

基于深度学习的牙齿病变自动检测算法

刘丰¹, 韩民^{1*}, 万军¹, 刘超^{2**}¹ 山东大学信息科学与工程学院, 山东 青岛 266237;² 山东大学齐鲁医院口腔颌面外科, 山东 济南 250012

摘要 龋病、牙周病等口腔疾病是影响人民健康的常见病和多发病, 不仅影响口腔器官功能的发挥, 还常常影响全身健康, 导致生活质量下降。X 光片是牙科疾病诊断过程中的重要依据之一, X 光片的智能化诊断对于快速准确诊断牙齿病变具有重要作用。为了实现龋齿病变和牙根尖周病变的自动检测, 本团队创建了牙齿 X 光片数据集, 并采用 YOLOV5 算法对牙齿 X 光片中的病变区域进行了检测。检测结果表明, 该算法可以有效识别牙齿中的龋齿病变和牙根尖周病变, 并能检测出这两种病变的区域, 检测准确率超过 95%, 可以满足牙齿病变智能化诊断的临床需求。

关键词 图像处理; 深度学习; 牙齿病变; 目标检测

中图分类号 TP181 文献标志码 A

DOI: 10.3788/CJL202249.2007207

1 引言

自 2019 年新冠病毒肺炎疫情暴发以来, 医院就成为人员聚集程度最严重的地点。在高风险地区, 部分医院一直在超负荷运转, 大量病人前往医院就诊。这样的人员聚集不仅加重了病人之间交叉感染的风险, 更是导致了公共医疗系统的超负荷运转, 轻症病和慢性病无法得到及时有效处理, 最终发展为重症疾病, 其中就包括牙科疾病。牙科数字化 X 光片判读作为牙科疾病诊断过程中的重要一环, 对于牙科疾病的诊断具有重要意义^[1]。因此, 基于深度学习准确、高效地利用 X 光片的图像特征实现病变的自动诊断, 对于缓解疫情期间就诊压力、减少人员聚集具有重要意义。

随着患者数量增加以及医学成像技术的发展, 牙科 X 光片数量骤增, 为深度学习^[2]在牙科领域的应用提供了数据保障。按照任务的不同, 可将深度学习应用于牙科疾病诊断的研究分为两类: 1) 病变类型分类, 即, 根据图像特征判断该图像中病变的所属类型; 2) 病变区域目标检测, 即, 在判断病变类型的同时对病变发生的位置进行标定^[3]。

目前, 对病变类型进行分类使用的主流网络是卷积神经网络 (CNN)^[4]。2018 年, Lee 等^[5]利用 InceptionV3 网络^[6]对牙科 X 光片中的龋齿病变进行识别, 分类识别准确率可达 89%, 取得了良好的实验结果。同年, 该课题组^[7]使用 VGG-19 网络^[8]对牙科 X 光片中的牙周病变进行识别, 识别准确率为

82.8%。2019 年, Krois^[9]等使用 7 层 CNN 对牙周骨缺损病变进行了检测, 分类识别准确率可达 81%。针对病变区域目标进行检测的主流网络结构包括 Region-CNN^[10] (R-CNN) 系列网络、Single Shot MultiBox Detector^[11] (SSD)、U-Net^[12]。2019 年, Chen 等^[13]使用 Faster RCNN^[14]进行牙齿编号, 以判断是否存在牙齿缺损, 识别精确率和召回率均为 80% (这是因为 Faster RCNN 使用 RoI Pooling 算法取整时会带来一定的精度损失)。2022 年, Zhang 等^[15]首次使用 SSD 对口腔摄影图像中的龋齿病变进行了检测, 但 prior box 的基础大小和形状不能直接通过学习获得, 需要手工设置。2021 年, Khan 等^[16]首次使用 U-Net 进行龋齿、牙槽骨吸收等病变的检测, U-Net 特有的 U 型网络使得分类精度和定位精度不可兼得。

You Only Look Once^[17] (YOLO) 算法具有检测速度快、检测效果好等优点。为了提高牙科病变区域检测的准确性, 本团队使用 YOLOV5 算法对牙齿数字化 X 光片上的病变区域进行检测。检测结果表明, YOLOV5 算法在牙齿 X 光片病变检测任务中表现出了更高的识别准确率, 使得患者自助诊断牙科疾病成为可能。

