

基于YOLOv8s的改进结核病病原体检测算法

乐振¹, 孙振², 鞠瑞文², 李庆党³

1. 青岛科技大学信息科学技术学院, 山东 青岛 266061; 2. 青岛科技大学数据科学学院, 山东 青岛 266061; 3. 青岛科技大学中德科技学院, 山东 青岛 266061

【摘要】针对结核病病原体在痰涂片图像背景复杂且目标小, 人工筛查成本较高的问题, 提出了一种基于YOLOv8s的结核病病原体检测方法。首先, 采用基于空间和通道重构卷积改进的结构来限制特征冗余。其次, 引入了坐标注意力来扩大模型的感受野。再者, 使用空间金字塔池化跨阶段局部网络来提取不同尺度上的特征信息。最后, 嵌入基于归一化的注意力机制抑制不太显著的特征。实验结果表明, 在公开数据集上, 改进网络模型与原YOLOv8s模型相比, 精确率和召回率分别提升2.7%和1.5%, 置信度为0.5时的平均精度均值提高了2.3%, 该模型能够有效辅助影像科医师进行诊断。

【关键词】结核病; 深度学习; YOLOv8; 注意力机制

【中图分类号】R318; TP391

【文献标志码】A

【文章编号】1005-202X(2024)07-0910-08

A novel tuberculosis pathogens detection algorithm based on YOLOv8s

YUE Zhen¹, SUN Zhen², JU Ruiwen², LI Qingdang³

1. School of Information Science and Technology, Qingdao University of Science and Technology, Qingdao 266061, China; 2. School of Data Science, Qingdao University of Science and Technology, Qingdao 266061, China; 3. College of Sino-German Science and Technology, Qingdao University of Science and Technology, Qingdao 266061, China

Abstract: To address the challenges of detecting tuberculosis pathogens in sputum smear images, such as complex backgrounds, small targets, and high costs of manual screening, a detection method based on YOLOv8s is presented. The structure is improved through spatial and channel reconstruction convolutions to limit feature redundancy, and a coordinate attention is introduced to expand the receptive field of the model. Furthermore, a spatial pyramid pooling cross-stage partial network is used to extract feature information at different scales, and a normalized attention mechanism is embedded to suppress less significant features. The experimental results on a public dataset show that compared with the original YOLOv8s model, the improved network model enhances precision and recall rates by 2.7% and 1.5%, respectively, and improved mean average precision at a confidence level of 0.5 by 2.3%, demonstrating that the improved model can effectively assist radiologists in diagnosis.

Keywords: tuberculosis; deep learning; YOLOv8; attention mechanism

前言

结核病是一种由结核分枝杆菌感染引起的慢性传染病, 最常见发生于肺部, 也称为肺结核^[1]。自发现结核分枝杆菌以来, 全球约有两亿人死于结核病。

结核病目前仍然是世界重大公共卫生问题之一, 是全球尤其是发展中国家危害最为严重的慢性传染病之一。《2023年全球结核病报告》显示, 2022年中国结核病新发患者数74.8万, 为全球第3大结核病高负担国家, 死于该病的人数约为3万^[2]。

早发现对此病的治疗具有重要意义^[3]。痰液结核分枝杆菌检查是确诊肺结核的主要方法之一, 也是制定医疗方案和评价治疗效果的主要依据^[4]。以光学显微镜检测一个痰液样本涂片至少需要统计300个视野下的结核分枝杆菌数量, 工作量较大, 容易出现漏检和误检的情况。因此使用人工智能技术辅助诊断显得尤为重要。

【收稿日期】2024-03-10

【基金项目】山东省泰山学者项目(tshw201502042)

