

基于改进YOLOv7网络模型的肺结节检测算法

刘涌涛, 王宝珠, 郭志涛

河北工业大学电子信息工程学院, 天津 300401

【摘要】针对当前肺结节检测领域中目标检测算法存在的小目标检测精度不高和模型预测框定位不准的问题,提出一种基于YOLOv7的肺结节检测方法,旨在更有效地获取小肺结节并实现目标检测框的持续收敛。在YOLOv7网络模型框架下,在以下3个方面进行改进:首先对头部网络嵌入有效SimAM通道注意力机制获取跨通道信息和目标空域信息,以突出目标特征,使模型能够更加精确地识别感兴趣区域。其次采用SIOU边界损失函数,在原损失函数上增加角度成本,重新定义距离成本和形状成本,以提高收敛速度,降低损失值。最后利用SIOU-NMS替换非极大抑制算法,缓解因目标遮挡而导致错误抑制的现象。实验结果表明,在自制肺结节数据集上,改进网络模型与原YOLOv7模型相比,准确率和召回率分别提升2.9%和3.1%,置信度为0.5时平均精度均值(mAP@0.5)提高3.7%,该模型能有效辅助诊断肺结节。

【关键词】肺结节;YOLOv7;注意力机制;SIOU;SIOU-NMS

【中图分类号】R318

【文献标志码】A

【文章编号】1005-202X(2023)12-1509-09

Lung nodule detection algorithm using improved YOLOv7 network model

LIU Yongtao, WANG Baozhu, GUO Zhitao

School of Electronic and Information Engineering, Hebei University of Technology, Tianjin 300401, China

Abstract: To address the issues in the current lung nodule detection for tuberculosis where the existing object detection algorithms have limited precision for small nodules and often predict bounding box locations inaccurately, a lung nodule detection method based on YOLOv7 is presented for obtaining small lung nodules more effectively and realizing the continuous convergence of target detection box. Based on the framework of YOLOv7 network model, the improvements are made in the following 3 aspects. (1) The cross-channel information and target airspace information are obtained with the effective SimAM channel attention mechanism embed in the Head network, so as to highlight the target features and enable the model to identify the regions of interest more accurately. (2) SIOU boundary loss function is used to increase the angle cost on the original loss function, and redefine the distance cost and shape cost to improve the convergence rate and reduce the loss value. (3) SIOU-NMS is used to replace the non-maximum suppression algorithm for reducing the error suppression due to target occlusion. The results of experiments on a custom lung nodule dataset show that compared with the original YOLOv7, the proposed method improves accuracy and recall rate by 2.9% and 3.1%, and the mean average precision at a confidence threshold of 0.5 is increased by 3.7%. The model can effectively assist in the diagnosis of lung nodules.

Keywords: lung nodule; YOLOv7; attention mechanism; SIOU; SIOU-NMS

前言

肺结节的形成可以有多种原因,其中一种是结核病。结核病是一种由结核分枝杆菌引起的传染

病,它通常会侵犯肺部,但也可以影响其他器官。早期辅助诊断可以有效帮助医师发现肺结节,从而降低感染率和死亡率,具有重要的临床意义。影像学检查是诊断肺结节的主要方法,诊断主要依靠胸部X线检查和计算机断层扫描(CT)^[1]。由于常规X线的检查局限性,CT图像检查已经成为目前不可替代的结核病诊断方法。因此,基于计算机辅助诊断的CT成像研究是医学成像领域的一个重要研究方向,具有工作效率高、准确性稳定的优点。对提高诊断的准确性和降低漏诊率有很大帮助。

【收稿日期】2023-07-25

【基金项目】河北省高等学校科学技术研究项目(ZD2022115)

【作者简介】刘涌涛,硕士,研究方向:智能信息处理、计算机视觉、医学图像处理,E-mail: 1046373227@qq.com

【通信作者】王宝珠,教授,研究方向:信息检测、图像处理、多媒体通信,E-mail: wbz_china@126.com

研究人员在过去采用传统的特征提取方法,随着机器学习和深度学习方法的发展,越来越多的研究人员将这些方法应用于肺结节识别。岳晴等^[2]通过改进 Cascade-RCNN 和 3D CNN,提高肺结节检测的有效性与可行性^[3]。李阳等^[4]提出一种结合残差结构的 ResAdd-UNet,并改进残差结构的 U-Net^[5],能够较精确地检测病灶位置。刘昊然^[6]通过滴水分割算法对肺结节图像进行预处理,并通过 YOLO^[7]网络对肺结节进行检测,但检测平均准确率有待提高。宋尚玲等^[8]采用 Faster-RCNN 网络检测大样本肺结节和 Adaboost 网络检测小样本肺结节,提高肺结节检测平均精度,但其参数数量和计算量过大,不利于实际临床诊断^[9-10]。蔡鸿宇^[11]通过注意力机制和 Ghost Net^[12]改进 YOLOv5 网络对病灶的自动检测,有效降低模型计算量和空间复杂度,但缺乏对图像深层次信息提取,不能有效检测小目标物体。

单次前向传播目标检测算法 YOLO 系列检测速度快,更适用于临床医学图像诊断,其中 YOLOv7^[13]拥有高效的网络结构和更准确的检测方法,因此本文提出一种基于 YOLOv7 的肺结节目标检测算法。经实验证明,改进后算法能提高网络的检测精度,同时解决高精度模型推理框架复杂而带来的无法有效识别小目标和预测框定位不准问题。

1 YOLOv7 目标检测算法

YOLO 是基于深度学习的实时目标检测算法,该算法的核心思想是将目标检测任务作为回归问题,通过网络模型同时预测目标的类别和位置。相比于其他算法,YOLO 检测算法简化算法流程,避免传统目标检测算法中复杂的管道流程,减少网络复杂度,从而提高效率,更适用于需要实时处理的场景。