2 基本原理

2.1 网络介绍

YOLOV5 是一种单阶段目标检测算法, 可以同时实现类别和目标区域的检测, 具有检测准确率

高、检测速度快、检测效果好等优点。网络分为 4 部分,分别是输入端(Input)、基准网络(Backbone)、

特征融合(Neck)和输出端(Head),网络结构如图 1 所示。

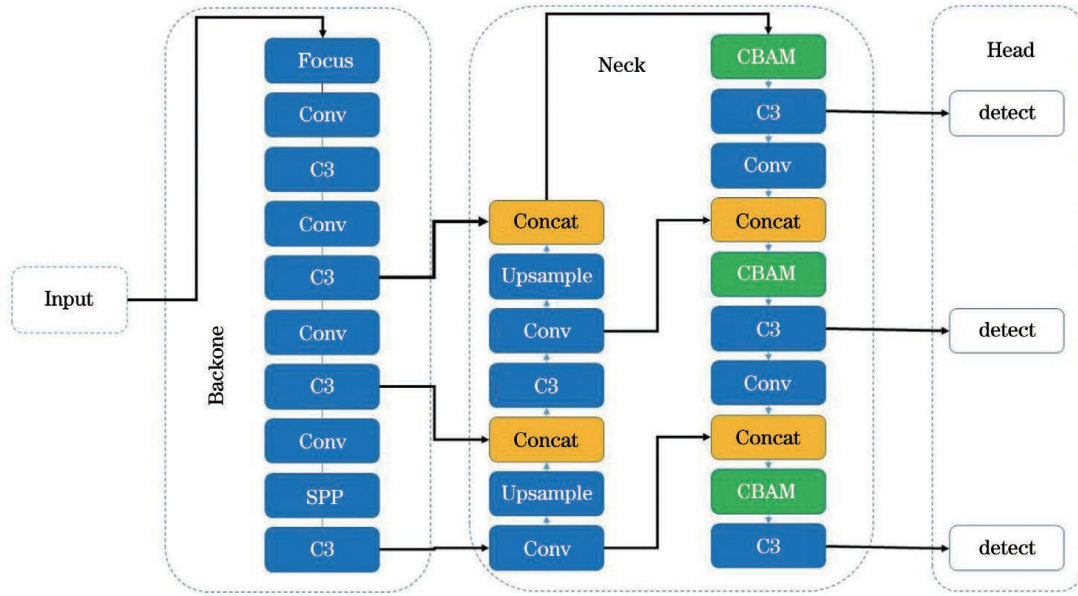


图 1 YOLOV5 结构

Fig. 1 YOLOV5 structure

Input 模块使用 Mosaic 数据增强和自适应锚框两种先进算法。其中:Mosaic 数据增强可以对输入网络的图像进行随机缩放、重新裁剪和排布,以便扩增数据集,提高网络模型的鲁棒性^[18];自适应锚框根据训练数据(图片)的尺寸自动计算出最佳 Anchors 的大小,以便更好地定位目标。

Backbone 模块新增了 Focus 部分,并新设计了一种 CSP(Cross Stage Partial)结构。Focus 部分用于图

片切片操作,实现图像特征的提取^[19]。CSP 结构首先将传入的图像特征分为两部分,一部分图像特征不做处理直接传入 Neck 模块,另一部分图像特征首先经过一次 CBL (Conv + Batch-Normlization + Leaky_ReLU)操作,接着进行降维的卷积和激活,最后再进行一次 CBL 操作,得到处理后的图像特征。下一阶段的 Neck 模块将两个特征进行融合,增强 CNN 网络的学习能力,同时保证准确率。CSP 结构图如图 2 所示。

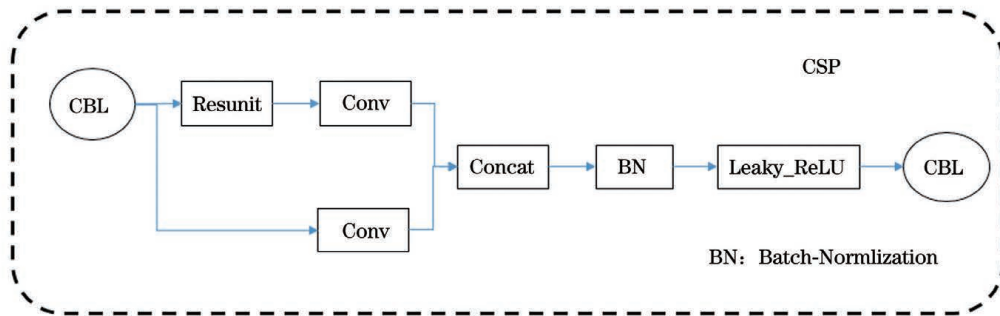


图 2 CSP 结构

Fig. 2 CSP structure

CBL 操作指 Conv 卷积、BN(Batch-Normlization)正则化和 Leaky_ReLU 函数激活操作;Resunit 指残差结构,其作用是增加网络反向传播时的梯度值,避免梯度消失,防止网络退化,提取更详细的图像特征;Concat 的作用是在 Conv 卷积后增加通道数;BN 正则化可以加快收敛速度;Leaky_ReLU 是激活函数,可以加快收敛和计算。

Neck 模块将 FPN(feature pyramid networks)^[20]结构与 PAN(pixel aggregation networks)^[21]结构相结合,用于实现网络特征融合,如图 3 所示。