【作者简介】乐振, 硕士研究生, 研究方向: 医学图像处理, E-mail: yzhellworld@163.com

【通信作者】李庆党, 教授, 研究方向: 人工智能及大数据技术, E-mail: lqd@qust.edu.cn

国内外有大量研究人员逐渐开始利用计算机辅助诊断肺结核^[5]。王旭^[6]使用3种色彩模型对结核分枝杆菌痰液涂片的显微镜图像进行特征提取,再使用朴素贝叶斯分类器对病菌进行分类识别。Alfadhli等^[7]使用加速鲁棒特征的空间金字塔,从大小不同的网格窗口中提取特征并连接在一起,然后用来训练基于支持向量机的分类器,最终实现了73%的受试者工作特征曲线下面积(Area Under Curve, AUC)。以上两种方法采用人工特征提取,不仅需要专业知识,还耗费时间和精力。张璇^[8]提出一种基于改进单阶段多方框检测器(Single Shot Multibox Detector, SSD)网络的结核分枝杆菌小目标检测算法^[9]。通过对结核分枝杆菌目标检测数据集进行训练和测试,平均精度均值(Mean Average Precision, mAP)达到79.8%。相较于SSD网络提高了2.3%,检测速度较快,但其模型平均精确度不够高。Yang等^[10]以肺结核CT图像为数据集,选用双阶段目标检测网络Faster-RCNN(Faster Region-CNN)^[11]为主体,结合了U-Net^[12]和残差块,一定程度上提升了检测精度,但其模型参数量较大,检测速度慢。

本文提出一种基于YOLOv8的结核病病原体检测算法。经实验证明,改进后的算法能提高网络的检测精度,并且检测速度较快。

1 YOLOv8s目标检测算法

YOLO(You Only Look Once)^[13]是目标检测领域的一种单阶段检测算法。由于其计算量小、实时性高的特性,应用非常广泛。YOLOv8由Ultralytics公司在2023年发布,是目前YOLO系列中最新的算法。相比于之前版本的YOLO以及其他主流的检测算法,其速度更快、精度更高、性能更好,在众多任务上取得了SOTA的成绩。YOLOv8融合了许多实时目标检测的优异技术,在保留了之前版本的架构情况下,其主要改进有:(1)通过参考YOLOv7^[14]中ELAN(Efficient Layer Aggregation Networks)设计思想将主干和颈部的C3(CSP bottleneck with 3 convolutions)模块换成了C2f(CSP bottleneck with 2 convolutions)模块。(2)检测头部分换成了解耦头结构,同时使用Anchor-Free替换了Anchor-Based。(3)使用BCE Loss作为分类损失函数,使用DFL Loss、CIOU Loss^[15]作为回归损失。(4)训练策略上,最后10轮关闭Mosaic数据增强操作,该操作可以有效提升精度。

根据计算量大小以及模型深度和宽度的区别,YOLOv8一共分为YOLOv8n、YOLOv8s、YOLOv8m、YOLOv8l、YOLOv8x 5个版本。考虑到参数量大小、精度等要素,本文选择了YOLOv8s作为基线模型。

2 基于YOLOv8s的改进检测算法

通过实验,提出了一种基于YOLOv8s的改进模型。考虑到模型对检测速度的要求,设计了一种C2f_ScConv结构代替主干特征提取网络中原有的C2f组件。C2f_ScConv能够有效地限制特征冗余,增强特征表示能力。结核分枝杆菌在图像中呈现尺度较小、形状多变的特性。由于卷积只能捕捉局部关系,因此将坐标注意力机制(Coordinate Attention, CA)^[16]嵌入网络。通过将位置信息嵌入通道注意力中,使得网络能够关注较大的区域。为了更好地在不同尺度上对图像进行特征融合,避免裁剪、缩放操作导致的图像失真问题,以便更好地检测不同大小的目标,笔者使用空间金字塔池化跨阶段局部网络(Spatial Pyramid Pooling Cross Stage Partial Network, SPPCSPC)替代网络中的快速金字塔池化(Spatial Pyramid Pooling Fast, SPPF)模块。部分结核分枝杆菌在图像中与空白部分较为相似,为了降低不太显著的特征权重,在颈部引入基于归一化的注意力机制(Normalization-based Attention Module, NAM)^[17],这使得网络在计算上更加高效,同时还能达到更高的准确率。

2.1 C2f_ScConv

卷积在各种计算机视觉任务中表现出色,但是卷积提取的特征之间会有空间和通道冗余。YOLOv8原有组件C2f的多分支设计提高了模型特征提取能力,但也增加了模型的参数量,降低了检测速度,同时还带有不少冗余的数据。