YOLOv7 是由 Han 等^[12]于 2020 年提出的基于单阶段的目标检测算法,其算法框架由输入端、主干网络、头部网络 3 部分组成。YOLOv7 网络模型输入端通过 Mosaic 方法,对 4 张图片进行随机缩减、随机裁剪和随机排布,对输入数据进行数据增强操作^[14]。主干网络采用若干 CBS (Cross Stage Binary)、E-ELAN (Efficient Layer Aggregation Networks) 和 MP (Max Pooling) 组成,作为特征提取模块,其 3 个输出结果传递给头部网络。头部网络采用“FPN (Feature Pyramid Network)+PAN (Path Aggregation Network)”双塔结构聚合当前的图像特征^[15-16],FPN 建立自上向下路径提取特征,PAN 基于 Mask R-CNN 网络采用自下往上路径提取特征,将下层显著特征信息传递到上层,头部网络实现预测目标的种类和位置。最后预测网络采用 CIOU 损失函数将目标检测任务转换

为回归任务,并利用非极大抑制 (Non-Maximum Supression, NMS) 对目标分类预测结果输出^[17]。

尽管 YOLOv7 有着训练速度快和自适应锚框等优点,但是其对小目标检测、预测框收敛速度和真实框与预测框重叠等问题效果较差。本文将注意力机制融入网络,替换损失函数和 NMS 算法,以满足对肺结节目标的更精确的检测。改进后的 YOLOv7-SSN 模型结构如图 1 所示。

2 基于 YOLOv7 的改进肺结节检测算法

2.1 头部 (Head) 网络的改进

Head 网络连接主干网络输出层,起到加强网络特征融合能力和提高网络检测精度的作用。Head 网络采用 FPN+PAN 结构,FPN 是高维度向低维度传递语义信息以增强多个尺度的语义表达,PAN 是低维度向高维度传递位置信息以增强多个尺度的位置表达,从而实现图像在不同尺寸下对同一物体的识别。肺结节只占特征图的较小部分,其余均为背景信息,经过层次较深的多次网络卷积后,背景信息迭代产生大量的冗余,淹没有效信息,从而导致小目标检测精度下降。本文为了精确定位肺结节小目标,通过一种有效通道注意力机制 SimAM 获取跨通道信息和目标空域信息^[18],使模型能够更加精确地识别感兴趣区域并对小目标进行识别,其具体结构如图 2 所示。

图 2 分别展示在通道维度上对特征进行加权和空间维度上对特征进行加权,而本文采用的 3D 注意力机制 SimAM,基于神经科学研究提出能量函数,用于计算注意力权值,解决前两种注意力机制对特征位置信息和通道信息不兼容的缺点。其能量函数表达式为:

$$e_i(w_i, b_i, y, x_i) = (y_i - \hat{t})^2 + \frac{1}{M-1} \sum_{i=1}^{M-1} (y_0 - \hat{x}_i)^2 \quad (1)$$

其中, t 代表目标神经元, x 代表其他神经元, y_i 代表目标神经元的值, y_0 代表其他神经元的值, \hat{t} 代表目标神经元的线性变换, \hat{x}_i 代表其他神经元的线性变换, M 代表通道数,在不失泛化性的前提下用二进制将 y_i 赋值为 1, y_0 赋值为 0 优化式子,优化后式子为:

$$e_i(w_i, b_i, y, x_i) = \frac{1}{M-1} \sum_{i=1}^{M-1} (-1 - (w_i x_i + b_i))^2 + (1 - (w_i x_i + b_i))^2 + \lambda w_i^2 \quad (2)$$

表达式中 w_i, b_i 的解析式为:

$$w_i = -\frac{2(t - \mu_i)}{(t - \mu_i)^2 + 2\sigma_i^2 + 2\lambda} \quad (3)$$

$$b_i = -\frac{1}{2}(t + \mu_i)w_i \quad (4)$$

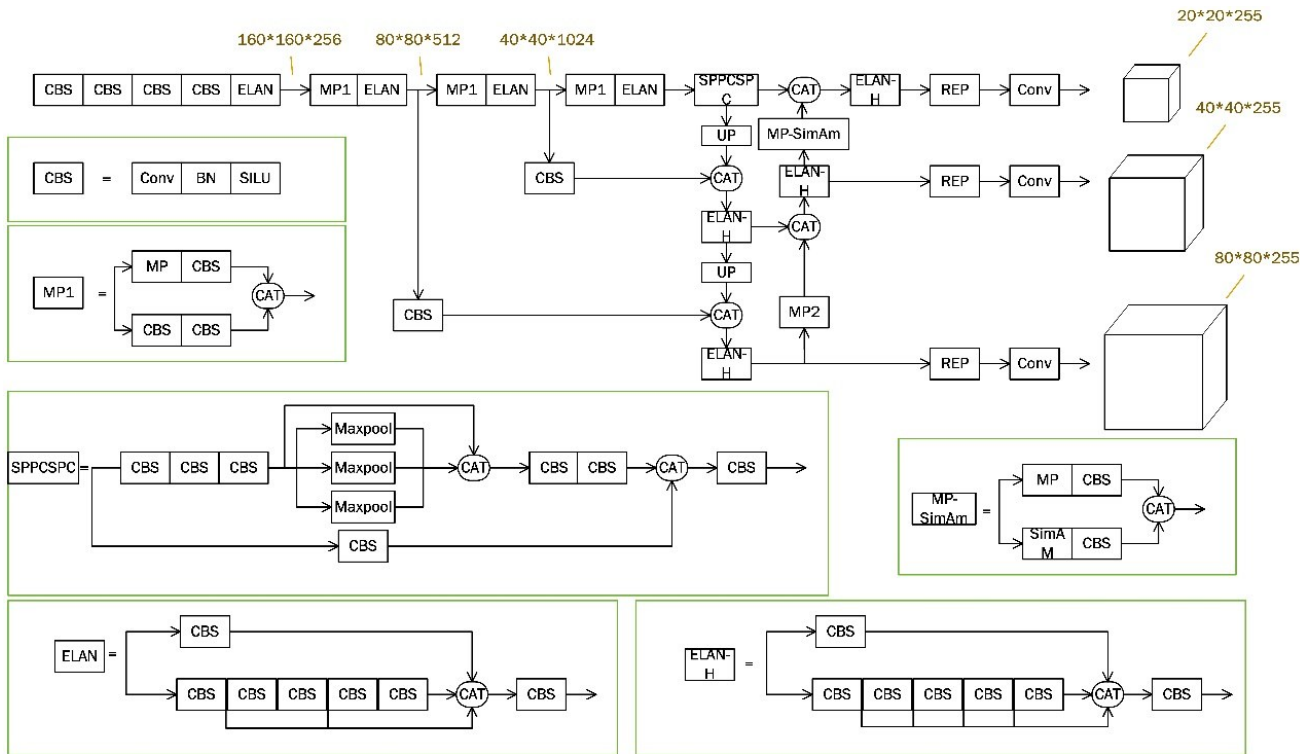


图1 改进后YOLOv7-SSN结构
Figure 1 Improved YOLOv7-SSN structure

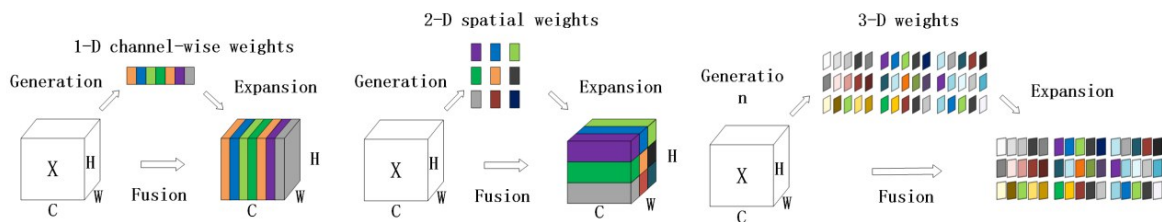


图2 SimAM结构
Figure 2 SimAM structure

其中, μ_i 和 σ_i^2 代表除目标神经元以外所有神经元的平均值和方差, 根据以上公式可以推导出低能量公式为:

$$e_i^* = \frac{4(\hat{\sigma}^2 + \lambda)}{(t - \hat{\mu})^2 + 2\hat{\sigma}^2 + 2\lambda} \quad (5)$$

上述式子表明能量越低, 目标神经元和周围神经元的区别越大, 基于空间抑制现象, 低能量 e_i^* 更能代表目标神经元, 通过推导出能量函数, 进一步证实神经元的重要性, 最终输出为:

$$\tilde{x} = \sigma\left(\frac{1}{E}\right) \otimes x \quad (6)$$

其中, x 表示输入特征张量, \otimes 代表逐个元素的乘法, E 代表特征图上所有 e_i^* 的能量值。

本文为了精确定位小目标, 提升小目标的检测精度, 在通道 MP 模块后添加有效 SimAM 注意力机

制模块, 并与输出网络相连接, 不仅减少了由于背景信息迭代而引起有效信息丢失, 也缓解了由于特征通道减少而导致的高层信息丢失的现象。

2.2 损失函数改进

损失函数作用于计算预测与实际数据的差距, 是衡量模型预测和实际效果的重要依据。在 YOLOv7 模型的损失函数包括置信度损失 L_{object} 、边界损失函数 L_{box} 和分类损失函数 L_{class} 3 项损失^[19], 因此 YOLOv7 模型的损失函数为 3 个损失加权求和所得。其公式可以表达为:

$$L = a \times L_{\text{object}} + b \times L_{\text{box}} + c \times L_{\text{class}} \quad (7)$$

其中, 默认值为 $a = 0.7$ 、 $b = 0.05$ 和 $c = 0.3$, 原模型边框损失函数为 CIOU, 即 $L_{\text{box}} = L_{\text{CIOU}}$, 当边界框预测越准确时, L_{box} 和 L_{CIOU} 将越小, 模型的性能也越好。CIOU 函数^[20]数学表达式为:

$$L_{\text{CIOU}} = L_{\text{IOU}} + L_{\text{dis}} + L_{\text{asp}} = 1 - I_U + \frac{\rho^2(b, b_{gt})}{c^2} + \alpha v \quad (8)$$

其中,预测锚框和真实锚框的交并比损失 $L_{\text{IOU}} = 1 - I_U$,预测锚框和真实锚框的中心距离损失 $L_{\text{dis}} = \rho^2(b, b_{gt})/c^2$,纵横比损失 $L_{\text{asp}} = \alpha v$ 。

$$I_U = \left| \frac{B \cap B_{gt}}{B \cup B_{gt}} \right| \quad (9)$$

$$v = \frac{4}{\pi^2} \left(\arctan \frac{w_{gt}}{h_{gt}} + \arctan \frac{w}{h} \right)^2 \quad (10)$$

$$\alpha = \frac{v}{(1 - I_U) + v} \quad (11)$$

其中, v 代表预测框和真实框的宽高比相似度, α 代表 v 的权重函数。此方法的重要变量是 v , 但纵横比是一种相对指标, 并未考虑到复杂样本的平衡问题, 且没有考虑到所需真实框与预测框之间角度不匹配问题。这种不足导致预测锚框收敛速度慢且不能持续收敛, 因为预测框可能在训练过程中发生旋转而不能和真实锚框进行匹配, 并最终产生更差的模型。针对这一问题, 提出 SIOU 边界损失函数^[21], 该函数在原基础损失函数上提出角度成本, 并重新定义距离成本和形状成本。其角度成本计算公式可以表达为:

$$A = 1 - 2 \sin^2(\arcsin(x) - \pi/4) \quad (12)$$

其中,

$$x = c_h / \sigma \quad (13)$$

$$\sigma = \sqrt{(b_{c_x}^{gt} - b_{c_x})^2 + (b_{c_y}^{gt} - b_{c_y})^2} \quad (14)$$

$$c_h = \max(b_{c_y}^{gt}, b_{c_y}) - \min(b_{c_y}^{gt}, b_{c_y}) \quad (15)$$

通过角度成本, 对距离成本重新定义, 计算表达式为:

$$\Delta = \sum_{t=x,y} (1 - e^{-\gamma \rho_t}) \quad (16)$$

其中,

$$\rho_x = (b_{c_x}^{gt} - b_{c_x}) / c_w \quad (17)$$

$$\rho_y = (b_{c_y}^{gt} - b_{c_y}) / c_h \quad (18)$$

$$\gamma = 2 - A \quad (19)$$

重新定义后的形状成本计算公式为:

$$\Omega = \sum_{t=w,h} (1 - e^{-w_t})^\theta \quad (20)$$

其中,

$$w_w = \frac{|w - w^{gt}|}{\max(w, w^{gt})} \quad (21)$$

$$w_h = \frac{|h - h^{gt}|}{\max(h, h^{gt})} \quad (22)$$

最终 SIOU 边界损失函数表达式为:

$$L_{\text{box}} = L_{\text{SIOU}} = 1 - \text{IOU} + (\Delta + \Omega)/2 \quad (23)$$

其中,

$$\text{IOU} = \left(\frac{|B \cap B^{GT}|}{|B \cup B^{GT}|} \right) \quad (24)$$

将 YOLOv7 中 CIOU 边界损失函数替换为 SIOU 损失函数, 即 $L_{\text{box}} = L_{\text{SIOU}}$, 通过计算角度成本, 重新定义距离成本和形状成本, 模型在训练时损失函数的收敛速度更快, 更有助于模型在学习过程中降低损失值, 提高定位精度, 从而提高模型的识别度。

2.3 NMS 算法改进

NMS 算法是对目标检测后进行目标框操作^[22], 目的是对目标框进行筛选, 抑制其他目标候选框, 找到最佳的物体检测位置。通过对比所有检测框置信度, 获取最大置信度目标检测框, 计算其余检测框与当前最大置信度检测框交并比 (Intersection Over Union, IOU) 来对重叠的预测候选框进行筛选, 当 IOU 值高于置信度时, 一组检测框被标识为同一对象。其公式可以表达为:

$$\text{box} = B_{\arg \max C_i} \quad (25)$$

其中, B 代表检测框; C 代表第 i 个检测框的置信度; box 代表具有最高置信度的检测框, 作为输出结果。但此方法将同一目标附近交并比进行置信度对比, 得分低的检测框将被强制置零, 在此情况下, 出现两个或者多个目标相邻时, 则将导致无法实现有效检测, 导致物体检测的平均检测率降低。同时置信度的选择过低导致误删检测框, 过高则会增大误检^[23]。针对医学图像领域中肺结节检测, 目标肺结节存在密集和相互邻近等复杂情况, 定位不准确的检测框会严重影响到目标肺结节与其余肺结节的相对位置关系判断, 针对以上问题, 本文以 SIOU 来考虑边框之间的相似度 (SIOU-NMS), 对检测框进行筛选, 从而提高算法整体性能。以 SIOU 来考虑边框之间的相似度评价表达式为:

$$\begin{cases} S_i = 0, f_{\text{SIOU}}(M, B_i) \geq \text{thresh} \\ S_i = S_i, f_{\text{SIOU}}(M, B_i) < \text{thresh} \end{cases} \quad (26)$$

其中, thresh 表示 IOU 置信度; $f_{\text{SIOU}}(M, B_i)$ 表示以 SIOU 计算边框相似度结果; 当 $S_i = S_i$ 则视为不同的目标框, 当 $S_i = 0$ 则将该检测框视为冗余框。

以 SIOU 来考虑边框之间的相似度, 选值更为精准, 更有助于缓解肺气管、肺泡等目标遮挡而被错误抑制的现象。

3 实验设计与结果分析

3.1 数据集制作

本研究采用北京胸科医院影像科所提供的真实数据集进行研究, 该数据集包含 300 例肺结核导致结

节病人的CT图像。同时在数据集勾画标准中,通过将医学影像数据导入相同勾画软件、定位勾画区域、大致勾画、精细修正、医师复核最终确定。为确保数据集的一致性和有效性,使用Labellmg对肺结节进行标注和校准。在校准规则中,所有图像都经过色彩和对比度的校准,以确保图像质量的一致性,同时由3位专业人员对标注结果进行独立审核。在标注完成后,肺结节的具体位置等信息被保存在XML文件中,这些XML文件作为后续实验的标签信息。同时对每张CT图像进行预处理,将原本的DICOM格式转换为方便使用的PNG格式,分辨率调整为418×418,并且编码位数为24位。数据集中共包含1348张直径范围为5~30 mm的肺结节CT切片,采用随机抽样的方式选取其中80%用作训练集,10%用作验证集,10%用作测试集。

3.2 实验配置及其初始参数

本实验的运行系统为Windows 11,处理器为Intel(R) Core(TM) i5-12600KF, GPU 为 12 GB 显存 GeForce RTX 3060。本文选用PyTorch开源框架进行训练、测试,汇编语言采用Python3.8。为适应模型训练,初始网络超参数设置如表1所示。并在网络模型中提供超参数优化的方法(Hyperparameter Evolution)进行学习 and 更新网络参数,进行300次的迭代训练。

表1 模型初始超参数设置
Table 1 Model initial hyperparameter setting

训练参数名称	参数值
初始学习率	0.01
余弦退火超参数	0.1
权重衰减系数	0.000 5
预热学习率动量	0.8
批处理大小	8
训练批次	300
图像输入尺寸	418×418

3.3 模型评价指标

本文肺结节检测模型主要评价指标有准确率、召回率和置信度为0.5时平均精度均值(mAP:0.5)^[24]。其中准确率表达式为:

Accuracy = $\frac{TP}{TP + FP}$

(27)

其中,TP为正样本且被判定为正样本,FP为假样本且被判定为正样本,准确率代表正样本预测准确程度。召回率表达式为:

Recall = $\frac{TP}{FN + TP}$

(28)

