

DOI:10.3969/j.issn.1005-202X.2023.08.009

医学影像物理

基于改进Faster RCNN的舌部多纹理检测

冀常鹏, 杨梦晗, 代巍

辽宁工程技术大学电子与信息工程学院, 辽宁 葫芦岛 125105

【摘要】为更加高效、并行地实现舌部多纹理的识别,提出一种基于Faster RCNN的改进舌部多纹理检测方法。首先,使用可变形卷积重塑主干提取网络中的卷积层并进行可变形池化,通过实际情况调整自身形状提取目标特征,以降低漏检率;其次,引入注意力机制scSE,通过增强有意义特征提高纹理表达能力;在目标相互掺杂且目标尺度差异较大的背景下,使用加权双向特征金字塔网络进行特征融合,以提升目标检测的准确率,最后进行迁移学习。实验结果显示该方法使所有类别平均精度达到0.935,较原始Faster RCNN模型提高了0.457,说明改进后的模型能有效解决目标掺杂和多尺度差异问题,具有较高的检测精度。

【关键词】舌部多纹理;Faster RCNN;深度学习;可变形卷积;注意力机制;迁移学习;多目标特征识别

【中图分类号】R318

【文献标志码】A

【文章编号】1005-202X(2023)08-0977-08

Tongue multi-texture recognition using improved Faster RCNN

Ji Changpeng, Yang Menghan, Dai Wei

School of Electronic and Information Engineering, Liaoning Technical University, Huludao 125105, China

Abstract: A Faster RCNN based method is proposed to realize the recognition of tongue multi-texture more efficiently and in parallel. Deformable convolution is used to reshape the convolutional layer in the backbone extraction network and perform deformable pooling, in which the region shape can be adjusted according to the actual condition to extract target features, thereby reducing the missed detection rate. Then, the attention mechanism scSE is introduced to improve the texture expression ability by enhancing meaningful features. In the context of multi-target hybrid and large multi-scale differences, a weighted bidirectional feature pyramid network is used for feature fusion to improve the accuracy of target detection, and finally transfer learning is conducted. The experimental results show that the method achieves an average accuracy of 0.935 for all categories, which is 0.457 higher than the original Faster RCNN model, indicating that the improved model can effectively solve the problems of multi-target hybrid and multi-scale differences, and has high detection accuracy.

Keywords: tongue multi-texture; Faster RCNN; deep learning; deformable convolution; attention mechanism; transfer learning; multi-target feature recognition

前言

中医舌部观察是中医望诊中关键的一步,中医十分侧重观舌辨病,因为人体的舌部能令医生对病情进行动态监测。生活习惯生成的多种舌部纹理,可以直接有效地揭示人体各部位功能状态,如舌边上有齿痕——脾虚湿气重,舌苔上有裂痕——脾胃

不足,舌苔厚腻——有积食,舌头有点刺——体内有热,此类身体功能状态对临床研究治病有重要作用^[1]。随着医学与科技的快速发展,舌诊与计算机进行融合,传统医学与现代科技相辅相成,可以实现使用计算机对中医舌诊进行记录并分析^[2]。目前机器学习算法已被广泛运用在舌诊方面^[3],利用图像处理技术对收集到的舌部图像进行特征提取^[4],并对其颜色与纹理等多个标签实现了可视化与数据化处理。文献[5]将多特征图像检索算法应用到舌部图像分析中,提出了20个特征值作为图像内容的特征向量。文献[6]是增强裂纹舌图像的假颜色变换程度,并对裂纹进行提取。文献[7]将舌图像划分为固定格式的特征块,辨别不同纹理特征的同时提出特征算法,

【收稿日期】2023-03-29

【基金项目】辽宁省教育厅基本科研项目(LJKMZ20220677)

【作者简介】冀常鹏,硕士,教授,研究方向:信号检测与估计、计算机通信与网络,E-mail: ccp@lntu.edu.com

【通信作者】代巍,博士,讲师,研究方向:微弱信号检测与信息处理,E-mail: daiwei0084@126.com

以此判断分析图像色彩与纹理特征。文献[8]使用道格拉斯-普克法提取中医舌诊齿痕舌图像特征,通过对比与医师观察的准确程度,判断舌部齿痕的症状轻重。文献[9]使用舌图像样本分类器令其重要样本增加,并采用两类加权SVM方法对舌象图像进行分类与识别。深度学习在目标识别等场景上能被大量应用,其优势是计算机在大量所提供的数据中自主学习并提取目标相应的特征,有效准确地表达目标特征且大幅提高目标检测的精确率^[10]。文献[11]利用深度卷积神经网络与迁移学习技术中的微调构建模型,实现对中医舌部图像局部特征提取与识别。

