

基于深度学习的牙根尖片病症识别算法

崔文君¹
CUI Wenjun

摘要

为提高牙根尖片中龋齿和修复体识别的准确性,在传统的YOLOv8模型的基础上,将可变形大核注意力(deformable large kernel attention, D-LKA)引入Head结构中的C2f模块,从而增强模型对目标特征的捕捉能力。同时,将PreLU激活函数代替原模型的SiLU激活函数,解决原模型存在的梯度消失问题,进一步提升了模型的拟合性,使算法运行更加高效。实验结果显示,改进后的算法平均精度mAP0.5比传统算法提高了2.6%,达到了89%;召回率R提高了2.3%,达到了84.8%,为口腔医学领域的进一步研究提供了一定的技术支持。

关键词

牙根尖片;深度学习;目标检测;YOLOv8;C2f_deformable_LKA

doi: 10.3969/j.issn.1672-9528.2025.02.031

0 引言

根据世界卫生组织相关报告和研究表明,口腔疾病在300种常见疾病中发病率位居前列,成为最为普遍的疾病之一,全球约有35亿人受到口腔疾病的困扰^[1]。目前龋齿是最常见的牙齿疾病之一,这是一种慢性病变过程,在早期阶段无明显症状。如果不及时发现治疗,可能会导致疼痛、感染和牙齿脱落等^[2],对身体健康构成威胁。而医学影像技术在牙科病变的早期发现中起着至关重要的作用。

目前对于口腔影像病症的诊断主要有两种方式,第一种是传统的医生人工诊断,但是这种方式由于医生的主观判断可能会导致误诊或漏诊的情况,且成本高、效率低。第二种是现在流行的深度学习方法,这种方法通过对大量已标注的牙齿病症数据进行学习和训练,从而能够在短时间内自动检测并标记目标病症区域,为医生的初步诊断提供线索。例如,Lee等人^[3]在检测根尖周X光片中的龋齿方面使用GoogLeNet Inception v3 CNN算法取得了较好的效果。Thanh等人^[4]应用了Faster R-CNN、YOLOv3、RetinaNet和SSD四种深度学习模型来检测口内非空洞龋齿和空洞龋齿。Ayhan等人^[5]提出了一种改进的CNN-YOLOv7模型在咬翼牙片上进行牙齿编号和龋齿检测,突出CNN同时用于这两种检测的潜在用途。Chen等人^[6]提出一种基于CNN算法的新深度学习集成模型,其中牙位检测准确率为88.8%,牙形检测准确率为86.3%。近来,Ayhan等人^[7]通过使用改进的YOLOv7(YOLOv7+CBAM)模型分析全景X光片来评估在牙科修复体下检测龋齿的有效性,其精确率

达到了83.4%。

虽然这些模型在牙齿影像检测方面取得了不错的成果,但是相较于其他医学影像检测,深度学习在牙科领域的研究却很少,特别是对于龋齿检测和修复体识别方面也更为有限。所以,本文基于YOLOv8模型进行改进,研究改进的YOLOv8算法在牙根尖片上对龋齿和牙根管修复体识别的性能。

1 原YOLOv8算法概述

YOLOv8作为Ultralytics公司在2023年推出的创新模型,相较于YOLOv5,在准确性方面得到了显著提升^[8],这使得YOLOv8成为一个功能更为全面的模型,能够胜任目标检测、图像分割以及分类等多种任务。为满足不同应用场景的需求,YOLOv8提供了5个不同规模的模型。

为进一步优化计算效率和减少网络流量负担,YOLOv8在架构设计上进行了多项改进。具体而言,YOLOv8在Backbone部分采用结合了CSPNet(cross stage partial network)思想的C2f模块^[9]替代YOLOv5中的C3模块,C2f模块通过将不同尺度的特征图进行融合,这有助于模型同时获取高分辨率和丰富的语义信息。此外,YOLOv8还保留了YOLOv5中的快速空间金字塔池化(SPPF)模块,这也有助于实现多尺度特征的有效融合。

在Neck部分,YOLOv8借鉴YOLOv5中的FPN+PAN结构^[10],并通过自上向下与自下向上的跨层连接,巧妙融合高层强语义特征和底层定位信息。YOLOv8通过进一步简化上采样阶段的卷积结构,降低了计算成本。另外,YOLOv8头部使用了解耦头部技术将检测与分类分开,使模型更加专