其中,FN为正样本且判定为假样本,召回率是指正确识别出的实际正样本占有所有实际正样本的比例。设置置信度为0.5,构建准确率和召回率生成曲线,平均精度均值等于准确率和召回率曲线下面积的平均值,mAP:0.5表达式为:

mAP:0.5 = $\frac{\sum^n \int E_{pr} \cdot E_{re} d(E_{re})}{n}$

(29)

3.4 主流模型对比分析

为验证本文所提出的YOLOv7-SSN网络模型在肺结节识别的有效性,将该模型与双检测阶段模型Faster RCNN和近几年YOLO系列模型进行对比实验。通常高平均准确率、低参数量、低运算量的肺结节检测模型更适合临床肺结节诊断。本文实验以模型复杂度(参数量和计算量),以及mAP:0.5作为评价指标。各模型在训练集上的检测结果如表2所示。

表2 各主流检测模型对比
Table 2 Comparison with various mainstream detection models

模型	mAP:0.5	参数量/M	计算量/GFLOs
Faster RCNN	0.887	136.65	370.4
YOLOv5l	0.860	46.50	109.1
YOLOv6l	0.871	59.60	150.7
YOLOv7	0.880	37.19	105.1
YOLOv7-SSN	0.917	36.18	103.9

通过表2可知,Faster-RCNN有较好的平均准确率,但参数量和计算量较高,模型内存占用率高和运算量大。YOLOv5l和YOLOv6l在结节检测中表现较好,但是由于在测试集中有较多的小型结节,容易与肺区内正常气管和肺泡混淆,导致定位不准确。具体检测效果如图3~图8所示。从图3~图8可以看出YOLOv7效果较好,主要原因是YOLOv7网络拥有更强大的网络结构。本文提出的YOLOv7-SSN网络相对于YOLOv7网络表现进一步提升,其平均准确率更高。

3.5 各模块改进对比分析

为验证本文改进点的有效性,本文采用递增的方式逐步修改每个改进模块,观察其对网络性能影响。实验结果如表3所示。YOLOv7-SimAM只对网络结构中Head进行改进,将SimAM模块添加下采样通道中的MP模块,网络特征提取能力更强,相对于YOLOv7网络准确率提高1.0%,召回率提高0.5%,

