

基于改进 YOLOv5 的舌面特征检测

张德龙, 金春阳, 张志东, 曹溪源, 薛晨阳

(中北大学 仪器科学与动态测试教育部重点实验室, 太原 030051)

摘要: 针对传统中医舌诊视觉诊断存在主观性强且耗费精力的问题, 提出一种基于改进 YOLOv5 的舌面齿痕和裂纹特征自动检测模型: 该模型在 YOLOv5 模型的骨干网络中引入 SimAM-CSP 模块以增强网络的特征提取能力; 在瓶颈层和预测部分之间加入瓶颈注意力模块, 进一步聚焦关键信息; 通过调整 YOLOv5 的特征融合结构, 增加对图像细节的感知能力, 提高网络性能; 将定位损失函数 GIoU 替换为 EIoU, 提升模型的训练收敛速度和预测回归精度; 利用 K-Means 聚类算法对 YOLOv5 的初始锚框进行调整, 使算法更加契合舌面齿痕和裂纹特征检测; 将改进后的 YOLOv5 模型在自制舌象数据集中进行训练, 得到的平均检测精度 (mAP) 为 79.5%, 较原算法提升了 6.3 个百分点。实验结果表明改进 YOLOv5 模型能够有效提高舌面齿痕和裂纹特征检测精度, 有助于辅助医生诊断。

关键词: YOLOv5; 深度学习; 目标检测; 异常舌; 舌诊

Tongue Feature Detection Based on Improved YOLOv5

ZHANG Delong, JIN Chunyang, ZHANG Zhidong, CAO Xiyuan, XUE Chenyang

(Key Laboratory of Instrumentation Science & Dynamic Measurement of Ministry of Education,
North University of China, Taiyuan 030051, China;)

Abstract: Visual diagnosis has the shortages of subjectivity and energy-consuming in traditional Chinese medicine tongue diagnosis, a detection model for tongue tooth mark and fissure features based on improved YOLOv5 is proposed. In the backbone of YOLOv5, this model introduces the SimAM-CSP module to enhance the feature extraction capability of the network. The Bottleneck Attention Module is added between Neck layer and Head layer to further focus on critical information. The feature fusion structure of the YOLOv5 is adjusted to enhance the perception ability of image details and improve the performance of the network. The localization loss function GIoU is replaced with the EIoU to improve training convergence speed and predict regression accuracy. The initial anchor frames of the YOLOv5 are adjusted by using the K-Means algorithm to make the model more suitable for tongue tooth mark and fissure detection. The improved YOLOv5 model is trained on the self-built tongue image dataset, the obtained mean average precision (mAP) is 79.5%, which is 6.3% higher than that of the original algorithm. Experimental results show that the improved YOLOv5 model can effectively improve the accuracy of tongue tooth mark and fissure detection, which assists doctors in diagnosis.

Keywords: YOLOv5; deep learning; object detection; abnormal tongue; tongue diagnosis

0 引言

舌诊作为中医望诊的重点内容之一, 具有无接触、无创、实时、便捷等优势^[1], 主要是通过观看舌质和舌苔的颜色、形状来做出病理判断^[2-4]。脏腑通过经脉与舌相联, 其病变可在舌质和舌苔上反映出来, 因此舌诊可通过舌象判断脏腑的虚实、气血津液的盛衰。齿痕舌即

舌体的边缘见牙齿的痕迹^[5-7], 一般出现在舌体的两侧, 舌体形状呈锯齿状, 受牙齿挤压处的舌体色泽暗红, 未受挤压处的舌色偏淡白^[4-6]。裂纹舌即舌面存在深度和数量不同的裂纹^[8-9], 各式各样的裂纹一般都预示着患有某种疾病。齿痕舌和裂纹舌是舌诊中常见的异常舌象, 在舌诊的研究中引起众多研究者的关注。中医认为齿痕舌多由气虚、脾虚或阳虚导致舌体胖大^[7], 而裂纹舌则由

收稿日期: 2024-11-21; 修回日期: 2024-11-25。

基金项目: 山西省重点研发计划(202102130501011); 中北大学研究生科技立项(20231940)。

作者简介: 张德龙(2000-), 男, 硕士研究生。

张志东(1985-), 男, 博士, 教授。

薛晨阳(1971-), 男, 博士, 教授。

引用格式: 张德龙, 金春阳, 张志东, 等. 基于改进 YOLOv5 的舌面特征检测[J]. 计算机测量与控制, 2025, 33(4): 89-94, 146.