FPN 结构自顶向下通过上采样进行强语义特征融合,而 PAN 结构自底向上进行强定位特征融合,两者相互补充,加强了网络特征的融合能力。

Head 模块采用 GIOU_Loss^[22]损失函数和非极大值抑制(NMS)^[23]两种方式来实现结果的回归。其中:GIOU_Loss 损失函数引入目标框和预测框的最小外接矩形面积参与计算,使网络能够在目标框和预测框不相交时也可以继续优化,并能够判断两个框的位置信息,得到更好的检测锚框;非极大值抑制可以判断框内是否存在多个物体重合。

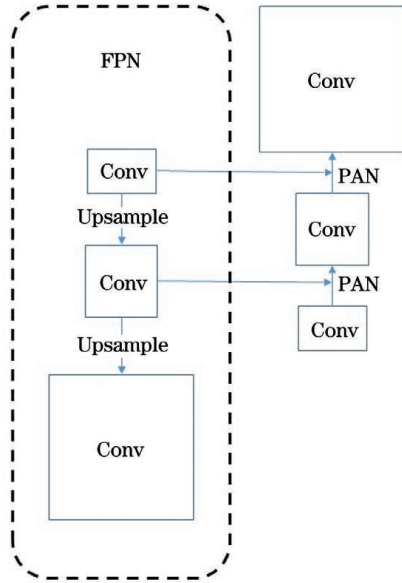


图 3 FPN+PAN 实现网络特征融合

Fig. 3 Network feature fusion realized with FPN+PAN

2.2 网络设置优化

为了获得效果更好的检测目标区域,针对构建的牙科病变数据集,本文通过以下 4 种措施来设置网络。

2.2.1 名义批次大小调整

将网络中的名义批次大小(NBS)设置为 64,同时将超参数 batch_size 设置为 32。此时,由这两个数值的比值为 2 可以得出,每迭代两次,网络才进行一次反向传播。这样,GPU 上的显存利用率提高了,矩阵乘法的利用率随之提高。

2.2.2 权重参数调整

YOLOV5 网络中的 PyTorch 框架存在一个缺陷,如果使用其自带的优化器进行权重参数的正则化,它就会将偏置参数一起正则化,这会降低模型的灵活性,而且更易引起欠拟合。所以,将模型可学习的参数分为三组,即偏置参数、权重参数和 BN 参数,这样在对权重参数进行正则化时,就可以避免同时对偏置参数正则化。

2.2.3 优化器替换

将原有的随机梯度下降(SGD)优化器替换为自适

应矩估计(Adam)优化器,这样可以更容易找到合适的权重,使权重的衰减更加平滑。这是因为 Adam 优化器基于一阶矩估计和二阶矩估计自适应调节每个权重参数的大小。一阶矩估计和二阶矩估计的计算公式分别为

$$m_t = \mu \times m_{t-1} + (1 - \mu) \times g_t, \quad (1)$$

$$n_t = \nu \times n_{t-1} + (1 - \nu) \times g_t^2, \quad (2)$$

式中: m_t 和 n_t 分别表示对梯度的一阶矩估计和二阶矩估计; g_t 表示梯度的大小; μ 和 ν 分别表示一阶矩估计和二阶矩估计的动量因子。对学习率进行偏置校正,校正公式为

$$\hat{m}_t = \frac{m_t}{1 - \mu^t}, \quad (3)$$

$$\hat{n}_t = \frac{n_t}{1 - \nu^t}, \quad (4)$$

最终得到学习率范围的计算公式为

$$\Delta\theta_t = -\frac{\hat{m}_t}{\sqrt{\hat{n}_t} + \epsilon} \times \eta, \quad (5)$$

式中: \hat{m}_t 和 \hat{n}_t 分别表示偏置校正后的 m_t 和 n_t ; η 为学习率; $\Delta\theta_t$ 为权重动态变化的范围; ϵ 是为避免分母为 0 而设置的数,其取值为 10^{-8} ,不影响算法表现。

2.2.4 学习率衰减策略调整

将学习率衰减策略更改为一维线性插值算法和余弦退火算法相结合的策略,其中:在 warm-up 阶段用一维线性插值算法预热学习率,使学习率从很小的数值开始增大;warm-up 阶段结束后,采用余弦退火算法更新学习率,使之保持下降趋势。余弦退火算法的表达式为

$$\eta_t = \eta_{\min}^i + \frac{1}{2}(\eta_{\max}^i - \eta_{\min}^i) \left[1 + \cos\left(\frac{T_{\text{cur}}}{T_i} \pi\right) \right], \quad (6)$$

式中: i 表示索引值; η_t 表示新得到的学习率; η_{\min}^i 和 η_{\max}^i 是第 i 次训练的最小学习率和最大学习率; T_{cur} 是当前的 epoch; T_i 是总的 epoch。