1. 天津职业技术师范大学电子工程学院 天津 300222

注于不同的任务，以提高目标检测精度。并且 YOLOv8 采用 Anchor-Free，可以减少模型对先验框的依赖，提高模型对不规则形状的检测精度。

2 基于改进的 YOLOv8 算法介绍

针对传统 YOLOv8 算法存在漏检高、准确率低等问题，提出一种改进的 YOLOv8n 网络模型。在原 YOLOv8 模型中将 Head 结构中的 C2f 结合可变形大核注意力机制，即 C2f_deformable_LKA 模块，能够动态调整卷积核的形状和大小来灵活地改变采样网格，从而增强模型对不同牙齿病症形状和大小的适应性。同时，通过将原 YOLOv8 的 SiLU 激活函数替换为 PreLU 激活函数，有效解决了原模型中存在的梯度消失现象。改进后的 YOLOv8 模型结构如图 1 所示。

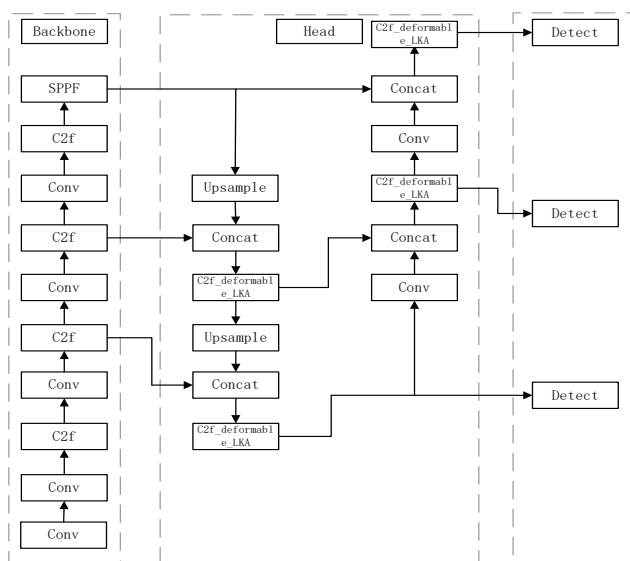


图 1 改进的 YOLOv8 模型结构

2.1 可变形大核注意力机制（D-LKA）原理

在处理目标检测、语义分割等任务时需要模型能够准确的捕捉到图像中的关键信息。然而，传统的卷积神经网络在处理这些任务时往往受限于固定的感受野大小和卷积核形状，这在一定程度上限制了模型捕捉图像中关键信息的能力。特别是牙根尖片中病症目标大小不一、标签分布不平衡等问题，需要模型具备更强的局部特征捕捉能力。因此，通过引入 D-LKA^[11]，以增强模型对牙齿病症局部特征的提取和关注。D-LKA 是一种简化的注意力机制，这种注意力的核心思想在于将大卷积核与可变形卷积的注意力机制结合在一起^[12]。

2.1.1 大核注意力机制

为克服受限于固定的感受野大小问题，大核注意力机制（LKA）^[13]采用更大的卷积核来捕捉更广泛的上下文信息。然而，直接使用大卷积核会导致计算量剧增，因此大核注意

力机制通过使用深度卷积、深度空洞卷积和 1×1 卷积，用更少的参数和更低的计算量来构建大卷积核。

2.1.2 可变形卷积

可变形卷积的核心思想是引入额外的偏移量，使卷积核在采样时能够根据图像特征自适应调整，增加模型对几何变换的鲁棒性。这些偏移量通常通过额外的卷积层生成，并与输入特征图具有相同的空间分辨率。在卷积过程中，模型会根据偏移量对卷积核的采样位置进行调整，从而使卷积核能够更准确地捕捉目标的形状和边界信息。这种自适应的采样方式显著提升了模型对不规则形状目标的检测能力。D-LKA 网络结构如图 2 所示。