心火上炎或阴血亏虚所致^[9]，通过异常舌形态的严重程度可以判断患者的身体状况。中医对于异常舌一般通过视觉诊断，但是医生没有太多的时间和精力去对齿痕进行标记和记录，而且传统舌诊主要依据医师个人经验，易受环境影响，因此一个可以辅助医生进行齿痕舌诊断的自动检测模型对于人们的健康管理至关重要。

舌体容易和嘴唇混淆，舌体的姿势和不对称也会影响判断^[4]，因此异常舌的自动检测是一个具有挑战性的任务。近年来已经有研究人员进行异常舌特征的识别任务。文献 [5] 和文献 [6] 等采用传统图像处理的方法，利用彩色空间来分析齿痕舌特征，但是这种方法存在一定的噪声导致提取不准确。文献 [10] 和文献 [11] 等则采用深度学习的方法进行舌面齿痕和裂纹特征的检测，利用卷积神经网络 (CNN) 提取特征并进行分类。文献 [12] 以 YOLO 模型为基础提出一种弱监督方法，可以利用粗糙标注的数据集来训练齿痕和裂纹检测模型，但是检测效果较差。

本文提出了基于改进 YOLOv5 算法的舌面齿痕和裂纹特征自动检测模型。首先，在骨干网络中引入了 SimAM-CSP 模块以增强网络的特征提取能力；其次，针对齿痕和裂纹的特征尺寸，在 Neck 层加入了瓶颈注意模块，提高信息利用率；然后优化了 YOLOv5 的特征融合结构，增加对图像细节的感知能力；另外，为了加快模型训练的收敛速度，进一步提高整体预测回归精度，在初始 YOLOv5 的基础上优化定位损失函数，将 GIoU^[13] 损失函数替换为 EIoU^[14]；最后，利用 K-Means 算法对自制舌象数据集的真实框进行聚类，调整模型初始锚框的大小，使 YOLOv5 模型更加契合异常舌的检测。最终训练结果 *mAP* 为 79.5%，能够较为精确检测齿痕和裂纹的位置。优于 Faster RCNN^[15]、SSD^[16] 和 YOLOv3^[17] 等算法，且较初始模型提高了 6.3 个百分点，验证了优化措施的有效性。

1 改进 YOLOV5 模型设计

随着深度学习与卷积网络研究的逐渐成熟，目标检测已经成为计算机视觉领域的一个重要热点^[18]，旨在定位图像中的特定目标同时进行类别判断^[19]。现有的检测算法可以分为 two-stage 方法和 one-stage 方法两大类。其中 two-stage 方法开发较早，主要包括 RCNN 算法以及由此延伸的 Faster-RCNN。

Two-stage 方法需要进行两个步骤，首先利用 RPN 网络提取特征信息生成候选框，然后对目标的类别和位置进行检测。这种方法具有很高的检测精度，但是由于结构比较复杂，模型训练和检测的速度较慢。One-stage 方法的特点是输入端到输出端一步到位，速度较快，非常适合实时性检测的任务，代表算法有 SSD 和

YOLO 系列。YOLOv1 是 one-stage 方法的早期版本，后来的 YOLO 系列算法都是沿用了最初的检测思想，即把目标检测任务转化为回归问题，将整个图片作为网络的输入，经过特征提取后直接获取预测结果。

YOLOv5 于 2020 年首次提出，直至今日仍不断地进行升级迭代，是目前应用最广泛的检测模型之一。YOLOv5 拥有 5 个架构，分别是 YOLOv5n、YOLOv5s、YOLOv5m、YOLOv5l 和 YOLOv5x。它们之间的区别在于架构的网络深度和宽度不同，而且参数量也随之不同。基于训练时间和模型大小的同步考虑，我们最终选择了综合性能较好的 YOLOv5s。