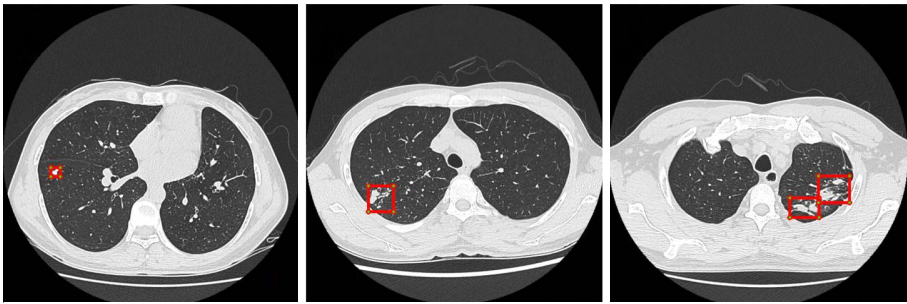


图 3 真实目标所在位置
Figure 3 Location of the real target

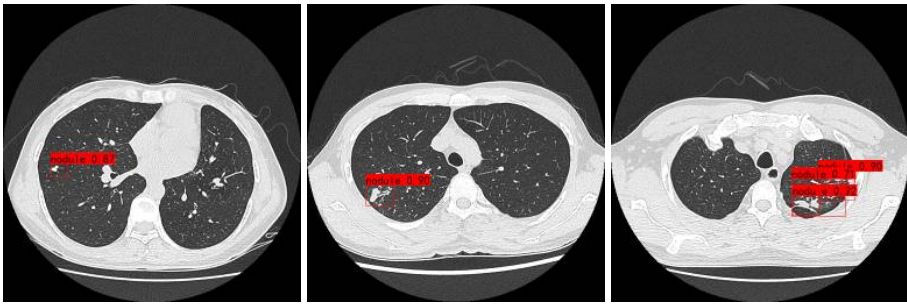


图 4 Faster-RCNN 目标检测结果
Figure 4 Faster-RCNN target detection results

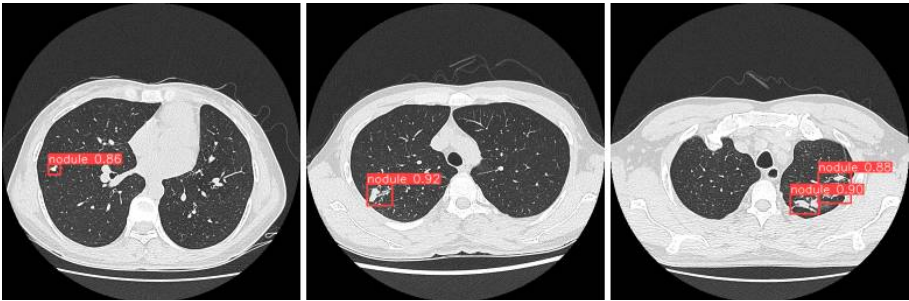


图 5 YOLOv5l 目标检测结果
Figure 5 YOLOv5l target detection results

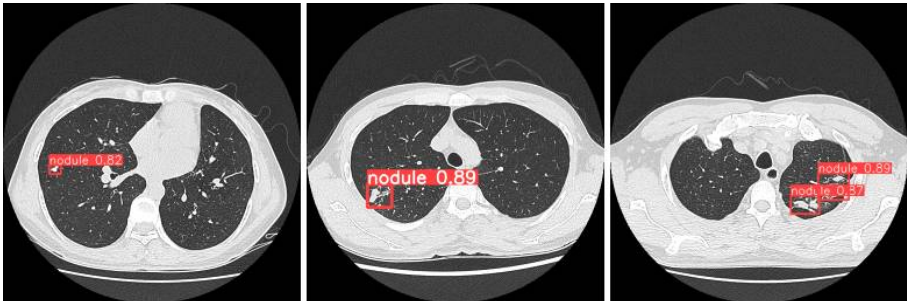


图 6 YOLOv6l 目标检测结果
Figure 6 YOLOv6l target detection results

mAP:0.5 提高 0.5%。YOLOv7-SimAM-SIOU 对网络 Head 部分和 Detect 部分都进行改进,将 Detect 部分损

失函数替换为 SIOU 处理,使检测框持续收敛,相对于 YOLOv7-SimAM 网络准确率提高 1.7%,召回率提

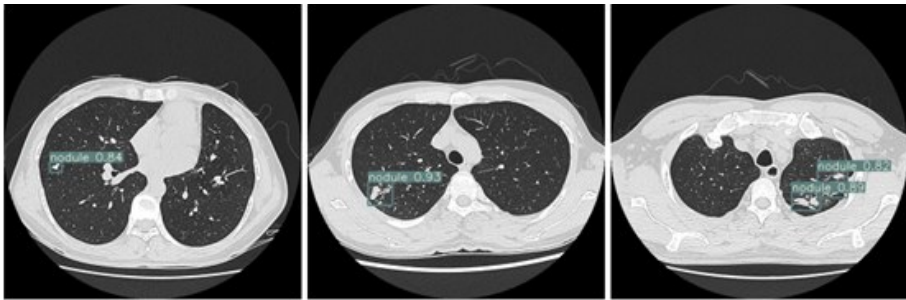


图7 YOLOv7目标检测结果
Figure 7 YOLOv7 target detection results

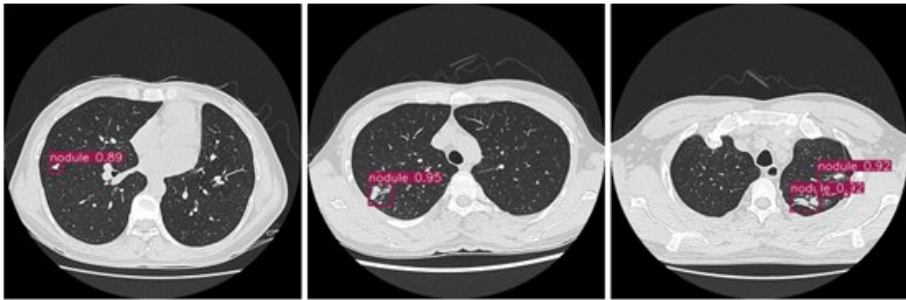


图8 YOLOv7-SSN目标检测结果
Figure 8 YOLOv7-SSN target detection results

高 1.8%, mAP: 0.5 提高 2.5%。YOLOv7-SSN 在 YOLOv7-SimAM-SIOU网络的基础上使用SIOU-NMS处理,使检测框定位更精确,相对于YOLOv7-SimAM-SIOU,准确率提高 0.2%,召回率提高 0.8%,mAP:0.5 提高 0.7%。

表3 改进模块对肺结节检测模型影响
Table 3 Effects of each improved module on the lung nodule detection model

模型	准确率	召回率	mAP:0.5
YOLOv7	0.881	0.862	0.880
YOLOv7-SimAM	0.891	0.867	0.885
YOLOv7-SimAM-SIOU	0.908	0.885	0.910
YOLOv7-SSN	0.910	0.893	0.917

从表3数据可以看出,改进YOLOv7-SSN性能与原网络差异,对改进前后训练结果进行对比。随着对网络每个模块进行改进,准确率、召回率、mAP:0.5值都得到逐步提升,其中损失函数改进对网络性能提升最大。改进网络模型YOLOv7-SSN准确率和召回率相对比YOLOv7分别提升2.9%和3.1%,mAP:0.5达91.7%。进一步对改进后的网络模型YOLOv7-SSN与原始模型YOLOv7进行比较,在300次迭代过程中

以准确率、召回率以及mAP:0.5为指标,探究改进模型的性能提升。

从准确率(图9a)和召回率(图9b)对比曲线可以观察到,在训练批次增加的过程中,模型在大约迭代200次后趋于稳定。在网络达到稳定状态之后,改进模型的整体准确率和召回率都优于YOLOv7网络。此外,在mAP:0.5对比曲线上(图9c)也可以看出,在网络趋于稳定之后,本文模型的mAP:0.5也超过了YOLOv7网络。改进后模型的准确率-召回率对比曲线(图9d)明显优于原YOLOv7网络曲线。实际检测图如图10~图12所示。

图10~图12中,针对实际情况中小目标、目标周围有肺气泡干扰以及目标检测框相邻肺结节3种类型的图像进行评估。本文检测模型可以在小目标区域实现更高的检测精度,如图10中改进后模型检测精度提高5%。在检测区域存在肺气泡干扰时(图11),改进后模型检测精度提高2%。同时在邻近的目标区域(图12),相邻目标检测精度分别提高10%和3%。整体效果来看,本文提出模型拥有更高的检测准确率。