上述方法在中医舌诊领域分别取得了一些成果。目前卷积神经网络的训练还依赖大量样本^[12-13],但医学图像数据的标注存在一定困难,难以得到大规模数据集^[14],所以采用迁移学习^[15]训练网络。舌部纹理检测对舌部信息特征提取的要求较高,易受外部环境的影响,且通常采集的舌部数据中包括多标签与互相掺杂的目标,难以达到理想的检测效果^[16]。现阶段对此类问题的研究往往局限于通用特定检测的场景^[17],缺乏检测多尺度目标和目标掺杂的针对性方法^[18]。考虑到以上问题,本文提出一种基于改进Faster RCNN的中医舌部多纹理检测方法,该方法在高精度两阶段目标检测Faster RCNN模型的基础上做以下改进:不断调整卷积层的权重值进行迁移学习,同时用可变形卷积重塑卷积块,引入注意力机制scSE,采用加权双向特征金字塔网络。

1 改进的Faster RCNN检测模型

1.1 Faster RCNN网络结构

Faster RCNN是目标识别领域中具有高精度检测率的网络模型,其训练过程如图1所示。

首先将舌苔图片进行数据预处理;处理好的图片经过主干特征提取网络(Backbone)进行特征提取获得特征图^[19];在区域建议网络中使用滑窗生成对应候选框,同时将特征图通过一组3×3的卷积层,再使用两个并行的1×1卷积层分别进行分类(cls)与边界框回归(reg),获得含有目标概率候选图像;特征图和候选图像到区域池化层(ROI Pooling)进行池化,规格化含有候选区域的特征图;最后,调整建议框截取到的特征层的尺寸,并进一步卷积以完成目标物体的分类与回归^[20]。

1.2 网络设计

1.2.1 可变形卷积 在传统卷积中,特征图分割成与卷积核相同的大小进行卷积^[21],由于卷积区域固定,不易判断复杂变形的物体。可变形卷积能在感受野

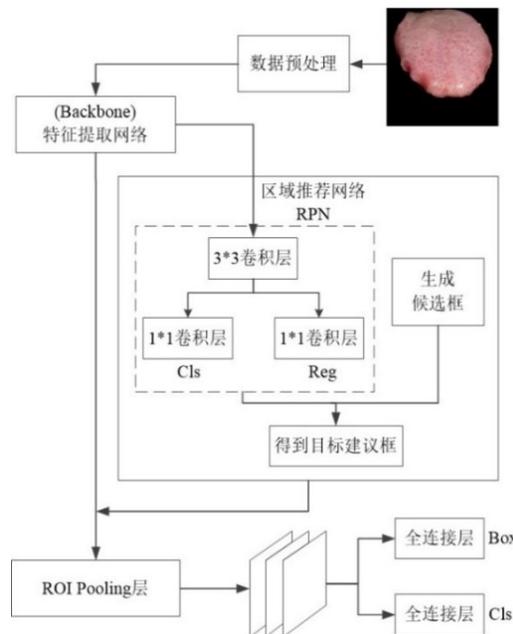


图1 Faster RCNN模型训练过程
Figure 1 Faster RCNN model training process

上引入可学习的偏移量,卷积区域能覆盖物体的实际形状边缘。

可变形卷积在原有传统卷积上引入一个标准的3×3卷积层,生成N个x,y方向上的二维偏移变量^[22],该卷积层的卷积核定义为:

$$R = \{(-1, -1), (-1, 0), \dots, (0, 1), (1, 1)\} \quad (1)$$

标准卷积操作与可变形卷积得到的特征矩阵分别为:

$$y(p_0) = \sum_{p_n \in R} \omega(p_n) \cdot x(p_0 + p_n) \quad (2)$$

$$y(p_0) = \sum_{p_n \in R} \omega(p_n) \cdot x(p_0 + p_n + \Delta p_n) \quad (3)$$

其中, p_n 是对 R 中所有位置的列举; Δp_n 代表位置偏移向量。

普通卷积操作的输出特征通道数是 N , 可变形卷积输出的偏差特征与其保持一致, 则叠加后通道维数是 $2N$ 。特征卷积核参数通过双线性插值后反向传播法计算得到, 可变形卷积过程如图2所示。

可变形卷积被卷积矩阵的尺寸有概率被扩展过大, 出现卷积核的感受野大于目标区域的情况, 进而引入无效信息。为保证有效信息的准确提取, 可变形卷积二代则新增加了一个权重矩阵, 为可变形卷积一代寻找有效信息的区域位置赋予权重, 其权重的获取方式与偏移矩阵的获取办法相同, 改进后的公式如下:

$$y(p) = \sum_{p_n \in R} \omega(p_n) \cdot x(p_0 + p_n + \Delta p_n) \cdot \Delta m_n \quad (4)$$