空间和通道重建卷积(Spatial and Channel reconstruction Convolution, ScConv)包括两个单元,一个为空间重构单元(Spatial Reconstruction Unit, SRU),另一个为通道重构单元(Channel Reconstruction Unit, CRU)^[18]。SRU通过先分离后重建的方法抑制空间维度上的冗余,增强特征表示能力,其结构如图1所示。分离操作的目的是将信息丰富的特征图与空间内容对应信息较少的特征图分离开来。笔者利用组归一化(Group Normalization, GN)^[19]层中的比例因子来评估不同特征图的信息内容,其表达式如下:

$$X_{\text{out}} = \text{GN}(X) = \gamma \frac{X - \mu}{\sqrt{\sigma^2 + \varepsilon}} + \beta \quad (1)$$

其中, μ 和 σ 是 X 的均值与方差, ε 是为了除数不为零而加入的很小的整数, γ 和 β 是可训练的参数。

再利用GN中的训练参数计算归一化相关权重 W ,表示不同特征映射的重要性,其表达式如下:

$$W = \frac{\gamma_i}{\sum \gamma_i} \quad (2)$$

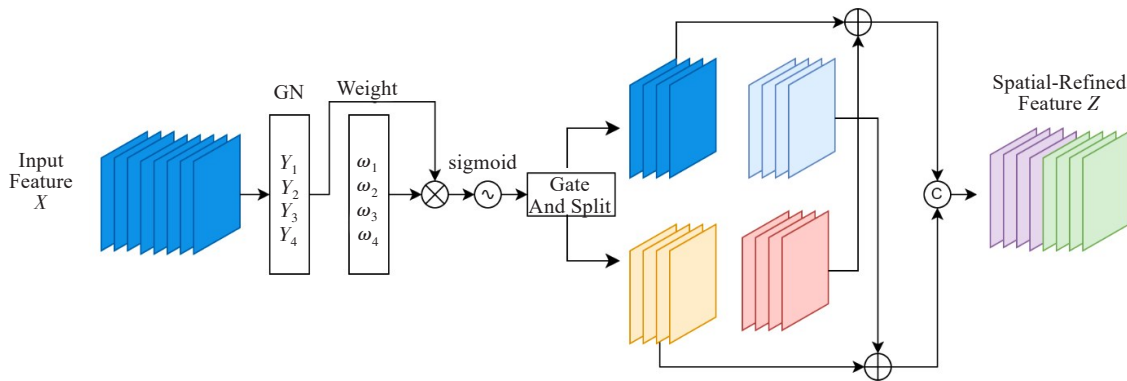


图1 空间重构单元结构
Figure 1 Spatial reconstruction unit structure

然后将经过 W 加权的特征通过 sigmoid 函数映射到 $(0,1)$ 范围^[20],随后通过门控机制,设定了一个阈值。按照 $>$ 阈值和 \leq 阈值的标准将数据划分成两个部分。其中一部分可以被认为是空间信息丰富的内容,而另一部分则是空间信息匮乏的内容。随后这两部分信息经过分割、相加、连接的重建操作得到空间重构特征 Z 。重建操作将信息丰富的特征与信息较少的特征相加,生成更丰富的特征。

CRU 采用分割、变换和融合策略来减少信道维

度的冗余以及计算成本和存储,其结构如图2所示。分割操作将经过 SRU 得到的的空间细化特征分割成两个部分,一部分通道数为 αC ,另一部分通道数是 $(1-\alpha) \cdot C$,随后对两组特征进行压缩。变换操作将经过分割后的两部分特征输入分别进行卷积、相加、连接。融合操作将变换得到的结果使用平均池化把全局空间信息和通道信息结合起来,做 Softmax 后得到权重向量,再与原结果相乘,随后按通道数分割成两部分再相加,就得到了通道提炼的特征。