图 4 为学习策略的收敛性比较。采用余弦退火策略后,学习率的衰减更加平滑,模型的收敛速度更快,收敛效果更好;采用一维线性插值策略后,学习率表现为

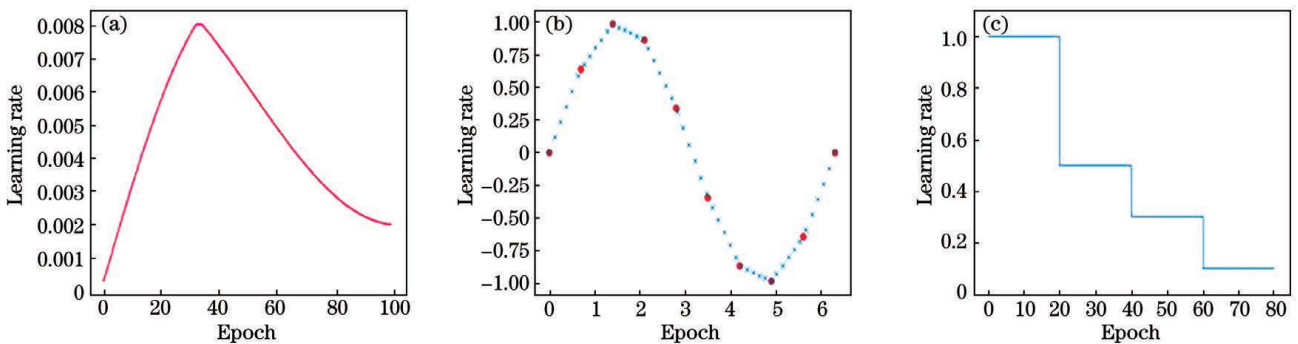


图 4 学习策略收敛性比较。(a)余弦退火策略;(b)一维线性插值策略;(c)阶梯下降策略

Fig. 4 Comparison of learning rate strategy convergence. (a) Cosine annealing strategy; (b) one-dimensional linear interpolation strategy; (c) step reduction strategy

离散的数据点;采用阶梯下降策略后,学习率曲线可以看成是分段函数。对比三者可以发现,余弦退火策略可以使学习率衰减得更加平滑,收敛速度更快。

2.3 评价指标

目标检测算法的评价指标有很多种,本文使用精确率(P)、召回率(R)、准确率(A)、F1 分数(F_1)以及均值平均精度(mAP)作为评价指标。

精确率 P 是用来评价二分类网络的指标。对于 YOLO 算法,精确率 P 更直观的解释是:在所有检测出的边界框中,检测正确的边界框所占的比例。该计算指标的计算公式为

$$P = \frac{S_{TP}}{S_{TP} + S_{FP}}, \quad (7)$$

式中: S_{TP} 表示真阳性样本的数量; S_{FP} 代表假阳性样本的数量。

在目标检测算法中,召回率 R 更直观的解释是:检测正确的边界框的数量与真实边界框数量的比值。该指标的计算公式为

$$R = \frac{S_{TP}}{S_{TP} + S_{FN}}, \quad (8)$$

式中: S_{FN} 代表假阴性样本的数量。

准确率 A 是指在所有样本中检测正确的边界框的数量所占的比值,计算公式为

$$A = \frac{S_{TP} + S_{TN}}{S_{TP} + S_{TN} + S_{FP} + S_{FN}}, \quad (9)$$

式中: S_{TN} 表示真阴性样本的数量。

F_1 被定义为精确率和召回率的调和平均数,综合了两者的产出结果,其计算公式为

$$F_1 = \frac{2PR}{P + R}. \quad (10)$$

mAP 是目标检测网络训练中用来观察网络性能的指标之一,mAP 的值越接近 1,说明目标类别越符合实际,目标框越贴合物体边缘。mAP(记为 I_{mAP})的计算公式为

$$I_{mAP} = \frac{1}{N} \int_0^1 P(R) dR, \quad (11)$$

式中: N 表示样本的类别数; $P(R)$ 表示当召回率为 R 时对应的精确率。

3 实验结果分析

3.1 实验准备

本次实验均在 Windows 10 系统下运行完成,编程语言使用 Python,编译器使用 Pycharm,网络框架选择 Pytorch1.7.1。实验所用计算机处理器型号为 intel CORE i5-10400,GPU 型号为 NVIDIA RTX 2060。

本次实验使用的数据集是由专业医师收集的牙齿 X 光片,而且这些 X 光片均使用 LabelImg 工具进行标注,分为龋齿区域、牙根尖周病变区域和健康牙齿区域。训练数据包含 23884 张图像,测试数据包含 5970