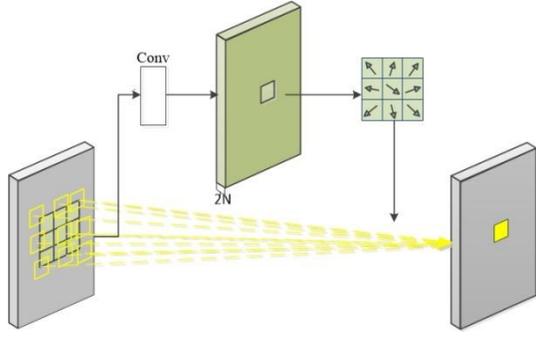


图2 可变形卷积过程

Figure 2 Deformable convolution

普通感兴趣区域池化层 (Region of Interest Pooling, RoI Pooling) 是把一张特征图分为若干个形状规则、位置固定的区域, 这给特征映射过程带来了局限性, 进而降低检测结果的准确性。池化操作也可借鉴可变形思想, 在其原有结构上增加了偏移特征, 公式如下:

$$y(k) = \sum \frac{x \cdot (p_n + \Delta p_n) \cdot \Delta m_n}{k} \quad (5)$$

1.2.2 scSE 注意力机制 为了提高舌部特征表达能力, 提出 scSE 注意力机制, 将空间压缩通道模块^[23] (sSE) 和通道压缩空间模块 (cSE) 结合, 主要目的是突出有用特征, 抑制无用特征。具体算法步骤如下:

设输入和输出特征图分别为 U 、 \hat{U} 。

步骤 1: 将 U 在通道上进行压缩, 设特征层 $U = [u^{1,1}, u^{2,2}, \dots, u^{i,j}, \dots, u^{H,W}]$, (H, W) 为特征图尺寸, (i, j) 为特征图的空间位置, 通过权重为 W_{sq} 的 1×1 卷积对空间进行压缩, 得到通道数为 1 的特征图 $q = W_{sq} \times U$;

步骤 2: 经过 Sigmoid 函数对 q 进行归一化处理, $\hat{U}_{sSE} = [\sigma(q_{1,1})u^{1,1}, \dots, \sigma(q_{i,j})u^{i,j}, \dots, \sigma(q_{H,W})u^{H,W}]$ 被得到, 其中 $\sigma(q_{i,j})$ 为特征图空间位置坐标的重要性;

步骤 3: 将 U 在空间上进行处理, 设其特征层 $U = [u_1, u_2, \dots, u_c]$, 通过全局池化后得到向量 z ;

步骤 4: 将向量 z 经过两次权重为 (W_1, W_2) 的全连接层进行信息的处理, 经过 ReLU 激活函数增强各通道独立性^[24], 得到 C 维度的向量 $\hat{z} = W_1(\delta(W_2 z))$;

步骤 5: 最后经过 Sigmoid 函数对 \hat{z} 进行归一化处理, 得到 $\hat{U}_{cSE} = [\sigma(\hat{z}_1)u_1, \sigma(\hat{z}_2)u_2, \dots, \sigma(\hat{z}_c)u_c]$, 其中 \hat{z} 代表第 i 个通道的重要性程度;

步骤 6: scSE 模块主要实现输入特征图 U 在空间与通道注意力模块中的作用, 将输出结果相加, 最终得到输出特征图 $\hat{U}_{scSE} = \hat{U}_{sSE} + \hat{U}_{cSE}$ 。实现过程如图 3 所示。

1.2.3 加权双向特征金字塔网络 舌部不同纹理样式

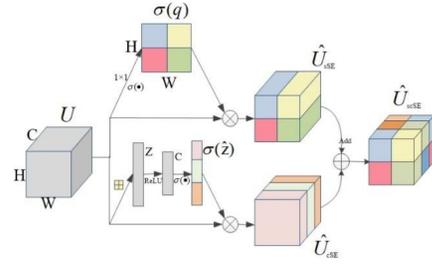


图3 scSE注意力机制实现过程

Figure 3 scSE attention mechanism

多变, 极易造成漏检问题, 为防止影响模型的识别效果, 特征需要进行多尺度融合^[25]。最初使用特征融合的卷积模块是 FPN, 通过连续的上采样并采用自顶向下的通道来融合多层特征。后来的 PANet 版本以 FPN 为基础, 增加了一个自下而上的通道弥补信息的单向性。而最新的版本加权双向特征金字塔 (BiFPN) 在简化的 PANet 增加了一条额外的线, 在不增加成本的情况下纳入更多的特征。上述金字塔模型如图 4 所示。