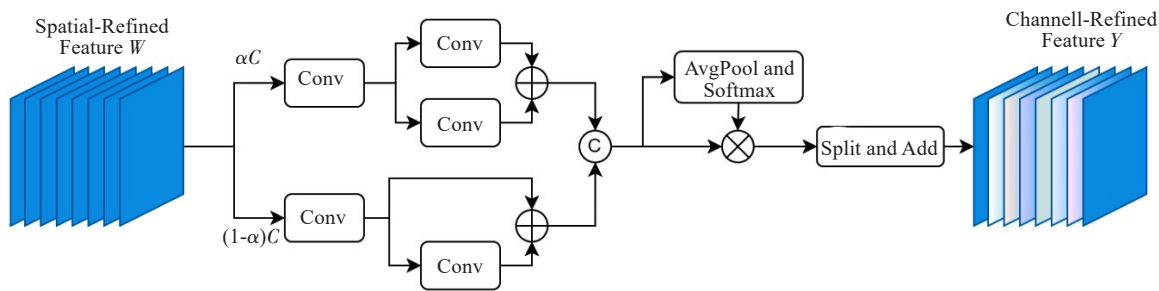


图2 通道重构单元结构
Figure 2 Channel reconstruction unit structure

结合 ScConv, 在 C2f 的基础上提出一种新的 C2f_ScConv 结构。将 C2f 的 Bottleneck 中的第二个卷积模块优化为 ScConv, 目的是减少密集模型参数中的固有冗余来提高性能, C2f_ScConv 结构如图3所示。

2.2 坐标(CA)注意力机制

由于卷积操作的局部感知特性,对图像中的位置和尺度相对敏感,导致对于不同位置和尺度的目标需要额外的处理和设计。在目标检测任务中,注意力机制能够让模型在观察目标时更具有针对性,使得目标识别与分类的精度都有所提升。它通过为输入的每个部分赋予不同的权重,抽取出更加关键及重要的信息,使模型做出更加准确的判断。同时,注意力机制不会

对模型的计算带来很大的开销。因此,笔者在原主干网络中嵌入CA注意力机制,CA注意力机制同时考虑了通道信息和空间信息,通过将位置信息嵌入到通道注意力中,学习到自适应的通道权重,使得模型更加关注有用的通道信息。CA注意力机制将通道注意力分解为两个一维特征编码过程,分别沿着两个空间方向聚合特征。通过这种方式,可以捕捉到一个空间方向上的长距离依赖关系,并同时保留另一个空间方向上的精确位置信息。然后,得到的特征图分别编码为一对方向感知和位置敏感的注意力图,可以互补地应用于输入特征图,以增强感兴趣对象的表示,CA注意力机制结构如图4所示。