张图像,在训练数据中选取 20474 张图像作为训练集,余下的 3410 张图像作为验证集。部分图像如图 5 所示。

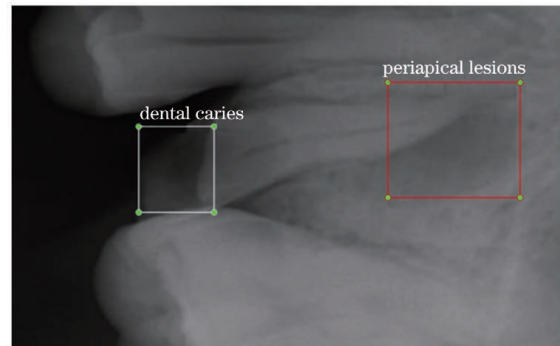


图 5 牙齿 X 光片和两种病变类型

Fig. 5 Dental X-ray image and two types of lesions

3.2 实验结果与评价

3.2.1 实验结果

将本文算法(YOLOV5)、SSD、Faster-RCNN 目标检测算法在创建的数据集上进行训练和测试,测试结果如表 1 所示。本文算法的精确率为 98%,召回率为 95%,准确率为 95%,F1 分数为 96%,各项评价标准均高于 SSD、Faster-RCNN 目标检测算法的评价指标,说明本文算法明显优于其他算法。

表 1 不同算法的评价指标对比

Table 1 Evaluation indices comparison of different algorithms

Algorithm	$P / \%$	$R / \%$	$A / \%$	F1 score / %
YOLOV5	98	95	95	96
SSD	91	65	69	76
Faster-RCNN	66	70	63	68

3.2.2 结果评价

算法性能的好坏通常体现在训练过程中。本文算法的训练过程用边界框损失、分类损失、置信度损失、精确率、召回率来评价。通过观察图 6 可知,边界框损失和分类损失分别在第 90 次和第 20 次 epoch 时开始平滑,训练过程已经收敛且趋于 0。这说明网络模型已学习到必要的特征,模型预测值和真实值的差距很小,性能良好。

由式(11)可知,如果以精确率 P 为纵坐标、以召回率 R 为横坐标构建 P - R 曲线,那么对应的 mAP 值就是每个类别的 P - R 曲线下方面积的平均值。本文算法的 P - R 曲线如图 7 所示。由图 7 可知网络训练的 mAP 值为 0.985,验证了网络训练符合本研究要求。实际的检测结果如图 8 所示,caries 表示检测为龋齿区域,cyst 表示检测为牙根尖周病变区域。可见,检测效果比较理想。

3.2 热力图可视化

神经网络在学习过程中会根据图片不同区域的不同特征来调整各个阶段参数的大小,进而实现目标检

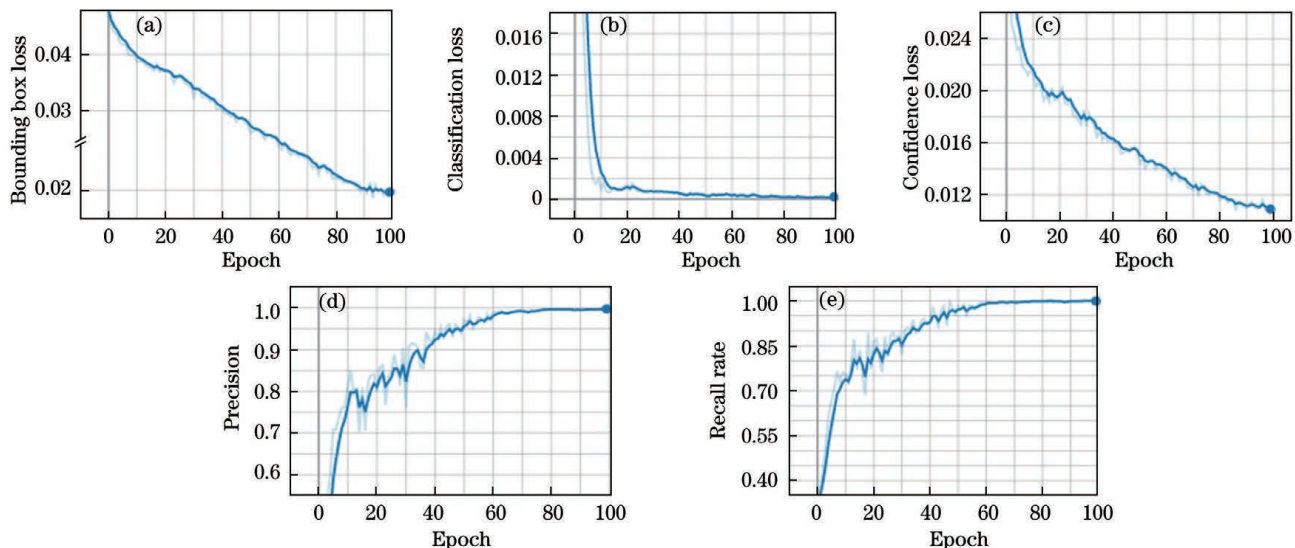


图 6 训练结果展示。(a)边界框损失;(b)分类损失;(c)置信度损失;(d)精确率;(e)召回率

Fig. 6 Training results display. (a) Bounding box loss; (b) classification loss; (c) confidence loss; (d) precision; (e) recall rate