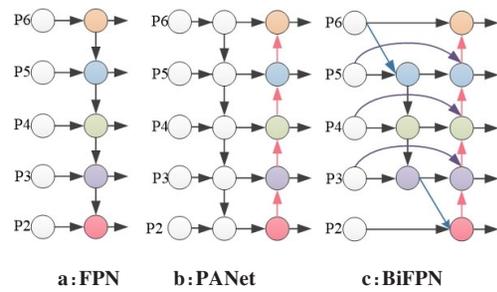


图4 特征金字塔模块

Figure 4 Feature pyramid module

多尺度特征融合的目的是聚合不同分辨率与语义的特征。在特征融合过程中, 对输入特征加额外的权重, 以避免最终特征对网络的输出不一致。采用加权融合能让网络学习不同输入特征的重要性, 增加额外权重的不同输出公式如下:

$$P_6^{out} = \text{Conv}(P_6^{in}) \quad (6)$$

$$P_5^{out} = \text{Conv}(P_5^{in} + \text{Resize}(P_6^{out})) \quad (7)$$

$$P_2^{out} = \text{Conv}(P_2^{in} + \text{Resize}(P_3^{out})) \quad (8)$$

选取的加权融合为快速及归一化特征融合, 在通过 RELU 激活函数之后来保证数值的稳定。图 4c 中 P_5^{id} 、 P_5^{out} 最终输出公式如下:

$$P_5^{id} = \text{Conv}\left(\frac{\omega'_1 P_5^{in} + \omega_2 \text{Resize}(P_6^{in})}{\omega'_1 + \omega'_2 + \omega'_3 + \varepsilon}\right) \quad (9)$$

$$P_5^{out} = \text{Conv} \left(\frac{\omega_1' P_5^{in} + \omega_2' P_5^{td} + \omega_3' \text{Resize}(P_4^{out})}{\omega_1' + \omega_2' + \omega_3' + \varepsilon} \right) \quad (10)$$

其中 P_5^{td} 、 P_5^{out} 分别表示 level5 在自顶而下与自底向上的中间特征和最终输出,其他特征都使用相同的方法构建。重复双向跨尺度连接与带权重是快速归一化特征融合结合,形成最终的加权双向特征金字塔网络。

1.2.4 微调迁移学习 迁移学习是将从源数据集学到的知识用到新目标进行学习。微调是其中常用的技术之一,当目标数据集样本较少时,是提升模型泛化能力的必要手段。具体流程为:在现有基础模型后增加全连接层,为使其发挥效果,令现有模型此时不进行训练,将现有网络结构冻结,只训练自定义添加的全卷积网络,训练后再将现有模型解冻加入全卷积网络共同进行训练。微调使整个模型不需要重新训练,训练效率就能得到提高,如果数据集不大,可以在相对较少的迭代次数内取得一个比较好的结果^[26]。

1.3 模型结构

为了更加高效地进行舌部多纹理研究,对 Faster RCNN 模型进行相应改进,在特征提取网络中,常规卷积块采用的是目前性能较好的 Resnet50,共 5 个阶段。骨干网络后引入可变形卷积,对原有特征提取网络中 Conv2-Conv5 常规卷积进行重塑,在提取特征时调整自身形状,同时使用可变形 RoI Pooling 使模型更准确地拟合纹理的形状和位置特征。在原有 Conv1 与替换后的 Conv2~Conv5 卷积块后引入 scSE 注意力机制,同时使用加权双向金字塔网络模型对添加注意力机制的卷积块进行特征融合。将最后一层全连接层前的所有层进行迁移,修改最后一层全连接层的分类信息,并为最后一层重新赋予随机权重,在最后的特征提取中尽可能最大化神经网络的性能。所使用的网络模型整体结构如图 5 所示。