如图 1 所示，YOLOv5 模型由 3 个部分组成，分别是主干网络 (Backbone)、融合层 (Neck)、检测头 (Yolo-Head)。

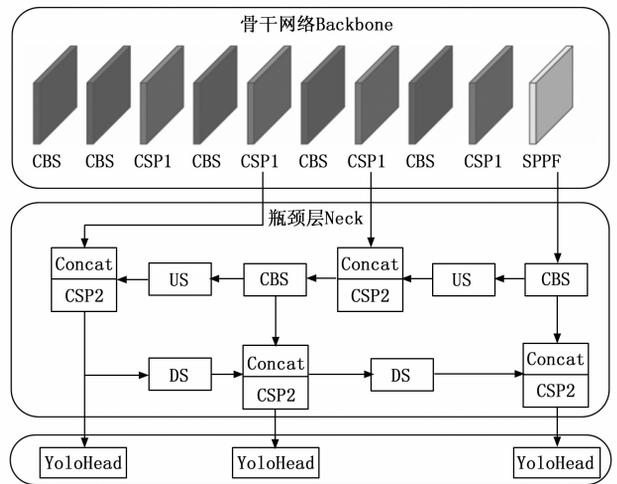


图 1 YOLOv5 框架

1) Backbone 部分的核心是 CSP- Darknet53 特征提取网络，由如图 2 所示的若干个 CBS 和 CSP1 模块构成，通过残差网络处理后输出 20×20 、 40×40 、 80×80 三种尺度特征。另外，采用 SPPF 模块替换空间金字塔池化层 SPP 结构，并应用在特征提取的最后一步。SPPF 模块利用 3 个相同尺度的最大池化层依次进行特征处理，并将结果拼接在一起，能够进一步融合特征，从而加快特征提取的速度，并解决目标的多尺度问题。

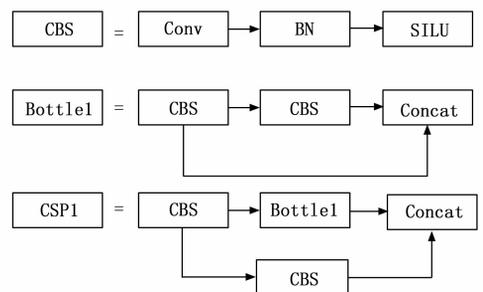


图 2 CBS 和 CSP1 模块结构

2) Neck 部分采用 FPN (Feature Pyramid Network, 特征金字塔网络) 结构和 PAN (Path Aggregation Network, 路径聚合网络) 结构相结合的方式特征融合。FPN 结构融合高层特征和低层特征以提取语义信息, PAN 结构将定位信息从低层传递到高层以增强不同尺度的定位能力。另外, Neck 部分中还添加了如图 3 所示的 CSP2 模块, 可以配合网络进一步融合特征。

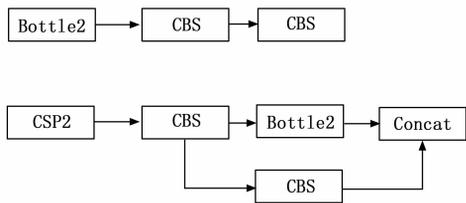


图 3 CSP2 模块结构

3) YoloHead 部分结合 Neck 融合的特征信息预测舌图像中的齿痕和裂纹, 生成边界框参数、置信度参数和类别参数来进行定位和分类, 并且会在最后使用非极大值抑制 (NMS) 的方法来筛选目标框。

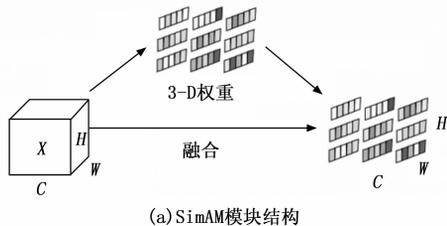
2 优化机制

本研究通过对 YOLOv5 目标检测模型进行改进, 提出了一种可用于舌面齿痕和裂纹检测的模型。主要改进工作如下: 1) 提出 SimAM-CSP 模块以增强特征提取能力; 2) 在 Neck 层中加入瓶颈注意模块 (BAM) 进一步提高信息的融合能力; 3) 改进了 Neck 结构的 concat 拼接节点; 4) 优化了损失函数以提高回归精度; 5) 调整了初始锚框的尺寸。

