界像素考虑其邻域像素间的关联性，而边界像素则从标签松弛出发。针对区域相关性，李渊彤等^[20] 基于局部关联性给单像素的损失分配不同的权重，同时额外定义基于颜色的损失，考虑区域像素之间关联关系的同时有效利用像素标签语义的监督信息。

$$w_i = 1 + \sum_{j \in K} \text{label_sim}_i (1 + \text{color_sim}_j + \text{pred_prob}_j)$$
$$\text{color_loss}_i = - \sum_{j \in K} \log(\text{pred_sim}_j + \varepsilon) \times (\text{color_sim}_j - \beta) \times \text{label_sim}_j$$
$$L_{\text{LC},i} = -w_i \times \log(p_i) + \alpha \times \text{color_loss}_i \tag{1}$$

其中，pred_prob_j 表示邻域像素被预测为正确标签率，color_sim_j 表示 RGB 空间内目标像素*i* 与邻域像素*j* 的余弦相似度，label_sim_i 指目标像素*i* 与邻域像素*j* 相等与否，K 表示大小为 3×3 的邻域窗口，α 为可调节参数，采用 0.1。其在像素标签语义监督信息的基础上充分利用区域像素关联性，然而其在边界像素标签错误时表现不佳，因此引入标签松弛技术^[21] 缓解边界标签错误情况，使模型具有一定标签容错能力。

$$L_{\text{boundary}} = -\log \sum_{i \in N} P(C) \tag{2}$$

其中 *N* 为 3×3 大小的松弛窗口内像素标签的集合，P(C) 为类别的 softmax 概率，*i* 表示某一像素点的标签，错标边界像素点在交叉熵损失函数下会贡献大量损失代价，不易于模型的正确收敛，而标签松弛处理则可以提升模型对这类训练数据的容错能力，易误标的舌象边界区域及其标签松弛前后，见图 2。

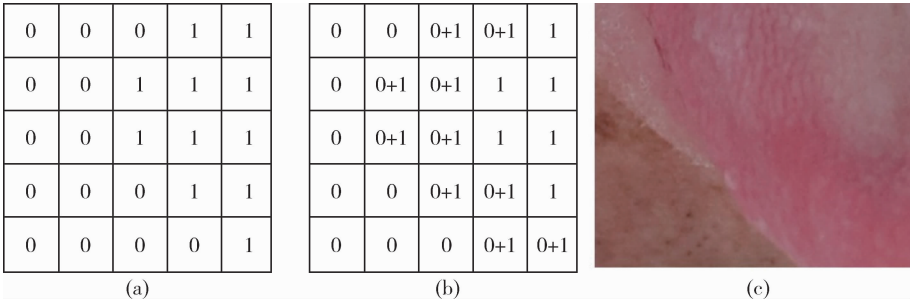


图2 松弛前后像素标签

注：0 表示背景，1 表示舌体；(a) 表示标签松弛前，(b) 表示标签松弛后，(c) 表示舌象易误标的边界区域。

最终损失函数为改进局部关联损失（improved local correlation loss, ILCLoss），调节参数采用 0.5：

$$L_{ILC} = L_{LC} + \lambda L_{boundary}$$

(3)

4 舌象分析模型构建

本文舌象分析整体架构由4部分组成，其中前3部分与文献^[20]一致，而第4部分从颜色是否相关角度出发，由两个多标签分类头组成，极大降低计算复杂度。第1部分，语义分割层对舌象进行像素级分类，获得舌体分割掩码，再以分割掩码的最小外围框为中心，随机扰动扩大最小框，最后将其缩放到固定尺寸获得舌象分析网络的输入（3×300×300）。第2部分，空间变换层（来自空间转换网络^[22]），纠正数据空间增广（镜像、旋转及平移等）带来的图像目标空间位置扰动现象，还可解决采集过程中舌体所处部位、姿势不相同问题。第3部分，语义共享层，由语义分割层模型的主干网络VGG16和两个卷积块组成，在两个分类头间共享语义，在降低计算量的同时挖掘更深层特征空间的共同语义特征。第4部分，任务语义层，由卷积块、全连接块、Sigmoid 激活函数组成，负责挖掘不同下游任务的特定语义。对于舌象特征分析，本文从颜色是否相关角度特征先验出发，将舌象特征分为两类：与颜色高度相关的特征（瘀点舌以及苔色）；与颜色不相关的特征（裂纹舌、芒刺舌、齿痕舌、胖大舌、瘦薄舌、腐苔、腻苔、厚苔）。

5 数据集收集与整理

5.1 数据集 A 舌象语义分割数据集

舌象语义分割数据集来自 Github 开源数据集（<https://github.com/BioHit/TongueImageDataset>），由标准舌象采集装置完成，数据集内包含分辨率为768×76的舌象300张，及其对应的分割掩码图像，见图3。