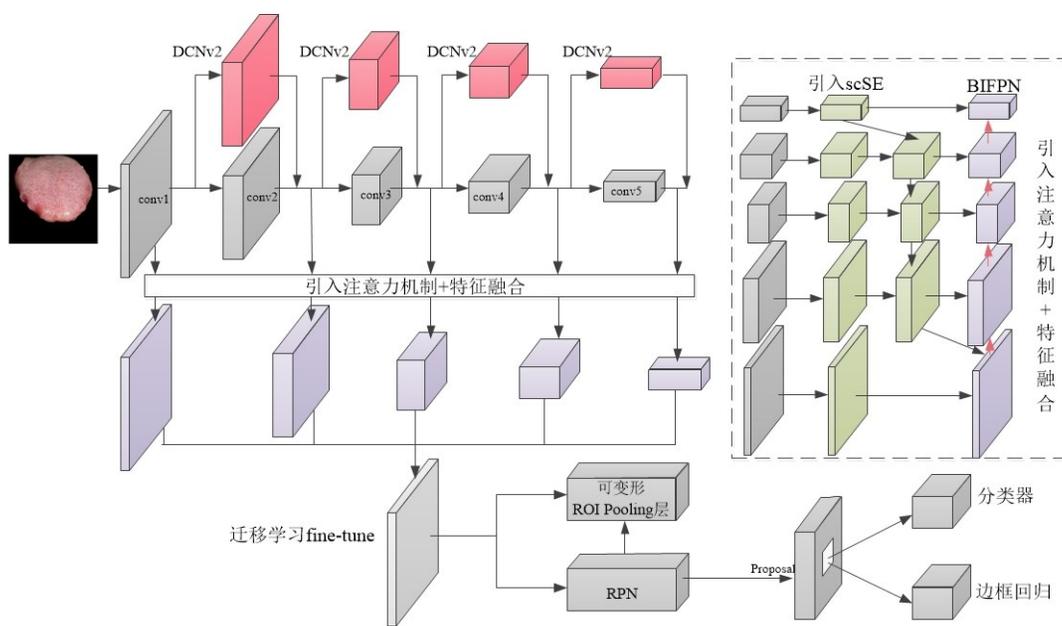


图5 本文网络模型整体结构

Figure 5 Overall structure of the proposed network model

2 实验及结果分析

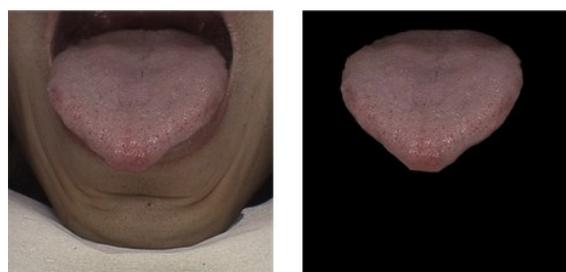
2.1 数据集与预处理

数据集采用的是哈尔滨工业大学提供的公开数据集^[27]与中医医疗机构提供的数据集。哈尔滨工业大学公开图像数据集包括 300 张舌图像,中医机构中获取的数据集包括 186 张舌图像。图像要求舌体暴露完整,舒展无遮挡。对数据集使用 labeling 工具进行标注,标注结果由多位专业中医学从业人员进行校验,最后对筛选后的舌部图像进行预处理。

数据集共分为齿痕舌、裂纹舌、腻苔舌、点刺舌等 4 类,数量分别为 142、213、358、189。获取数据集后将图像标签文件转为 JPG 格式,并以 8:1:1 的比例将数据集划分为训练、验证与测试集。

舌部纹理特征中的特征提取易受外部环境影响,而数据集除含有舌体目标外,仍存在与目标不相关的信息^[28]。因此数据集使用 Unet 神经网络模型对舌部区域进行自动定位与分割,有效避免传统分割方法产生的问题,必须针对不同数据分别提取图像特征^[29]。由于图像的空间分辨率差异较大,对舌部

数据在横断面上使用降采样和重采样,图片大小调整到256×256。对比预处理前后的舌部图像数据,如图6所示。



a: 预处理前

b: 预处理后

图6 舌部预处理前后对比图

Figure 6 Before and after tongue image preprocessing

2.2 评价指标

为了评估舌部纹理整体的检测能力,采用表1的混淆矩阵进行评估,并衍生了精确率(Precision)、召回率(Recall)、平均精度均值(mAP)等一系列指标。

表1 混淆数据
Table 1 Confusion matrix

混淆矩阵		预测值	
		Positive	Negative
真实值	Positive	TP	FN
	Negative	FP	TN

精准率、召回率分别为判断的正样本与实际正样本中正确的比例,公式如下:

$$\text{precision} = \frac{TP}{TP + FP} \quad (11)$$

$$\text{recall} = \frac{TP}{TP + FN} \quad (12)$$

平均精度(AP)代表反映各类别性能的评价指标,为PR曲线的面积值,所有指标取值均在0~1之间,越接近1表明效果越好。mAP是所有AP的平均值,用于衡量模型所有类别性能的评价指标,公式如下:

$$AP = \int_0^1 P(R) dR \quad (13)$$

$$mAP = \frac{\sum_{i=1}^M AP_i}{M} \quad (14)$$

其中*i*代表类别,*M*为总类别。检测准确率评价指标使用平均精度均值 mAP_{50%},即目标检测边界框与真实边界框的交集和并集的比值大于0.5时,视为预测正确,反之则预测错误。

2.3 实验细节

实验在 Windows 系统下采用 Pytorch1.4.0 深度学习算法框架并利用 GPU 进行加速计算。实验系统的处理器为 Core(TM) i7-12700H 2.50 GHz,显存大小为 16 GB,GPU 为 NVIDIA GeForce RTX 3060。消融实验使用标准随机梯度下降法进行训练, batchsize 为 2, 每个实验训练 60 个 epoch。