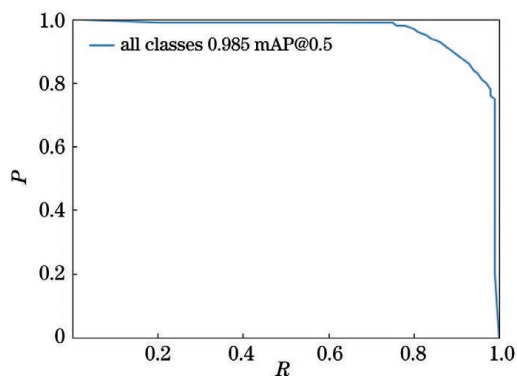


图 7 P-R 曲线

Fig. 7 P-R curve

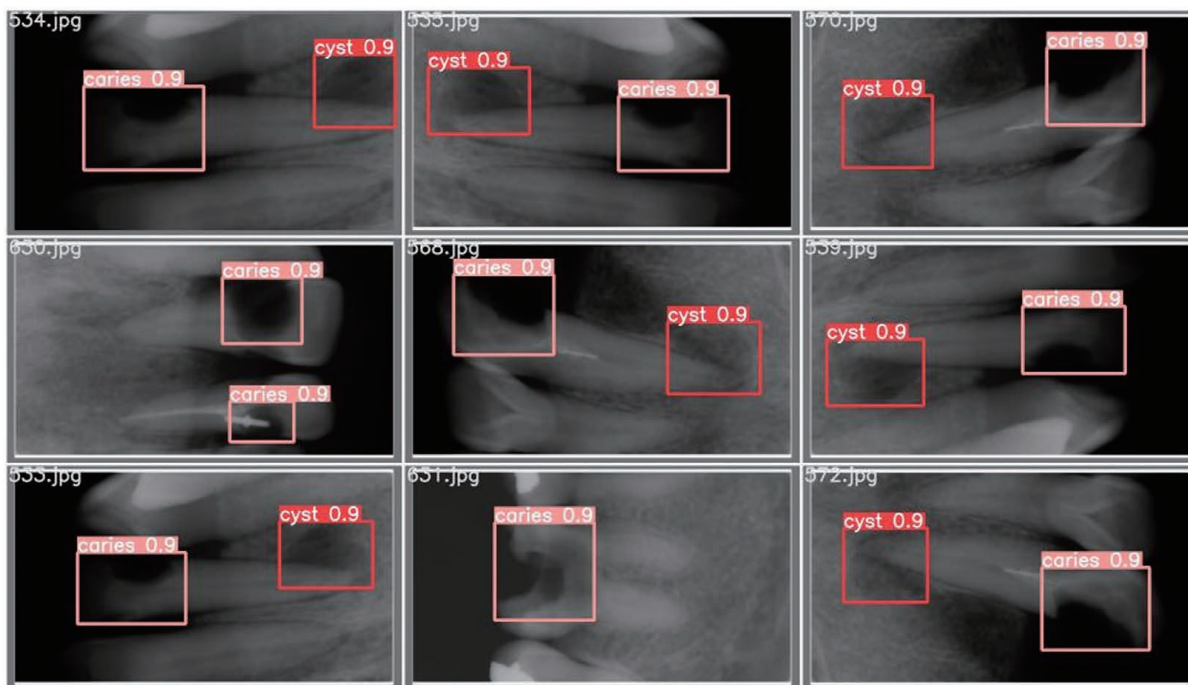


图 8 结果示例, caries 标签表示检测为龋齿部分, cyst 标签表示检测为牙根尖周病变部分

Fig. 8 Examples, where caries label represents the are is caries and cyst label represents the area is periapical lesion

测。但由于迭代过程中的不可观测性,无法得知图片中的哪一部分最能刺激网络。因此,可视化在网络评价中起到了重要作用,而热力图可以将网络模型对图片不同部分的感兴趣程度用颜色表示出来。网络模型越感兴趣的部分,网络参数越大,颜色越偏向于红色。图 9(b)为牙齿图片的可视化热力图,红色部分为网络模型感兴趣区域,同时也为目标检测区域。图 9(b)右半部分有类似特征部分也有激活网络,但区域较小,特征并不明显,所以没有检测为目标区域。

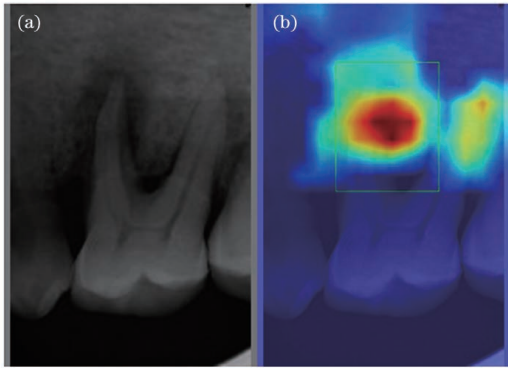


图 9 可视化热力图。(a)原图;(b)热力图

Fig. 9 Visualization of thermodynamic diagram. (a) Original drawing; (b) thermodynamic diagram