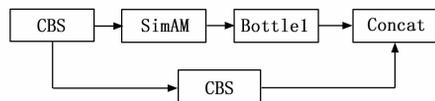
2.1 SimAM-CSP 模块

舌象图片中舌部易与唇部和其他面部混淆, 并且背景的重复迭代会产生冗余信息, 导致出现漏检和错检的问题。如图 4 所示, 我们提出了 SimAM-CSP 模块, 在 CSP1 模块的基础上, 加入了 SimAM 注意力机制模块^[20], 增强 Backbone 的特征提取能力, 并抑制其他背景因素的干扰。

SimAM 模块是 Yang 等人^[20]提出的一种 3D 注意力模块, 巧妙地运用一种能量函数机制, 旨在发掘并量化每个神经元的重要性, 进而精准地分配相应的权重。它首先计算每个神经元与邻域内其他神经元的差异 (通常是通过计算平方差), 然后根据这些差异来推断每个神经元的重要性。最后, 通过一个简单的激活函数将这些重要性转换为注意力权重, 并将其应用于原始特征图, 以增强或抑制某些区域的信息。SimAM 模块在处理机制上显著区别于传统的通道和空间注意力机制, 它更加聚焦于局部相似性的分析, 而非全局上下文信息的捕捉, 没有向原始网络添加额外的参数, 可以随意地嵌入到卷积神经网络中。



(a) SimAM 模块结构



(b) SimAM-CSP 模块结构

图 4 模块结构图

2.2 BAM 注意力机制

经过 Backbone 层特征提取后, 图像会变得更加抽象, 特征图的分辨率也会随之降低, 而 FPN 和 PAN 相结合的结构能够帮助模型提高信息利用率。由于舌齿痕和裂纹的特征大多为中型或大型尺寸, 本文针对该特点在 Neck 层与 YOLOHead 相应的节点设计添加了瓶颈注意力模块 (BAM), 进一步提高信息的利用率。

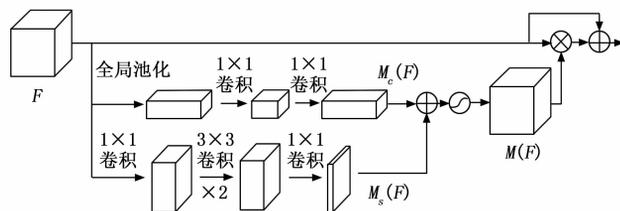


图 5 BAM 注意力模块结构

BAM^[21]是一种简单高效的注意力模块, 能够通过生成三维注意力图聚焦重要元素。如图 5 所示, 输入一个特征图 F , 依靠 BAM 模块中两个独立的注意力分支——通道 $M_c(F)$ 和空间 $M_s(F)$ 计算相应的特征图, 最终计算得到的特征图如式 (1) 所示:

$$F' = F + F \otimes M(F) \quad (1)$$

其中: \otimes 为按位相乘。具体地, $M_c(F)$ 利用全局平均池化和两次 1×1 的卷积得到通道注意力特征图, 而 $M_s(F)$ 则通过两次 1×1 卷积和两次 3×3 卷积得到空间注意力特征图, 然后再将特征图相加, 经过激活函数后生成最终的注意力特征图, 如式 (2) 所示:

$$M(F) = \sigma(M_c(F) + M_s(F)) \quad (2)$$

其中: σ 为 sigmoid 激活函数。

将 BAM 与 Neck 层相结合, 能够在目标预测时聚焦关键信息, 获取更多与目标有关的细节特点, 从而提高检测精度。

2.3 Neck 层结构优化

目标检测网络识别性能的关键因素在于如何巧妙地融合不同尺度下的特征图, 根本目的在于显著提升特征

的可判别能力，进而助力模型更加精准且深入地理解输入的复杂数据。在这一过程中，特征图的层次性扮演着至关重要的角色。低层特征往往直接源自原始图像数据，它们保留了图像的边缘、纹理等细节信息，但映射的语义层次相对较低，这些信息虽然基础却略显不足。相对而言，深层特征则是通过模型内部的多层次非线性变换和抽象得到的，它们能够更好地捕捉和表达数据的深层次语义信息，具有更为强大和丰富的表征能力，对于识别和理解图像中的目标至关重要。