2.4 实验结果

2.4.1 消融实验 在相同舌部数据集上和相同实验环境下,分别训练原 Faster RCNN 模型与进行迁移学习的 Faster RCNN 模型。通过图 7 混淆矩阵的对比,可以发现裂纹舌与腻苔舌的分类较为容易,分类效果较好,进行迁移学习微调后的 Faster RCNN 模型 (Faster RCNN-ft) 不仅提升舌部纹理小样本训练的性能,更快并有效地辨别目标形状,同时也可以避免过拟合纹理。从图 7 可知,加入迁移学习后,舌部目标被误判为背景 (FP) 的概率降低,网络的识别能力有一定提升。

舌部纹理普遍状态多样、细节差异过大,对舌部表面纹理的检测会受到形状因素的影响,因此在微调过 Faster RCNN 模型的特征提取网络将标准卷积重塑为可变形卷积 (DCNv2) 后,与原始模型检测准确率差别的对比实验结果如表 2 所示,说明可变形卷积比起常规卷积,进行采样点易学习到舌部纹理在多形态纹理目标,能精准地定位目标,拥有更好的检测效果。由于多任务纹理中的点刺与裂纹属于舌部图像中弱边缘目标,且大部分处于腻苔之中,提出一种提高特征的注意力机制 scSE 结构,可以看出 mAP 有适当提升,说明增强了特征图上的有意义像素。针对舌部图像中存在多尺度的目标特征图,通常网络使用特征金字塔网络 (FPN) 对底层空间位置特征和高层语义特征进行互补,最后对融合 FPN 和 BiFPN 的改进后模型整体性能分别进行验证。从表 2 中可看出,当优化后的模型引入 BiFPN 时,速度效率优于 FPN 特征融合网络,检测精度较 FPN 特征融合网络也有一小部分的提升。可以通过增添较少的计算量,融合不同语义信息与分辨率的特征图。

2.4.2 对比实验 为了验证改进模型的有效性,与现阶段主流无锚框检测模型 YOLOv4、YOLOv5、两阶段检测模型 Double-Head RCNN 和 SSD 模型进行对比分析。实验结果如表 3 所示,表中列出了各类别在不同先进算法上进行检测的平均精度与召回率。可以看出,改进后的算法在齿痕舌、腻苔舌、裂纹舌上的 mAP 均优于上述模型,在其他类别上的检测效果也较好,因为改进的舌部纹理检测算法充分考虑到

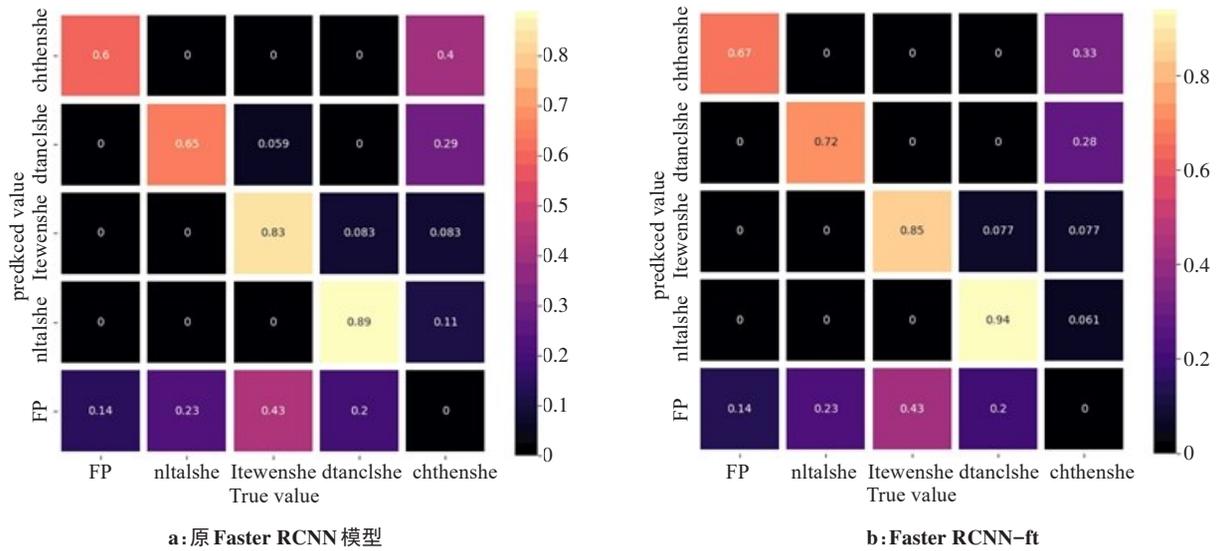


图7 混淆矩阵对比图

Figure 7 Confusion matrix comparison

表2 各模块对模型的性能影响

Table 2 Effect of each module on the model performance

实验	基础模型	DCNv2	scSE	FPN	BIFPN	平均精度均值/%	检测时间/s
1	Faster RCNN-ft	×	×	×	×	72.1	0.036
2	Faster RCNN-ft	√	×	×	×	75.9	0.045
3	Faster RCNN-ft	√	√	×	×	77.9	0.048
4	Faster RCNN-ft	√	√	√	×	88.3	0.069
5	Faster RCNN-ft	√	√	×	√	93.5	0.061