图3 舌象语义分割数据集示例

5.2 数据集 B 舌象特征数据集

舌象特征数据采用慧医谷中医药科技（天津）股份有限公司的便携中医全诊仪设备随机采集264例对象，其中女性160例，男性104例，共264张舌图，见图4。初始分辨率为2760×1840，经过数据增强处理得到528张，舌体提取以及随机扰动后缩放为固定分辨率300×300，数据分布，见表1。

表1 舌象特征数据集部分数据分布占比

类别		占比（%）
苔色	淡白	89
	白	6
	黄	2
	灰	2
	黑	1
腐腻苔	正常	48
	腻	37
	腐	15
...	...	...
裂纹舌	裂纹	86



图4 舌象特征数据集示例

6 实验设计及结果评估

6.1 舌象分割对比实验

在数据集A实验，共300张图片，其中训练集200张，测试集50张，验证集50张。评估标准：对类求IoU加权平均值的频权交并比（frequency weighted intersection – over – union, FWIoU）；预测值和真实值两个集合的交集与并集之比（mean intersection – over – union, MIoU）；计算总像素中每个类中被正确分类的像素数占比，在此基础上获取全部

类的平均像素精度（mean pixel accuracy, MPA）；正确标记的像素占总像素比例的像素精度（pixel accuracy, PA）。其中，MIoU 是最常见的语义分割评价准则，也是本文模型选择的依据。

基于 FCN-8s、改进 FCN-16s、FCN-32s、U-Net 4 种网络架构对负对数损失（negative logarithm, NLLoss）、局部损失（FocalLoss）、局部关联损失（local correlation loss, LCLoss）、改进局部关联损失（ILCLoss，本文方法）4 种损失进行对比实验，其中优化器为 Adam，最大学习率为 $2e-4$ ，学习策略为 1-cycle，训练步长为 40，实验结果，见表 2。在 MIoU 指标下，本文提出的改进局部关联损失在 FCN-16s 上表现性能最优，其他模型结构上

与最优损失 LCLoss 的表现基本持平。虽然对 FCN 系列模型的提升幅度不大，但对 UNet 模型有显著提升。相比最优的损失 LCLoss，MIou 指标仍有接近 3 个百分点的提升，其他指标的提升也在 2 个百分点左右，说明改进局部关联损失确实可以更好地指导模型学习，分析 FCN 系列提升不明显的原因可能为其初始权重迁移自大规模数据集的训练权重，损失函数对学习指导作用有限。同时各项指标相较于 LCLoss 均有改善，说明标签松弛技术在邻域信息关联的基础上进一步挖掘了模型的鉴别能力，提升了模型对错误标签的容忍力。综合实验结果，本文选用 FCN-16s 模型的编码器 VGG16 作为舌象分析的主干网络。

表 2 不同模型在不同损失下的性能评估

评价指标	损失	FCN-8s	FCN-16s	FCN-32s	U-Net
频权交并比	负对数损失	0.963 7	0.966 2	0.958 3	0.854 8
	局部损失	0.965 4	0.968 1	0.962 2	0.881 6
	局部关联损失	0.972 2	0.972 9	0.963 7	0.891 9
	本文方法	0.971 1	0.981 2	0.954 4	0.903 5
均交并比	负对数损失	0.949 3	0.952 9	0.941 7	0.793 0
	局部损失	0.951 8	0.955 5	0.947 2	0.837 6
	局部关联损失	0.961 3	0.962 1	0.949 3	0.849 5
	本文方法	0.961 0	0.965 7	0.954 9	0.877 0
平均像素精度	负对数损失	0.980 6	0.981 9	0.976 6	0.855 4
	局部损失	0.983 6	0.983 4	0.977 7	0.921 0
	局部关联损失	0.986 2	0.984 0	0.980 6	0.916 5
	本文方法	0.985 7	0.983 6	0.973 1	0.927 7
像素精度	负对数损失	0.980 7	0.982 1	0.977 8	0.921 2
	局部损失	0.981 6	0.983 1	0.980 0	0.934 3
	局部关联损失	0.985 2	0.985 6	0.980 7	0.941 1
	本文方法	0.981 2	0.991 3	0.981 1	0.959 8

6.2 舌象分析对比实验

采用 $\gamma_- = 4$ ， $\gamma_+ = 4$ 和 $m = 0.05$ 的非对称损失^[23]（asymmetric loss, ASL），处理数据中正负样本不平衡问题。优化器为 Adam，最大学习率为 $2e-4$ ，学习策略为 1-cycle，训练步长为 20。同