为了充分利用这两种特征的优势，并克服它们各自的局限性，设计了一种创新的融合策略。如图 6 所示，我们将最深层的网络所提取到的特征，在上采样后直接与 PAN 结构中的特征融合。这一步骤巧妙地绕过了 CSP2 和额外的卷积层，从而避免了在特征传递过程中可能导致的高级语义信息的损失。通过这种方式，我们不仅能够保留和增强特征中的高层语义信息，还能有效增加模型对图像细节信息的感知能力，使得目标检测网络在复杂多变的场景中能够表现出更加卓越的性能。改进后的模型整体结构如图 7 所示。

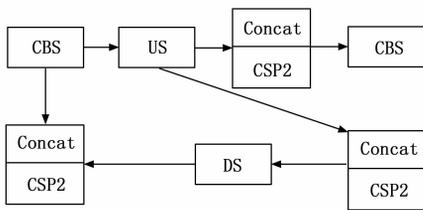


图 6 优化后的特征融合结构

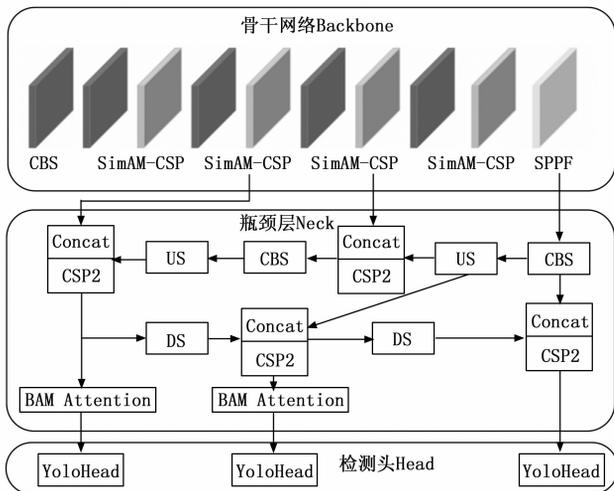


图 7 改进后的 YOLOv5 框架

2.4 损失函数优化

$IoULoss^{[22]}$ 是评价模型预测框与真实框重合度的常见损失函数，即预测框与真实框的重合面积与合并面积的比值。而 YOLOv5 采用 $GIoULoss$ 作为边界框损失函数，其计算方式如式 (3)：

$$L_{GIoU} = 1 - IoU + \frac{A^c - U}{A^c} \tag{3}$$

其中： A^c 为包含真实框和预测框的最小矩形的面积。 $GIoU$ 引进了最小外接矩形，具有了表示相对位置的信息，但是当真实框完全包裹预测框时， $GIoU$ 无法区分其相对的位置关系。因此，我们将 $GIoU$ 损失函数替换为 $EIoU$ ，其计算方式如式 (4)：

$$L_{EIoU} = 1 - IoU + \frac{\rho^2(b, b^{gt})}{(\omega^c)^2 + (h^c)^2} + \frac{\rho^2(\omega, \omega^{gt})}{(\omega^c)^2} + \frac{\rho^2(h, h^{gt})}{(h^c)^2} \tag{4}$$

其中： $\rho^2(b, b^{gt})$ 为预测框与真实框的中心点的欧氏距离， ω^c 和 h^c 为包含预测框和真实框的最小外接矩形的宽和高， ω ， h 和 ω^{gt} ， h^{gt} 分别为预测框和真实框的宽和高。 $EIoU$ 作为一个优秀的回归定位损失函数，同时考虑到了重叠面积、中心点距离和长宽比 3 种参数，可以加快网络收敛速度，提高回归精度。

2.5 初始锚框调整

目标检测算法通常会生成多个大小和高宽不一的边界框，即为锚框。遍历输入图像中所有可能存在目标的锚框，然后选出正确的目标框，并且对位置和大小进行回归调整。YOLOv5 算法的初始锚框大小是创作者经过很多次训练得到的，分别对应 80×80 、 40×40 、 20×20 不同尺度的预测层，通常适应于不同的检测场景。针对舌齿痕和裂纹检测任务，需要的真实框大多为竖直矩形框，然而初始锚框多为方框，起不到有效框选目标的作用。因此，利用 K-Means 聚类的方法产生随机聚类中心，对初始锚框的尺寸进行了调整，并且经过多次训练后得到 9 组锚框，具体尺寸如表 1 所示。