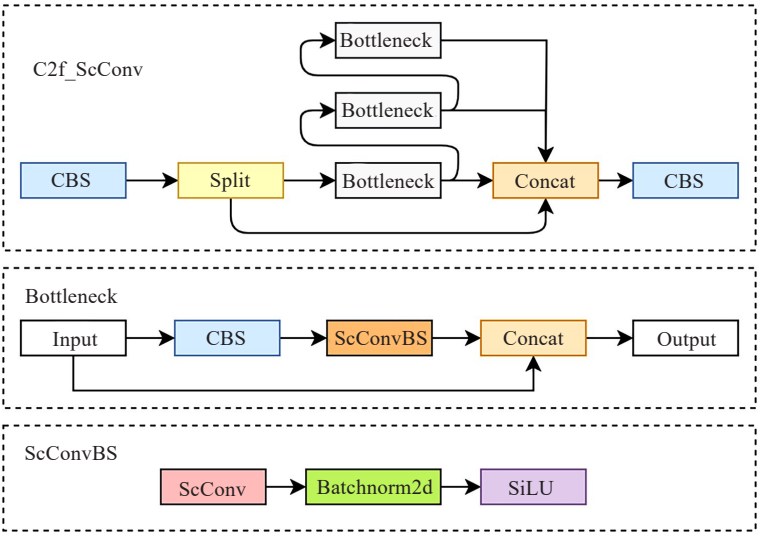


图3 C2f_ScConv 结构
Figure 3 C2f_ScConv structure

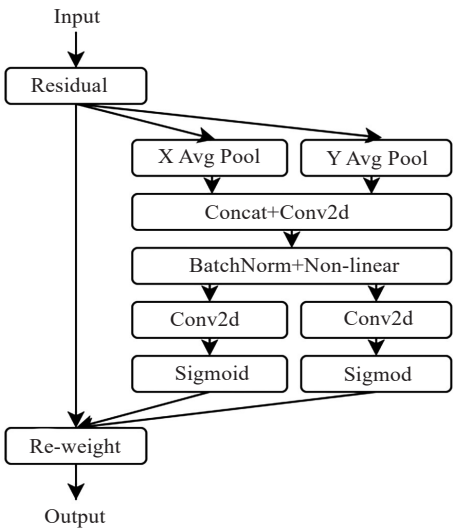


图4 坐标注意力结构
Figure 4 Coordinate attention structure

2.3 SPPCSPC

结核分枝杆菌在图像中的尺度较小,且在一定范围内变化。通过对数据集图像中的所有目标进行分析,发现在图像上呈现的长度大多在40像素左右,最小为11像素,最大为135像素。

空间金字塔池化(Spatial Pyramid Pooling, SPP)^[21]是一种在卷积神经网络中使用的技术,能够提取不同尺寸的空间特征信息,从而提升模型对于空间布局 and 物体变形的鲁棒性。跨阶段局部网络(Cross Stage Partial Network, CSP)^[22]结构通过将浅层的特征图在通道维度一分为二,一部分经由特征提取模块向后传播,另一部分则直接与特征提取模块的输出进行合并,实现了更加丰富的特征信息。

SPPCSPC 结合了 SPP 和 CSP 的结构,相较于原版 YOLOv8 中的 SPPF 能够更好地捕捉到尺度变化较大的结核分枝杆菌,所以笔者使用 SPPCSPC 替代 SPPF, SPPCSPC 结构如图 5 所示。

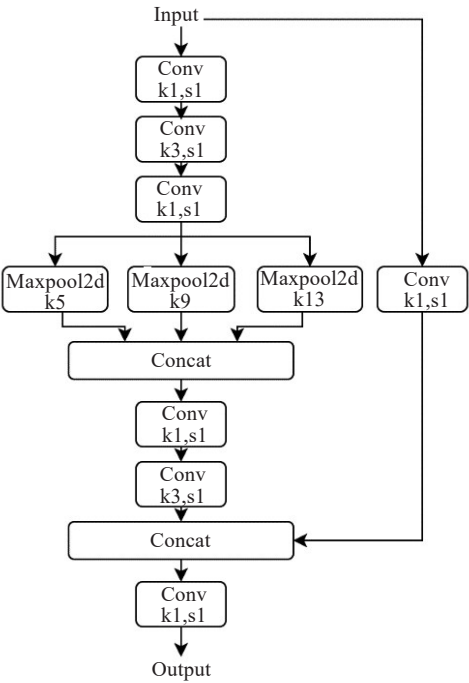


图5 SPPCSPC 结构
Figure 5 Spatial pyramid pooling cross-stage partial network structure

2.4 基于归一化的注意力机制

注意力机制帮助深度神经网络抑制较不显著的像素或通道。以往的许多研究都聚焦于通过注意力操作来捕捉显著特征。这些方法关注不同空间维度

或者通道之间的信息,但它们忽略了权重的贡献程度。NAM注意力机制采用卷积块注意力模块(Convolutional Block Attention Module, CBAM)^[23]的架构,无需额外的全连接、卷积等操作。重新设计了通道和空间注意力子模块,直接使用归一化中的缩放来计算注意力权重^[24],并通过增加正则化项来进一步抑制不显著的特征,其改进的子模块结构如图6所示。通过观察数据集我们了解到,部分结核分枝杆菌在图像中与空白部分较为相似。因此,笔者可以通过NAM注意力机制来抑制空白部分的权重,从而取得更好的检测效果。