了舌部数据目标掺杂的特点,对数据集样本小、舌纹理形态多变、边缘不清晰与多尺度目标问题进行针对性改进。本文模型的mAP为93.5%,较原始Faster RCNN模型提高45.7%,同时优于其余模型,进一步验证了本研究方法的有效性。综上所述,改进Faster RCNN模型是可行的,可以最大程度地提高舌部纹理检测的精确度。

表3 实验结果比较(%)

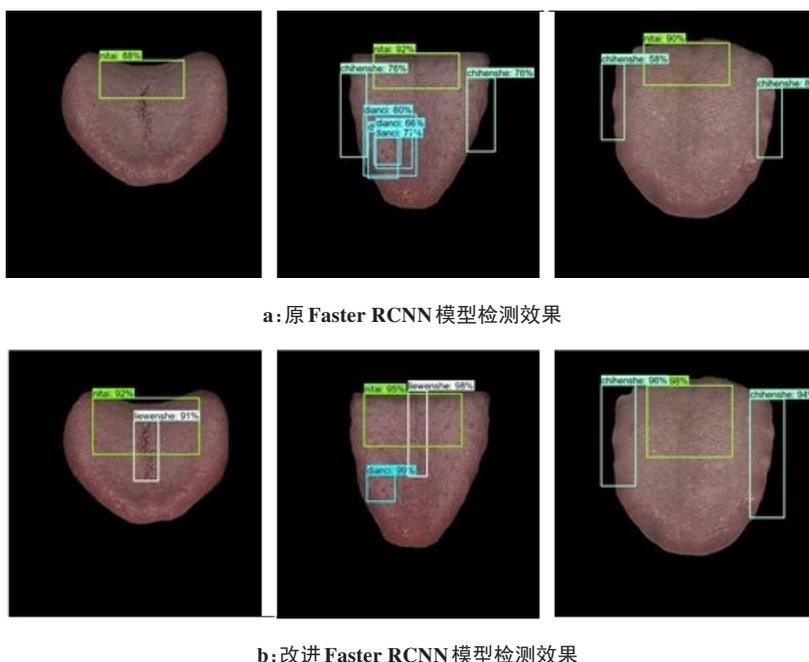
Table 3 Comparison of experimental results (%)

算法	齿痕	点刺	裂纹	腻苔	平均精度均值
YOLOv4	55.5	45.5	55.2	57.8	53.5
YOLOv5	59.9	55.3	57.9	60.5	58.4
Double-Head RCNN	62.9	59.6	61.2	65.5	62.3
SSD	58.6	54.1	59.7	57.6	57.5
本文方法	93.0	91.7	92.8	96.8	93.5

2.4.3 测试效果验证 为了更直观地区分改进Faster RCNN模型与原始模型在舌部纹理的检测效果,从舌部纹理数据集的测试集中随机抽取图像在相同实验环境下进行实验,其检测结果如图8所示。可以看出,改进后的模型相较于原始模型,舌部纹理重叠导致的漏检率降低,精确度也在一定程度上提升。

3 结论

为高效并行地提高舌部纹理检测精度,本文提出一种基于Faster RCNN网络并结合迁移学习、可变形卷积、scSE注意力机制和加权双向特征金字塔网络的舌部多纹理检测模型。通过与其他模型对比,改进模型较原基础模型在舌部数据集大幅度改善了精确率不高与纹理重叠导致漏检的问题。为进一步验证模型的泛化能力,未来会对舌部更多纹理进行扩充,使用更多种不常见舌部纹理进行深入研究,得到更全面的舌部类别。