4 结 论

本团队使用 YOLOV5 算法对牙齿 X 光片中的病变区域进行检测,在数据集上进行训练和测试,同时修改了网络名义批次的大小,使用了更合适的优化器,调整了权重参数,修改了学习率衰减策略。此外,将 YOLOV5 算法与他人所用算法进行了对比。最后,基于可视化的热力图分析了网络模型的特征提取效果。实验结果表明,YOLOV5 算法可以检测出 X 光片上的大多数病变区域,算法的各项性能指标均有良好表现,在牙科医学智能化临床诊断方面具有一定的实际应用价值。

参 考 文 献

- [1] 万阔, 景泉. 交流、融合、进步: 从美日两国牙科镇静镇痛发展现状看中国口腔舒适化未来[J]. 中国口腔医学继续教育杂志, 2021(1): 31-37.
Wan K, Jing Q. Communication, integration and progress—looking at the future of oral comfort in China from the current situation of dental sedation and analgesia in the United States and Japan [J]. Chinese Journal of Stomatological Continuing Education, 2021(1): 31-37.
- [2] Hinton G E, Salakhutdinov R R. Reducing the dimensionality of data with neural networks [J]. Science, 2006, 313 (5786): 504-507.
- [3] LeCun Y, Bottou L, Bengio Y, et al. Gradient-based learning applied to document recognition [J]. Proceedings of the IEEE, 1998, 86(11): 2278-2324.
- [4] Manjula K, Devi G D, Vijayarekha K. Detection of dental diseases using classification algorithms[J]. International Journal of Recent Technology and Engineering, 2019, 8(3): 4485-4489.
- [5] Lee J H, Kim D H, Jeong S N, et al. Detection and diagnosis of dental caries using a deep learning-based convolutional neural network algorithm [J]. Journal of Dentistry, 2018, 77: 106-111.
- [6] Szegedy C, Vanhoucke V, Ioffe S, et al. Rethinking the inception architecture for computer vision [C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 2818-2826.
- [7] Lee J H, Kim D H, Jeong S N, et al. Diagnosis and prediction of periodontally compromised teeth using a deep learning-based convolutional neural network algorithm [J]. Journal of Periodontal & Implant Science, 2018, 48(2): 114-123.
- [8] 舒朗, 张智杰, 雷波. 一种针对红外目标检测的 Dense-Yolov5 算法研究 [J]. 光学与光电技术, 2021, 19(1): 69-75.
Shu L, Zhang Z J, Lei B. Research on Dense-Yolov5 algorithm for infrared target detection [J]. Optics & Optoelectronic Technology, 2021, 19(1): 69-75.
- [9] Krois J, Ekert T, Meinhold L, et al. Deep learning for the radiographic detection of periodontal bone loss [J]. Scientific Reports, 2019, 9: 8495.
- [10] Girshick R, Donahue J, Darrell T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [C]//2014 IEEE Conference on Computer Vision and Pattern Recognition, June 23-28, 2014, Columbus, OH, USA. New York: IEEE Press, 2014: 580-587.
- [11] Liu W, Anguelov D, Erhan D, et al. SSD: single shot MultiBox detector [EB/OL]. (2015-12-08) [2021-02-05]. <https://arxiv.org/abs/1512.02325>.
- [12] Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation [M]//Navab N, Hornegger J, Wells W M, et al. Medical image computing and computer-assisted intervention-MICCAI 2015. Lecture notes in computer science. Cham: Springer, 2015, 9351: 234-241.
- [13] Chen H, Zhang K L, Lü P J, et al. A deep learning approach to automatic teeth detection and numbering based on object detection in dental periapical films [J]. Scientific Reports, 2019, 9: 3840.
- [14] Ren S Q, He K M, Girshick R, et al. Faster R-CNN: towards real-time object detection with region proposal networks [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.
- [15] Zhang X, Liang Y, Li W, et al. Development and evaluation of deep learning for screening dental caries from oral photographs [J]. Oral Diseases, 2022, 28(1): 173-181.
- [16] Khan H A, Haider M A, Ansari H A, et al. Automated feature detection in dental periapical radiographs by using deep learning [J]. Oral Surgery, Oral Medicine, Oral Pathology and Oral Radiology, 2021, 131(6): 711-720.
- [17] Redmon J, Divvala S, Girshick R, et al. You only look once: unified, real-time object detection [C]//2016 IEEE Conference on Computer Vision and Pattern Recognition, June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 779-788.
- [18] 侯高峰, 房丰洲. 基于深度学习的糖尿病眼底病变检测的研究 [J/OL]. 激光与光电子学进展:1-12 [2021-07-04]. <http://kns.cnki.net/kcms/detail/31.1690.TN.20211215.1119.008.html>.
Hou G F, Fang F Z. Study on detection of diabetes fundus disease based on deep learning [J/OL]. Laser & Optoelectronics Progress:1-12 [2021-07-04]. <http://kns.cnki.net/kcms/detail/31.1690.TN.20211215.1119.008.html>.
- [19] Mohiyuddin A, Basharat A, Ghani U, et al. Breast tumor detection and classification in mammogram images using modified YOLOv5 network [J]. Computational and Mathematical Methods in Medicine, 2022, 2022: 1359019.
- [20] Yao J, Qi J M, Zhang J, et al. A real-time detection algorithm for kiwifruit defects based on YOLOv5 [J]. Electronics, 2021, 10(14): 1711.
- [21] An Q Z, Pan Z X, Liu L, et al. DRBox-v2: an improved detector with rotatable boxes for target detection in SAR images [J]. IEEE Transactions on Geoscience and Remote Sensing,