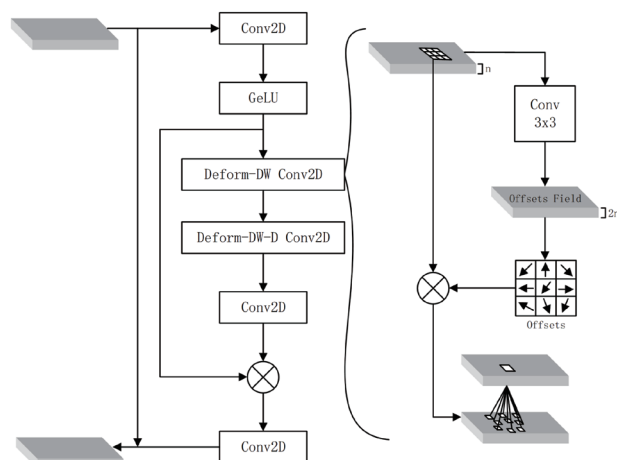


图 2 D-LKA 网络结构

D-LKA 模块公式分别为：

$$Attention = Conv_{1 \times 1} \left(Conv_{DDW-D} \left(Conv_{DDW}(F) \right) \right) \quad (1)$$

$$Output = Conv_{1 \times 1} (Attention \otimes F') + F \quad (2)$$

式中：输入特征用 $F \in \mathbb{R}^{C \times H \times W}$ 和 $F' = GeLU(Conv(F))$ 表示；Output 为输出特征。其中，GeLU 是一种激活函数， F' 为经过卷积和激活函数后的特征图^[14]。注意分量 $\mathbb{R}^{C \times H \times W}$ 表示为注意力图，其中每个值表示对应特征的重要性。Conv_{DDW-D} 和 Conv_{DDW} 表示不同类型的深度卷积操作，符号 $\otimes$ 表示逐元素乘积操作。

2.2 PreLU 激活函数

原始 YOLOv8 算法中使用 SiLU 激活函数，如式 (3) (4) 所示。SiLU 函数结合了 Sigmoid 和 ReLU 函数的优点，提供了平滑的非线性变换。虽然在广泛的数据集和应用上效果显著，但在本次实验使用的数据集中效果不佳，存在梯度消失现象，而 PreLU 激活函数能够有效缓解这一问题。因此，为了提升目标识别精度，将 SiLU 函数替换为参数化修正线性单元 PreLU。

$$SiLU(x) = x \cdot Sigmoid(x) \quad (3)$$

$$\text{Sigmoid}(x) = \frac{1}{1 + e^{-x}} \quad (4)$$

PReLU 激活函数是 ReLU 函数的优化, 由 He 等人^[15]在 2015 年提出, 旨在解决 ReLU 激活函数的一些局限性。PReLU 公式为:

$$\text{PReLU}(x) = \begin{cases} x, & x \geq 0 \\ kx, & x < 0 \end{cases} \quad (5)$$

由公式可以看出, PReLU 有一个可学习的负斜率参数 k , 相较于 ReLU, 其主要优势在于克服了 ReLU 的死神经元的问题, 使 PReLU 在输入为负值时也能拥有非零的导数, 从而促进神经元能够更好地学习, 并且网络能够根据数据自适应地调整 k ^[16], 在有效缓解梯度消失的同时增强了网络的灵活性。PReLU 示意图如图 3 所示。

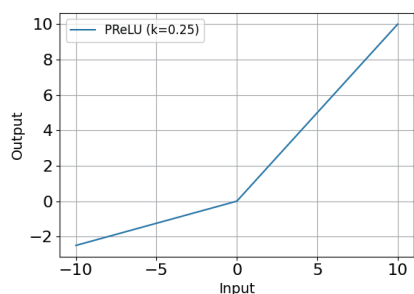


图 3 PReLU 激活函数示意图

3 实验及结果分析

3.1 实验环境

实验在 Ubuntu 20.04 操作系统上开展, 编程语言选用 Python 3.8, 同时使用 PyTorch1.11.0 来构建深度学习模型。模型训练上, 使用单显卡 NVIDIA GeForce RTX 3090 24 GB GPU 作为训练平台, 并通过 CUDA11.3 进行 GPU 加速。

3.2 数据集和预处理

为了验证本文牙齿病症检测模型的有效性, 从 kigggle 数据集上随机选取 910 张公开数据集, 涵盖多种牙齿病症。后使用 labeling 工具对数据集进行仔细标注, 本实验选择龋齿和牙根管修复体为研究目标, 并确定两种病症的标签分别为 Caries 和 RCT。最后将数据集按照 4:1 的比例, 采用动态分割方法, 将数据集划分为训练集、验证集。其中, 训练集 728 张图像, 用于模型的训练和学习; 验证集 182 张图像, 用于评估模型性能。