4 结 论

本文在网络YOLOv7的基础上进行改进使其更适合肺结节目标检测。在YOLOv7网络模型框架

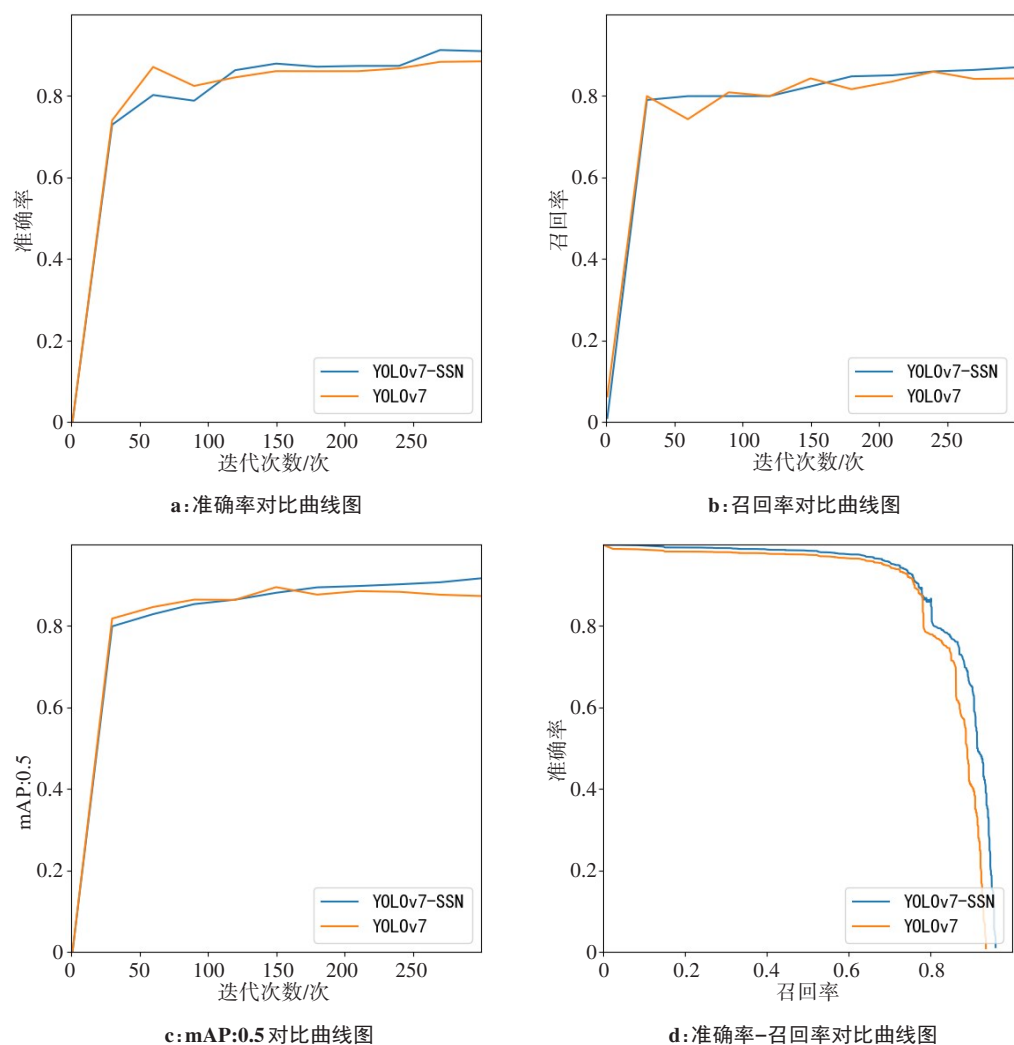


图9 模型改进前后准确率、召回率、mAP:0.5 对比曲线和准确率-召回率对比曲线图

Figure 9 Comparison of accuracy, recall rate, mean average precision at a confidence threshold of 0.5, and accuracy-recall curves before and after model modification

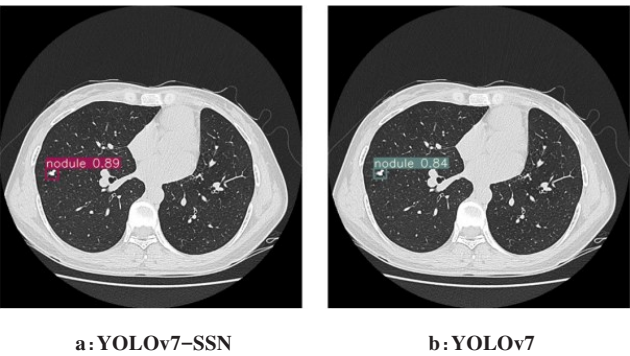


图10 小目标肺结节检测示例

Figure 10 Examples of small lung nodule detection

下,对头部网络嵌入有效注意力机制获取跨通道信息和目标空域信息,以突出目标特征,使模型能够更加精确地识别感兴趣区域。并采用SIOU边界损失函数(原损失函数上增加角度成本,重新定义距离成本和形状成本)提高检测框收敛速度,降低损失值。

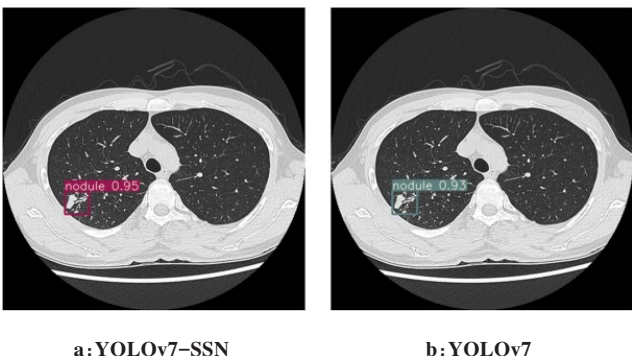


图11 目标周围有肺气泡干扰时肺结节检测示例

Figure 11 Examples of lung nodule detection interfered with pulmonary bubbles around the target

同时利用SIOU-NMS替换NMS算法,缓解因目标遮挡而导致错误抑制的现象。在自制数据集上进行实验表明,改进后网络准确率、召回率和mAP:0.5均高于原YOLOv7模型,具有优秀的肺结节病灶检测潜

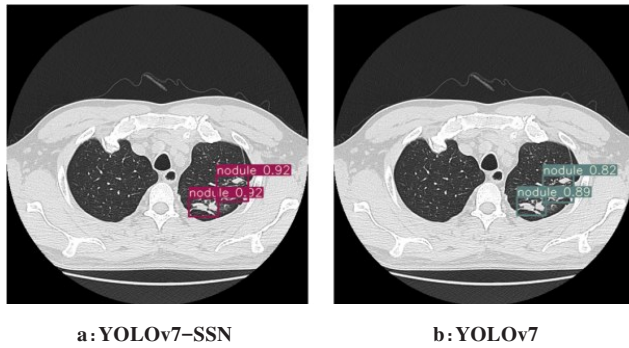


图12 目标检测框相邻肺结节检测示例

Figure 12 Example of adjacent lung nodule detection in the target detection box

力。由于数据获取的限制,目前研究中尚未包含更多类型的肺结节影像表现,如恶性肺结节和良性肺结节等。为了扩充数据集,后续计划引入其他类型的肺结节病变,这在未来对各类肺部疾病的辅助诊断具有一定的参考价值。