表 1 调整后的锚框尺寸

| 尺度 | 80×80 | 40×40 | 20×20 |
|------|----------------|-----------------|-----------------|
| 锚框尺寸 | 27×51 | 41×80 | 70×151 |
| | 44×53 | 67×90 | 68×178 |
| | 51×78 | 49×125 | 87×223 |

3 实验准备

3.1 数据集

本实验使用的自建数据集主要为使用统一的仪器采集患者的舌面图像，在采集过程中，保持拍摄环境标准且稳定，严格保证舌头图像的质量，并且在每次使用后对采集设备进行消毒通风，确保采集环境的卫生和安全。该数据集包含了 600 名以上患者的舌头图片，年龄分布在 18~70 岁之间，图片像素大小为 3264×2448 。我们对这些图片进行旋转、镜像、增加噪声等增强处理，最终得到了含有 1290 张图片的数据集，其中包含了 1091 个齿痕病灶和 1248 个裂纹病灶。同时使用 Make Sense 在线标注网站对采集到的舌头数据集进行

标注, 并如表 2 所示, 按照 7 : 2 : 1 的比例分为训练集、验证集和测试集。

表 2 数据集的图像分布情况

| 数据集 | 总数据集 | 训练集 | 验证集 | 测试集 |
|-----|-------|-----|-----|-----|
| 数量 | 1 290 | 928 | 233 | 129 |

3.2 实验环境

本实验的训练平台配置如表 3 所示。迭代次数设置为 250 个 epoch, 其中前 50 个 epoch 为冻结训练, 其中在第 170 个 epoch 时, 取消了 mosaic 训练机制, 加速收敛; 使用的优化器是随机梯度下降法 (SGD), 冻结训练和非冻结训练的 batch-size 均设置为 16。为了避免学习率过大错过最优解, 设置了学习率自动衰减机制。

表 3 实验环境

| 条目 | 版本 |
|--------|------------------------|
| 显卡 | Nvidia RTX4090 |
| CPU | Intel Xeon Silver 4316 |
| 系统 | Ubuntu 20.04 操作系统 |
| CUDA | 11.6 |
| Python | 3.7 |
| 框架 | Pytorch |

3.3 评价指标

本文通过均值平均精度 (mAP), 精确率 (Precision), 召回率 (Recall) 作为评价指标来评估舌齿痕和裂纹检测模型的有效性。AP 是平均精度, 在数值上等于 P-R 曲线下的面积。mAP 表示平均精度的平均值, 即各类的均值平均精度, 用以衡量训练模型在各个目标中的检测能力。mAP 计算方法如式 (5):

$$mAP = \frac{1}{|Q_R|} \sum_{q=Q_q} AP(q) \quad (5)$$

Precision 和 Recall 分别表示模型判断正确类别的准确率和分辨所有类别中正类的能力。二者的关系一般是负相关的, 即 Precision 提高时, Recall 往往会降低。二者的计算方式如式 (6) (7):

$$P = \frac{TP}{TP + FP} \quad (6)$$

$$R = \frac{TP}{TP + FN} \quad (7)$$

其中, TP (True Positives) 代表真实特征中成功检测的个数, FP (False Positives) 代表真实特征中错误检测的个数, FN (False Negatives) 代表漏检个数。

4 实验结果与分析

我们使用包括 129 张图片的测试集对改进前后的模型进行性能测试。如图 8 (a) 所示, 改进后的模型在 P-R 曲线图上呈现出更为优越的表现。具体而言, 改进后模型的 P-R 曲线所包裹的面积显著大于初始的 YOLOv5

模型, 这一显著差异直接反映了在舌齿痕和裂纹数据集中, 我们改进后的模型性能更加有效。图 8 (b) 则直观地呈现了模型改进前后在训练过程中的损失函数曲线对比, 虽然两种模型的初值基本相等, 随着训练轮数的不断增加, 改进后的模型展现出了更为出色的优化能力, 其损失值始终保持在低于初始模型的水平上, 且收敛后未出现过拟合。这表明我们的改进策略不仅提升了模型的泛化能力, 还确保了其稳定性和可靠性。