时为了对数据集 B 予以充分利用，在数据集 B 选取 400 张舌图数据当作训练集，以 5 折交叉验证方式为训练、验证方法，剩余 128 张舌图为测试集，予以实验对比。评估指标包括平均准确率（ Acc_{avg} ）、F1 分数（ $F1_{macro}$ ）、宏查全率（ R_{macro} ）、宏查准率（ P_{macro} ），计算方式如下。其中： TP_i 是

指5折交叉验证中当前折中预测为正样本的正样本数量； FN_i 是指5折交叉验证中当前折中预测为正样本的负样本数量； TN_i 是指5折交叉验证中当前折中预测为负样本的负样本数量； FN_i 是指5折交叉验证中当前折中预测为负样本的正样本数量； K 等于5。

$$Acc_{avg} = \sum_{i=1}^K \frac{TP_i + TN_i}{TN_i + TP_i + FN_i + FP_i}$$

(4)

$$R_{macro} = \sum_{i=1}^K \frac{TP_i}{TP_i + FN_i}$$

(5)

$$P_{macro} = \sum_{i=1}^K \frac{TP_i}{TP_i + FP_i}$$

(6)

$$F1_{macro} = \sum_{i=1}^K \frac{TP_i^2}{2TP_i^2 + TP_i \times FN_i + TP_i \times FP_i}$$

(7)

消融实验对比无知识迁移、预训练 VGG16 知识迁移、舌象分割语义知识迁移 3 种情况下舌象特征分析模型的学习能力，见表 3。知识迁移学习参与下的模型性能均比未知识迁移的性能更好，舌象语义分割的知识迁移带来的性能提升最大，分析原因在进行语义分割任务训练的过程中模型已经挖掘到一定的舌象特征。

表 3 使用不同知识迁移的 5 折交叉验证性能比较

模型	Acc _{avg}	P _{macro}	R _{macro}	F1 _{macro}
未迁移	0.816 5	0.876 5	0.836 5	0.846 5
VGG16 迁移	0.818 5	0.878 8	0.842 8	0.846 8
分割迁移	0.825 9	0.885 8	0.854 1	0.864 0

舌象分析模型对两个子任务的预测效果，见表 4。舌质齿痕、舌苔厚薄、舌质裂纹识别率均大于 85%，预测良好，但舌质胖瘦识别率、宏 F1 值分别为 79.06%、83.27%，预测效果一般，可能与胖瘦舌类别样本失衡、训练样本量少、缺少嘴巴数据信息比较等因素有关。舌质芒刺一般，舌苔腐腻、苔色与舌质瘀点效果较差，分析原因一是数据量少，二是虽然已经将颜色相关先验加入网络结构设计，但二者结构相同，而实际上二者特征挖掘难度不同。

表 4 舌象特征分析模型 5 折交叉验证性能对比

舌象特征	Acc _{avg}	P _{macro}	R _{macro}	F1 _{macro}
舌苔苔色	0.791 6	0.875 1	0.835 6	0.865 5
舌质瘀点	0.796 8	0.867 0	0.799 6	0.839 7
舌质胖瘦	0.790 6	0.864 1	0.803 6	0.832 7
舌质齿痕	0.858 7	0.911 1	0.891 1	0.885 8
舌苔厚薄	0.870 1	0.932 9	0.880 7	0.915 4
舌苔腐腻	0.772 2	0.778 4	0.821 9	0.799 5
舌质芒刺	0.815 2	0.898 3	0.912 2	0.847 1
舌质裂纹	0.911 8	0.959 2	0.887 1	0.926 3
平均值	0.825 9	0.885 8	0.854 1	0.864 0

7 结语

本文首先构建了一个新的以区域关联性为基础、引入标签松弛技术的语义分割损失函数，该函数显式约束模型学习局部区域各像素关联性的同时，对错误标签具有一定的容错能力，最后舌象语义分割模型 MIoU 指标在测试集为 96.57%。同时，提出从舌象特征中内含的是否颜色相关的潜在先验出发，在模型构建阶段将舌象特征仅解耦为两类下游多标签分类任务，在加速模型拟合的同时有效降低了模型的复杂度，并且舌象分析与舌象分割模型共用相同的解码器，便于知识迁移学习，在加速模型拟合的同时有效提升模型的性能。通过这一模型对舌象特征予以识别及提取，其宏 F1 值、平均准确率分别为 88.58%、82.59%。受限于数据采集成本，本研究数据量略有不足，因此一些特征的识别精度较差，数据不均衡现象很明显，未来考虑增加一些数据量。本研究对舌象分割模型和舌象特征进行了切实有效的综合分析，为日后临床实现舌象现代化分析奠定基础。本研究尚有不足，如分类识别不同舌象特征时，共用编码器可强化各舌象特征之间的关联性；与此同时如何保持特征的鉴别性是后续研究重点之一。