3.3 评估指标

实验选用准确率 P 、召回率 R 、平均精度均值 mAP (IoU 阈值取 0.5 和 0.5~0.95) 来评估模型。数值越高, 表明模型的检测效果越好, 其中 P 、 R 、mAP 的计算公式^[17]分别为:

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

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

$$AP = \int_0^1 P(R) dR \quad (8)$$

$$mAP = \frac{1}{C} \sum_{i=1}^C AP_i \quad (9)$$

式中: 真阳性 (TP) 指模型正确检测出牙齿病症区域; 假阳性 (FP) 指模型将非病症区域误检测为病症区域; 假阴性 (FN) 指模型未能检测到实际存在的牙齿病症区域^[18]; AP_i 表示类别索引值为 i 的 AP 值; C 为训练样本的总类别数。

3.4 实验结果分析

本次实验对数据集目标进行了 300 轮次训练, 如图 4 所示为改进前后 mAP@0.5 曲线对比图, 相较于原始的 YOLOv8, mAP 得到了明显的提升。表明本文算法在提升牙齿病症检测精度上的有效性。

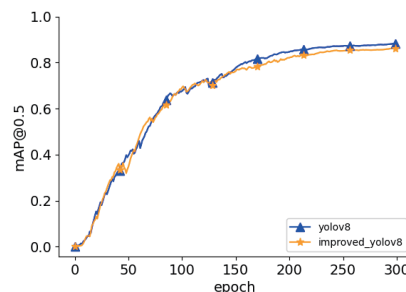


图 4 YOLOv8 与改进的 YOLOv8 的 mAP@0.5 性能对比图

消融实验结果如表 1 所示。由表 1 可知, 在修改 PReLU 激活函数后各项指标除召回率 R 外均有提升, 此激活函数确实能够解决梯度消失问题。当使用可变形大核注意力机制时, P 和 mAP 也均有所提高, 说明 D-LKA 能有效的根据图像特征自适应调整算法, 提高性能。而 PReLU 和 D-LKA 结合后的算法, 在 P 和 R 之间取得了更好的平衡, 其平均精度均值达到了 89%, 召回率达到了 84.8%, 分别提升了 2.6% 和 2.3%。

表 1 消融实验结果

单位: %

实验算法	mAp@0.5	mAp@0.5:0.95	P	R
YOLOv8n	86.4	55.7	85.4	82.5
YOLOv8n+P	88.3	57	94.2	76.5
YOLOv8n+D	86.4	56	86.8	78.8
YOLOv8n+P+D	89	56	85.2	84.8

图 5 为可视化效果对比图, 能够更直观地展示改进前后算法性能差异。其中左侧为原算法的检测图, 右侧为改进后算法的检测图。观察对比可知, 改进后的算法在每个类别检测中的准确率均有显著提升, 均优于原检测算法。

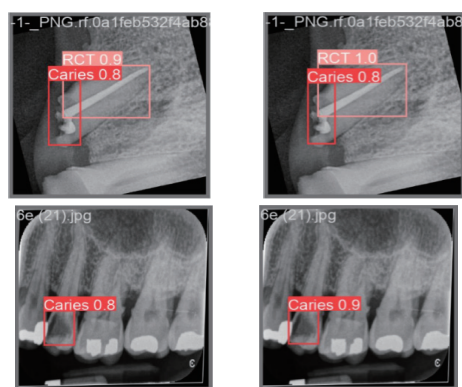


图5 改进前后可视化对比

4 结论

本文专注于口腔医学领域, 改进了基于 YOLOv8n 算法的网络结构。在数据收集阶段, 随机选取不同拍摄角度的牙片图像, 提高模型在真实临床环境中的适用性和鲁棒性。在网络结构改进阶段, 将 D-LKA 引入到 C2f 模块中, 能够更有效聚焦于目标病症的关键特征。同时, 采用 PReLU 激活函数解决梯度消失问题。实验结果表明, 改进后的算法 mAP0.5、mAP0.5:0.95 和 R 分别达到了 89%、56%、84.8%, 相较于原始算法分别提升了 2.6%、0.3% 和 2.3%。虽然针对病症识别准确率有显著提升, 但是改进后的算法运行速率有所降低, 下一步将研究如何使模型更具轻量化, 并应用于具体实际中。