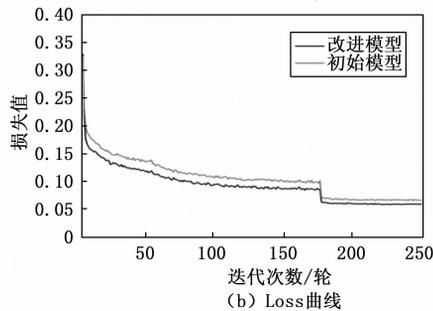
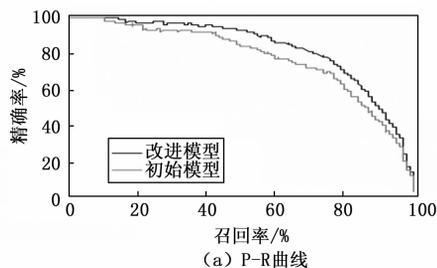


图 8 实验结果曲线图

为了验证各个改进方法对于 YOLOv5 算法的影响, 我们使用每次实验结果产生的 best.pt 文件作为基准权重, 对所采取到的优化措施进行消融实验, 使用 mAP 指标来衡量一个模型的优劣。优化前后 YOLOv5 模型的消融实验结果如表 4 所示。其中初始 YOLOv5 模型的 mAP 为 73.2%, 而优化后的为 79.5%, 综合提高了 6.3%。具体而言, SimAM-CSP 模块的改进使得 mAP 提高了 2.5%, BAM 注意力模块的添加则使得 mAP 提高了 2.63%, 特征融合结构的调整使得 mAP 提高了 0.9%,

表 4 消融实验的结果对比

| SimAM | BAM | Neck | 锚框 | Loss | mAP/% |
|-------|-----|------|----|------|----------------|
| | | | | | 73.2 |
| ✓ | | | | | 75.7 (↑ 2.5) |
| | ✓ | | | | 75.83 (↑ 2.63) |
| | | ✓ | | | 74.1 (↑ 0.9) |
| | | | ✓ | | 76.48 (↑ 3.28) |
| | | | | ✓ | 75.1 (↑ 1.9) |
| | | | ✓ | ✓ | 76.8 (↑ 3.6) |
| ✓ | | | ✓ | ✓ | 77.55 (↑ 4.35) |
| | | ✓ | ✓ | ✓ | 77.38 (↑ 4.18) |
| ✓ | | ✓ | ✓ | ✓ | 78.65 (↑ 5.45) |
| ✓ | ✓ | ✓ | ✓ | ✓ | 79.5 (↑ 6.3) |

将初始的损失函数替换为 $EIoU$ 后, mAP 提高了 1.9%, 而优化锚框后, mAP 提高了 3.28%。总体来看, 每个优化措施都提高了 YOLOv5 算法的性能, 其中初始锚框的优化最为出色。

模型的消融实验结果明确地表明了每一次优化的效果。相较于单独的优化机制, 不同的机制加在一起的效果也会不同。随着损失函数的替换、锚框尺寸的优化、Neck 层的优化、BAM 注意力机制的添加和 SimAM-CSP 的逐渐加入, 模型训练的效果随之逐步提升, mAP 指标最终达到 79.5%。

另外, 我们还将加入 Neck 层中的 BAM 模块与其他注意力机制在同一个数据集和相同训练参数下进行对比训练, 图 9 给出了在不同的注意力机制下的 mAP 指标的对比结果。添加 BAM 模块后的 mAP 比添加 CBAM 模块、ECA 模块、SE 模块和 CA 模块分别高了 1.61%、2.84%、1.49% 和 2.26%。显而易见, BAM 模块的效果比其他模块更加优秀。其中, 添加 ECA 注意力模块后的 mAP 值最低。

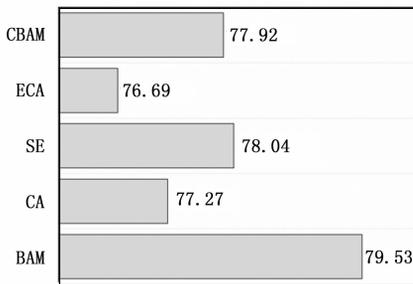


图 9 注意力机制性能对比结果

为了进一步验证我们的模型所具有的优势, 我们分别使用经典目标检测模型 Faster RCNN、SSD 和 YOLOv3 在同一个数据集上进行训练, 并且和优化后的 YOLOv5 模型进行各项指标的对比, 如表 5 所示。