2019, 57(11): 8333-8349.

- [22] 高梦婷, 孙晗, 唐云祁, 等. 基于改进 YOLOv5 的指纹二级特征检测方法[J/OL]. 激光与光电子学进展: 1-19[2022-04-04]. <http://kns.cnki.net/kcms/detail/31.1690.tn.20220317.1011.002.html>.

Gao M T, Sun H, Tang Y Q, et al. Fingerprint second-order

minutiae detection method based on improved YOLOv5[J/OL]. 激光与光电子学进展: 1-19[2022-04-04]. <http://kns.cnki.net/kcms/detail/31.1690.tn.20220317.1011.002.html>.

- [23] Wan J J, Chen B L, Yu Y T. Polyp detection from colorectum images by using attentive YOLOv5[J]. *Diagnostics*, 2021, 11(12): 2264.

Automatic Detection of Dental Lesions Based on Deep Learning

Liu Feng¹, Han Min^{1*}, Wang Jun¹, Liu Chao^{2**}

¹ School of Information Science and Engineering, Shandong University, Qingdao 266237, Shandong, China;

² Department of Oral and Maxillofacial Surgery, Qilu Hospital of Shandong University, Jinan 250012, Shandong, China

Abstract

Objective Since the outbreak of COVID-19, many hospitals have become overloaded with patients seeking examination, resulting in an imbalance between medical staff and patients. These high concentrations of people in hospital settings not only aggravate the risk of cross-infection among patients, but also stall the public medical system. Consequently, mild and chronic conditions cannot be treated effectively, and eventually develop into serious diseases. Therefore, the use of deep learning to accurately and efficiently analyze X-ray images for diagnostic purposes is crucial in alleviating the pressure on medical institutions during epidemics. The method developed in this study accurately detects dental X-ray lesions, thus enabling patients to self-diagnose dental conditions.

Methods The method proposed in this study employs the YOLOV5 algorithm to detect lesion areas on digital X-ray images and optimize the network model's parameters. When hospitals and medical professionals collect and label training data, they use image normalization to enhance the images. Consequently, in combination with the network environment, parameters were adjusted into four modules in the YOLOV5 algorithm. In the Input module, Mosaic data enhancement and adaptive anchor box algorithms are used to generate the initial box. The focus component was added to the Backbone module, and a CSP structure was implemented to determine the image features. When the obtained image features are input into the Backbone module, the FPN and PAN structures are used to realize feature fusion. Subsequently, GIOU_Loss function is applied to the Head module, and NMS non-maximum suppression is used to generate a regression of results.

Results and Discussions The proposed YOLOV5-based neural network yields satisfactory training and testing results. The training algorithm produced a recall rate of 95%, accuracy rate of 95%, and F1 score of 96%. All evaluation criteria are higher than those of the target detection algorithms of SSD and Faster-RCNN (Table 1). The network converges to smoothness after loss is reduced in the training process (Fig. 6), which proves that the network successfully learns the necessary features. Thus, the difference between predicted and real values is very small, which indicates good model performance. The mAP value of network training is 0.985 (Fig. 7), which proves that the network training meets the research requirements. Finally, an observation of the visualized thermodynamic diagram reveals that the network's region of interest matches the target detection region (Fig. 8).

Conclusions This study proposes the use of the YOLOV5 algorithm for detecting lesions in dental X-ray images, training and testing on the dataset, modifying the network's nominal batch size, selecting an appropriate optimizer, adjusting the weight parameters, and modifying the learning rate attenuation strategy. The model's training results were compared with those of algorithms used in previous studies. Finally, the effect of feature extraction was analyzed after the thermodynamic diagram was visualized. The experimental results show that the algorithm model detects lesion areas with an accuracy rate of more than 95%, making it an effective autonomous diagnostic tool for patients.

Key words image processing; deep learning; dental lesions; target detection