参考文献:

- [1] LI K C, MAO Y C, LIN M F, et al. Detection of tooth position by YOLOv4 and various dental problems based on CNN with bitewing radiograph [J]. IEEE access, 2024,12:11822-11835.
- [2] BAYATI M, SAVAREH B A, AHMADINEJAD H, et al. Advanced AI-driven detection of interproximal caries in bite-wing radiographs using YOLOv8[J]. Scientific reports, 2025, 15(1): 4641.
- [3] LEE J H, KIM D H, JEONG S, et al. Detection and diagnosis of dental caries using a deep learning-based convolutional neural network algorithm[J]. Journal of dentistry, 2018, 77(10): 106-111.
- [4] THANH M T G, TOAN N V, NGOC V T N, et al. Deep learning application in dental caries detection using intraoral photos taken by smartphones[J]. Applied sciences, 2022, 12(11): 5504.
- [5] AYHAN B, AYAN E, BAYRAKTAR Y. A novel deep learning-based perspective for tooth numbering and caries detection[J]. Clinical oral investigations, 2024, 28(3): 178.
- [6] CHEN C C, WU Y F, AUNG L M, et al. Automatic recognition of teeth and periodontal bone loss measurement in digital

radiographs using deep-learning artificial intelligence[J]. Journal of dental sciences, 2023, 18(3): 1301-1309.

- [7] AYHAN B, AYAN E, ATSÜ S. Detection of dental caries under fixed dental prostheses by analyzing digital panoramic radiographs with artificial intelligence algorithms based on deep learning methods[J]. BMC oral health, 2025, 25: 216.
- [8] SUN Z, ZHU L X, QIN S, et al. Road surface defect detection algorithm based on YOLOv8[J]. Electronics, 2024,13(12): 2413.
- [9] 组少杰. 基于深度注意力模型的 YOLOv8 煤矸石检测算法[J]. 煤矿现代化, 2025,34(2):64-69.
- [10] 杨博, 胡珍珍. 基于 YOLOv8n 改进算法的自动驾驶目标检测 [J/OL]. 控制工程, 1-7[2024-05-16].<https://doi.org/10.14107/j.cnki.kzgc.20240494>.
- [11] AZAD R, NIGGEMEIER L, HÜTTEMANN M, et al. Beyond self-attention: deformable large kernel attention for medical image segmentation[DB/OL].(2023-08-31)[2024-06-19].<https://doi.org/10.48550/arXiv.2309.00121>.
- [12] WANG T, ZHANG H, JIANG D. CSD-YOLO: a ship detection algorithm based on a deformable large kernel attention mechanism[J]. Mathematics, 2024, 12(11): 1728.
- [13] GUO M H, LU C Z, LIU Z N. et al. Visual attention network[DB/OL]. (2022-07-11)[2024-04-19].<https://doi.org/10.48550/arXiv.2202.09741>
- [14] 李营安, 彭龔. 基于改进 YOLOv8n 木材表面缺陷检测算法研究 [J]. 黑龙江工业学院学报 (综合版), 2024, 24(11): 86-93.
- [15] HE K M, ZHANG X Y, REN S Q, et al. Delving deep into rectifiers:Surpassing human-level performance on imagenet classification[C]//Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV), Piscataway:IEEE, 2015: 1026-1034.
- [16] 包从望, 朱广勇, 邹旺, 等. 基于 SimAM 注意力机制的轴承故障迁移诊断模型 [J]. 机电工程, 2024,41(5): 862-869.
- [17] CHEN H W, ZHOU G H, JIANG H X. Student behavior detection in the classroom based on improved YOLOv8[J]. Sensors, 2023, 23(20): 8385.
- [18] 刘晨雨, 徐健, 李轲, 等. 基于改进 YOLOv8n 的儿童肠套叠 B 型超声图像特征检测 [J]. 生物医学工程学杂志, 2024, 41(5):903-910.

【作者简介】

崔文君 (2000—), 女, 山东淄博人, 硕士研究生, 研究方向: 人工智能。

(收稿日期: 2025-02-13)