a:原Faster RCNN模型检测效果

b:改进Faster RCNN模型检测效果

图8 不同模型图像检测结果比较

Figure 8 Comparison of image detection results of different algorithms

【参考文献】

- [1] 肖二钢, 王益民. 中医舌诊研究的回顾与展望[J]. 天津中医药, 2006, 23(4): 348-350.
- [2] 周越, 杨杰, 沈利. 中医舌象信息的数字化方法研究[J]. 生物医学工程学报, 2004, 21(6): 917-920.
- [3] 张鑫, 朱明峰, 杜建强, 等. 中医数据挖掘算法研究进展[J]. 江西中医药大学学报, 2015, 27(1): 111-113.
- [4] Fu M, Xu P, Li X, et al. Fast crowd density estimation with convolutional neural networks[J]. Eng Appl Artif Intell, 2015, 43: 81-88.
- [5] 诸薇娜, 周昌乐, 徐丹, 等. 基于颜色纹理的图像多特征检索技术在中医舌诊中的应用研究[J]. 中国图象图形学报, 2005, 10(8): 57-63.
- [6] 杨朝辉, 张大鹏, 李乃民. 裂纹舌图像的核假彩色变换及其纹线提取[J]. 计算机辅助设计与图形学学报, 2010, 22(5): 771-776.
- [7] 朱洁华, 阮邦志, 励俊雄, 等. 舌诊客观化研究的一种图像处理方法[J]. 中国生物医学工程学报, 2001, 20(2): 132-137.
- [8] 朱穆朗玛, 陆萍, 夏春明, 等. 基于道格拉斯-普克法提取55例齿痕舌图像特征研究[J]. 中华中医药学刊, 2014, 32(9): 1673-1677.
- [9] 张新峰, 沈兰荪. 加权SVM在中医舌象分类与识别中的应用研究[J]. 中国生物医学工程学报, 2006, 25(2): 230-233.
- [10] Zhuang QB, Gan S, Zhang LY. Human-computer interaction based health diagnostics using ResNet34 for tongue image classification[J]. Comput Method Prog Biomed, 2022, 226: 107096.
- [11] 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.
- [12] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]//2016 IEEE CVPR. 2016: 770-778.
- [13] Gorur K, Bozkurt MR. GKP signal processing using deep CNN and SVM for tongue-machine interface[J]. Traitement Du Signal, 2019, 36(4): 319-329.
- [14] 王丽冉, 汤一平, 陈朋, 等. 面向舌体分割的两阶段卷积神经网络设计[J]. 中国图象图形学报, 2018, 23(10): 1571-1581.
- [15] 宋超, 王斌, 许家伦. 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.
- [16] Zhou H, Hu G, Zhang X. Constitution identification of tongue image based on CNN[C]//2018 11th International Congress on Image and Signal Processing, Biomedical Engineering and Informatics (CISP-BMEI). New York: IEEE, 2018: 1-5.
- [17] Lin TY, Dollar P, Girshick R, et al. Feature pyramid networks for object detection[C]//Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2017: 2117-2125.
- [18] Peng CD, Wang L. Establishing and validating a spotted tongue recognition and extraction model based on multiscale convolutional neural network[J]. Digital Chinese Medicine, 2022, 5(1): 49-58.
- [19] Jiang T, Hu XJ, Yao XH, et al. Tongue image quality assessment based on a deep convolutional neural network[J]. BMC Med Inform Decision Making, 2021, 21(1): 1-14.

- [20] 邹斌, 张聪. 基于Faster R-CNN的密集人群检测算法[J]. 计算机应用, 2023, 43(1): 61-66.
Zou B, Zhang C. Dense crowd detection algorithm based on Faster R-CNN[J]. Computer Applications, 2023, 43(1): 61-66.
- [21] Rawat W, Wang Z. Deep convolutional neural networks for image classification: a comprehensive review[J]. Neural Computation, 2017, 29(9): 2352-2449.
- [22] 赵楚, 段先华, 苏俊楷. 改进Faster RCNN的瓷砖表面瑕疵检测研究[J]. 计算机工程与应用, 2023, 59(3): 1-10.
Zhao C, Duan XH, Su JK. Improved Faster RCNN for tile surface defect detection[J]. Computer Engineering and Applications, 2023, 59(3): 1-10
- [23] Hu Y, Wen G, Luo M, et al. Fully-channel regional attention network for disease-location recognition with tongue images[J]. Artif Intell Med, 2021, 118(5): 102-110.
- [24] Prajapati MD, Namdev N. Prediction of heart disease risk based on deep learning[J]. Int Res J Modern Eng Tech Sci, 2022, 4(2): 1011-1016.
- [25] Zhang Y, Wang M, Li Z. An efficient object detection framework with modified dense connections for small objects optimizations[C]// Proceedings of the 17th ACM International Conference on Computing Frontiers. New York: ACM, 2020: 174-181.
- [26] Li J, Chen Q, Hu X, et al. Establishment of noninvasive diabetes risk prediction model based on tongue features and machine learning techniques[J]. Int J Med Inform, 2021, 149(2): 104429.
- [27] 崔振超. 计算机舌诊中舌体分割与纹理分类研究[D]. 哈尔滨: 哈尔滨工业大学, 2015.
Cui ZC. Research on tongue segmentation and texture classification in computerized tongue diagnosis[D]. Harbin: Harbin Institute of Technology, 2015.
- [28] Xu Q, Zeng Y, Tang W, et al. Multi-task joint learning model for segmenting and classifying tongue images using a deep neural network [J]. IEEE J Biomed Health Inform, 2020, 24(9): 2481-2489.
- [29] Li J, Zhang Z, Zhu X, et al. Automatic classification framework of tongue feature based on convolutional neural networks [J]. Micromachines, 2022, 13(4): 501-513.

(编辑:黄开颜)