表 5 与其他算法进行的对比

| 模型 | 类别 | AP | P | R | mAP |
|-------------|----|------|------|------|-------|
| Faster RCNN | 齿痕 | 61 | 81.5 | 45.8 | 71.9 |
| | 裂纹 | 82.9 | 83 | 57.9 | |
| SSD | 齿痕 | 68.7 | 86.2 | 30.9 | 72.3 |
| | 裂纹 | 75.9 | 86.8 | 64.2 | |
| YOLOv3 | 齿痕 | 47.6 | 75 | 14.8 | 55.2 |
| | 裂纹 | 62.7 | 89.1 | 33.4 | |
| Ours | 齿痕 | 78.3 | 82.4 | 53.1 | 79.5 |
| | 裂纹 | 80.7 | 85.9 | 69.8 | |

可以得出, 本文所提出的优化后的模型在针对齿痕和裂纹两种舌特征的平均检测精度都要优于其他几种算法, 实现了对 Faster RCNN、SSD 和 YOLOv3 三种经典算法的分别超越, 提升幅度分别达 7.6%、7.2% 和

24.3%。其中, 对于齿痕的检测, 改进后的 YOLOv5 模型对比三者的检测精度分别高了 17.3%、9.6% 和 30.7%, 而对于裂纹的检测, 改进后的 YOLOv5 模型比 SSD 和 YOLOv3 高了 4.8% 和 18%, 但略低于 Faster RCNN。综上所述, 改进后的 YOLOv5 更加适合齿痕和裂纹异常舌特征的检测。

最后, 为了更直观地展示本文模型的检测性能, 使用改进后的模型在测试集中预测, 对于舌齿痕和裂纹的识别效果如图 10 所示。可以看出, 本文的模型可以准确识别出舌头上存在的齿痕和裂纹, 检测不同大小的特征, 进一步验证了改进模型在特征检测任务中的优越性。

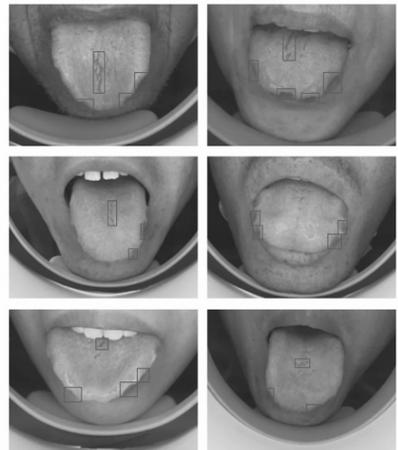


图 10 模型预测结果

5 结束语

我们针对舌面特征自动检测任务, 采取了一系列创新的策略以提升模型的性能, 通过数据增强处理扩充了数据集, 并且对初始的 YOLOv5 算法进行了优化。首先, 提出了 SimAM-CSP 模块, 增强了网络的特征提取能力; 其次, 在 Neck 层添加了 BAM 注意力机制, 聚焦关键信息, 进一步提高检测精度; 然后优化了特征融合结构, 更好地捕捉图像数据的语义信息; 另外使用 $EIoU$ 替换了初始的损失函数 $GIoU$, 提升了模型的训练收敛速度和预测回归精度; 最后, 使用 K-Means 算法对数据集的真实框进行聚类, 调整了初始锚框的尺寸, 使模型更契合齿痕和裂纹检测。实验结果表明, 优化的模型 mAP 要比初始 YOLOv5 的 mAP 高 6.3%, 并且综合优于其他算法, 可以较为准确地检测舌齿痕和裂纹的位置。我们将继续扩充数据集, 加强对其他异常舌形态的研究, 不断引入新的检测目标, 并进一步完善和优化该模型。通过持续的努力和创新, 我们期望能够构建出更加高效可靠的舌面特征自动检测系统, 为中医诊断和现代医学研究提供更加有力的技术支持。

(下转第 146 页)