【参考文献】

- [1] 卫旭瑛,白峥嵘,孔令福,等. MSCT薄壁囊腔型肺癌、薄壁空洞性肺结核影像学特征及鉴别诊断[J]. 中国CT和MRI杂志, 2022, 20(5): 68-70.
Wei XY, Bai ZR, Kong LF, et al. MSCT imaging features and differential diagnosis of lung cancer with thin-walled cavity and thin-walled cavitary tuberculosis[J]. Chinese Journal of CT and MRI, 2022, 20(5): 68-70.
- [2] 岳晴,尹健宇,王生生. 基于改进CNN的低剂量CT图像肺结节自动检测[J]. 计算机科学, 2022, 49(S1): 54-59.
Yue Q, Yin JY, Wang SS. Automatic detection of pulmonary nodules in low-dose CT images based on improved CNN[J]. Computer Science, 2022, 49(S1): 54-59.
- [3] Cai Z, Vasconcelos N. Cascade R-CNN: delving into high quality object detection[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2018: 6154-6162.
- [4] 李阳,宋长明,王浩,等. 基于残差结构的三维U-Net肺结节检测算法[J]. 中原工学院学报, 2022, 33(2): 72-78.
Li Y, Song CM, Wang H, et al. The-dimensional U-Net lung nodule detection algorithm based on residual structure [J]. Journal of Zhongyuan Institute of Technology, 2022, 33(2): 72-78.
- [5] Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation[C]//Medical Image Computing and Computer-Assisted Intervention. 2015: 234-241.
- [6] 刘昊然. 基于YOLOv4的肺结节辅助检测方法的研究与实现[D]. 重庆: 重庆邮电大学, 2021: 3-5.
Liu HR. Research and implementation of lung nodule auxiliary detection method based on YOLOv4[D]. Chongqing: Chongqing University of Posts and Telecommunications, 2021: 3-5.
- [7] Eedmon 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(CVPR). IEEE, 2016: 779-788.
- [8] 宋尚玲, 杨阳, 李夏, 等. 基于Faster-RCNN的肺结节检测算法[J]. 中国生物医学工程学报, 2020, 39(2): 129-136.
Song SL, Yang Y, Li X, et al. Pulmonary nodule detection algorithm based on Faster-RCNN [J]. Chinese Journal of Biomedical Engineering, 2020, 39(2): 129-136.
- [9] Ren S, He K, Girshick R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. IEEE Trans Pattern Anal Mach Intell, 2017, 39(6): 1137-1149.
- [10] Yoav F, Robert E. A decision-theoretic generalization of on-line learning and an application to boosting[J]. J Comput Syst Sci, 1997, 55: 119-139.
- [11] 蔡鸿宇. 基于YOLOv5的肺结节检测算法研究[D]. 昆明: 昆明理工大学, 2023: 5-7.
Cai HY. Research on lung nodule detection algorithm based on YOLOv5 [D]. Kunming: Kunming University of Science and Technology, 2023: 5-7.
- [12] Han K, Wang Y, Tian Q, et al. GhostNet: more features from cheap operations [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. IEEE, 2020: 1580-1589.
- [13] Wang CY, Bochkovskiy A, Liao HY. YOLOv7: trainable bag-of-freebies sets new state of the art for real-time object detectors[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. IEEE, 2023: 7464-7475.
- [14] Wang CY, Mark L, Wu YH, et al. CSPNet: a new backbone that can enhance learning capability of CNN[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). IEEE, 2020: 1571-1580.
- [15] Lin TY, Dollar P, Girshick R, et al. Feature pyramid networks for object detection[C]//2017 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2017: 2117-2125.
- [16] Liu S, Qi L, Qin H, et al. Path aggregation network for instance segmentation[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2018: 8759-8768.
- [17] Neubeck A, Gool LV. Efficient non-maximum suppression[C]//18th International Conference on Pattern Recognition (ICPR 2006). 2006: 850-855.
- [18] Yang L, Zhang RY, Li L, et al. SimAM: a simple parameter-free attention module for convolutional neural networks[C]//International Conference on Machine Learning. 2021: 11863-11874.
- [19] 黄凯变, 房宵杰, 梅林, 等. MPE-YOLOv5:面向边缘计算的轻量化YOLOv5手势识别算法[J]. 哈尔滨工业大学学报, 2023, 55(5): 1-13.
Huang KW, Fang XJ, Mei L, et al. MPE-YOLOv5: lightweight YOLOv5 gesture recognition algorithm for edge computing [J]. Journal of Harbin Institute of Technology, 2023, 55(5): 1-13.
- [20] Du S, Zhang B, Zhang P, et al. An improved bounding box regression loss function based on CIOU loss for multi-scale object detection[C]//2021 IEEE 2nd International Conference on Pattern Recognition and Machine Learning (PRML). IEEE, 2021: 92-98.
- [21] Gevorgyan Z. SIOU loss: more powerful learning for bounding box regression[J]. arXiv preprint arXiv:2205.12740, 2022.
- [22] Nayereh Z, Mahmood F, Seyed MJ, et al. The improvement in obstacle detection in autonomous vehicles using YOLO non-maximum suppression fuzzy algorithm[J]. J Supercomput, 2021, 77(11): 13421-13446.
- [23] 郝佳杰, 方赛银, 肖涵, 等. 改进YOLOv5的叶片黑斑病检测算法研究[J]. 南方农机, 2022, 53(16): 1-4.
Hao JJ, Fang SY, Xiao S, et al. Research on improved YOLOv5 algorithm for leaf spot detection[J]. Southern Agricultural Machinery, 2022, 53(16): 1-4.
- [24] Christine M, Daniel G, Tim H, et al. Anaphora resolution: analysing the impact on mean average precision and detecting limitations of automated approaches[J]. Int J Inform Retrieval Res, 2018, 8(3): 33-45.

(编辑:薛泽玲)