作者贡献: 崔涛负责实验设计与实施、论文撰写与修订; 何佳俊负责文献综述、论文撰写; 何华负责图表制作; 李睿负责论文修订; 赵亮负责数据收集与分析; 垢德双负责论文审核。

利益声明: 所有作者均声明不存在利益冲突。

参考文献

- 1 ZHANG L, QIN J. Tongue - image segmentation based on gray projection and threshold - adaptive method [J]. Journal of clinical rehabilitative tissue engineering research, 2010, 14 (9): 1638 - 1641.
- 2 HUANG Z P, HUANG Y S, YI F L, et al. An automatic tongue segmentation algorithm based on OTSU and region growing [J]. Lishizhen medicine and materia medica research, 2017, 8 (13): 296 - 302.
- 3 张希颖. 基于随机游走算法的舌体图像分割 [D]. 北京: 北京理工大学, 2016.
- 4 RUAN Q, WU Q, YAO J, et al. An efficient tongue segmentation model based on u - net framework [J]. International journal of pattern recognition and artificial intelligence, 2021, 35 (16): 2154035.
- 5 CAI Y, WANG T, LIU W, et al. A robust interclass and intraclass loss function for deep learning based tongue segmentation [J]. Concurrency and computation: practice and experience, 2020, 32 (22): e5849.
- 6 WANG L, TANG Y, CHEN P, et al. Two - phase convolutional neural network design for tongue segmentation [J]. Journal of image and graphics, 2018, 23 (10): 1571 - 1581.
- 7 WANG X, LIU J, WU C, et al. Artificial intelligence in tongue diagnosis: using deep convolutional neural network for recognizing unhealthy tongue with tooth - mark [EB/OL]. [2023 - 04 - 24]. <https://www.sciencedirect.com/science/article/pii/S2001037020300325>.
- 8 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.
- 9 杨晶东, 张朋. 基于迁移学习的全连接神经网络舌象分类方法 [J]. 第二军医大学学报, 2018, 39 (8): 897 - 902.
- 10 张林子, 周武, 张洪来. 自动舌诊技术的研究进展 [J]. 中国中医基础医学杂志, 2023, 29 (5): 871 - 876.
- 11 田琪, 吕琴, 任健, 等. 舌象采集及其特征分析的客观化研究进展 [J]. 中医药导报, 2023, 29 (3): 119 - 123.
- 12 高丽, 令晓明. 基于改进模糊算子和形态学的舌体分割方法 [J]. 兰州交通大学学报, 2006 (3): 89 - 91.
- 13 LI Q L, XUE Y Q, WANG J Y, et al. Automated tongue seg - mentation algorithm based on hyperspectral image [J]. Journal of infrared and millimeter waves, 2007, 7 (1): 77 - 80.
- 14 ZHOU C, FAN H, LI Z. Tonguenet: accurate localization and segmentation for tongue images using deep neural networks [EB/OL]. [2023 - 04 - 24]. <https://ieeexplore.ieee.org/abstract/document/8864990>.
- 15 HUANG X, ZHANG H, ZHUO L, et al. TISNet - enhanced fully convolutional network with encoder - decoder structure for tongue image segmentation in traditional Chinese medicine [J]. Computational and mathematical methods in medicine, 2020 (1): 6029258.
- 16 瞿婷婷, 夏春明, 王忆勤, 等. 基于 Gabor 小波变换的舌苔腐腻识别 [J]. 计算机应用与软件, 2016, 33 (10): 162 - 166.
- 17 WANG S, LIU K H, WANG L T. Tongue spots and petechiae recognition and extraction in tongue diagnosis images [J]. Computer engineering & science, 2017, 39 (6): 1126 - 1132.
- 18 刘斌, 胡广芹, 张新峰, 等. 一种改进的中医舌苔厚度自动描述方法 [J]. 北京生物医学工程, 2018, 37 (2): 157 - 163.
- 19 宋超, 王斌, 许家佗. 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 - 1496.
- 20 李渊彤, 罗裕升, 朱珍民. 基于深度学习的舌象特征分析 [J]. 计算机科学, 2020, 47 (11): 148 - 158.
- 21 张宜春, 徐鹏举, 向翔. 基于边缘预测的传统文化图案语义分割算法 [J]. 中国传媒大学学报 (自然科学版), 2022, 29 (4): 19 - 25, 56.
- 22 JADERBERG M, SIMONYAN K, ZISSERMAN A, et al. Spatial Transformer Networks [EB/OL]. [2023 - 07 - 29]. <https://doi.org/10.48550/arXiv.1506.02025>.
- 23 BEN - BARUCH E, RIDNIK T, ZAMIR N, et al. Asymmetric loss for multi - label classification [EB/OL]. [2023 - 02 - 04]. <https://doi.org/10.48550/arXiv.2009.14119>.