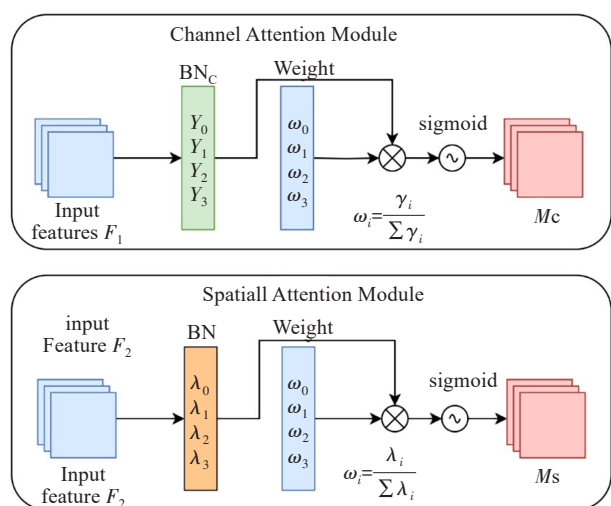


图6 改进的子模块结构

Figure 6 Improved submodule structure

对于两个子模块都使用了批归一化因子(Batch Normaztion, BN)的缩放因子,其表达式如下:

$$BN_{out} = BN(F) = \gamma \frac{F - \mu}{\sqrt{\sigma^2 + \epsilon}} + \beta \quad (3)$$

其中 BN_{out} 为输入特征经过BN后得到的输出, μ 为输入特征的均值, σ 为方差; γ 和 β 都是可训练的参数。同理,将BN的比例因子用在空间维度上。通道注意力模块和空间注意力模块所对应的输出表达式如下:

$$Mc = \text{sigmoid}\left(W_\gamma\left(BN(F_1)\right)\right) \quad (4)$$

$$Ms = \text{sigmoid}\left(W_\lambda\left(BN(F_2)\right)\right) \quad (5)$$

最终NAM注意力机制的结构图如图7表示。

综合以上改进思路,本文所提出方法使用的网络模型结构如图8所示。

3 实验设置与结果分析

3.1 实验环境

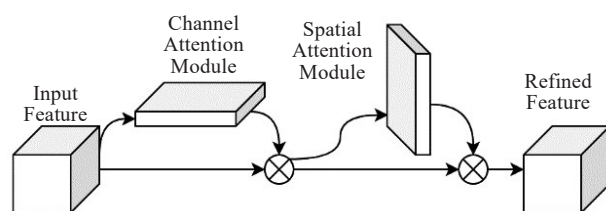


图7 NAM注意力机制结构

Figure 7 Normalization-based attention module structure

本文的所有实验环境为 Windows 11 操作系统,处理器是 Intel Core i9-10900X,内存 64 GB, NVIDIA GeForce RTX 3090 显卡,显存 24 GB。深度学习框架采用 PyTorch,编程语言采用 Python3.8, CUDA 版本 11.3。模型训练参数设置为:初始学习率 0.001,使用 Adam 优化器优化损失函数^[25],训练批次为 8,训练轮次为 200 轮,其余超参数使用 YOLOv8 的默认设置。

3.2 数据集

本文实验数据集采用 Kaggle 上公开的一个结核病图像数据集进行实验,且该数据集已采取 CC0 协议,贡献至公开领域。数据是原作者在实验室中使用专业的软件和硬件设备采集的光学显微镜下的痰样本涂片图像,共计 1 218 张图片,并且已经打好标签。将数据集按照 7:2:1 的比例划分为训练集、验证集和测试集。

3.3 评价指标

为了量化网络模型的检测性能,采用精确率(Precision)、召回率(Recall)和置信度为 0.5 时平均精度均值(mAP0.5)^[26]进行定量分析:

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

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

$$AP = \int_0^1 P(r) dr \quad (8)$$

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

其中,TP 为预测正确的正样本数量,FP 为预测错误的正样本数量,FN 为预测错误的负样本数量,AP 表示某一类别样本的 P-R(Precision-Recall)曲线下的面积,mAP0.5 是交并比(Intersection over Unit, IoU)阈值等于 0.5 时所有类别样本的 AP 均值,用于衡量模型对于所有类别样本的平均精准度。

3.4 消融实验

为了验证本文所提出的每种改进策略的有效性,以原始 YOLOv8s 为基础,采用逐步添加每个改进模块进行实验,观察其对网络性能的影响。实验结果如表 1 所示,从实验结果上来看 Cf2_ScConv 有效

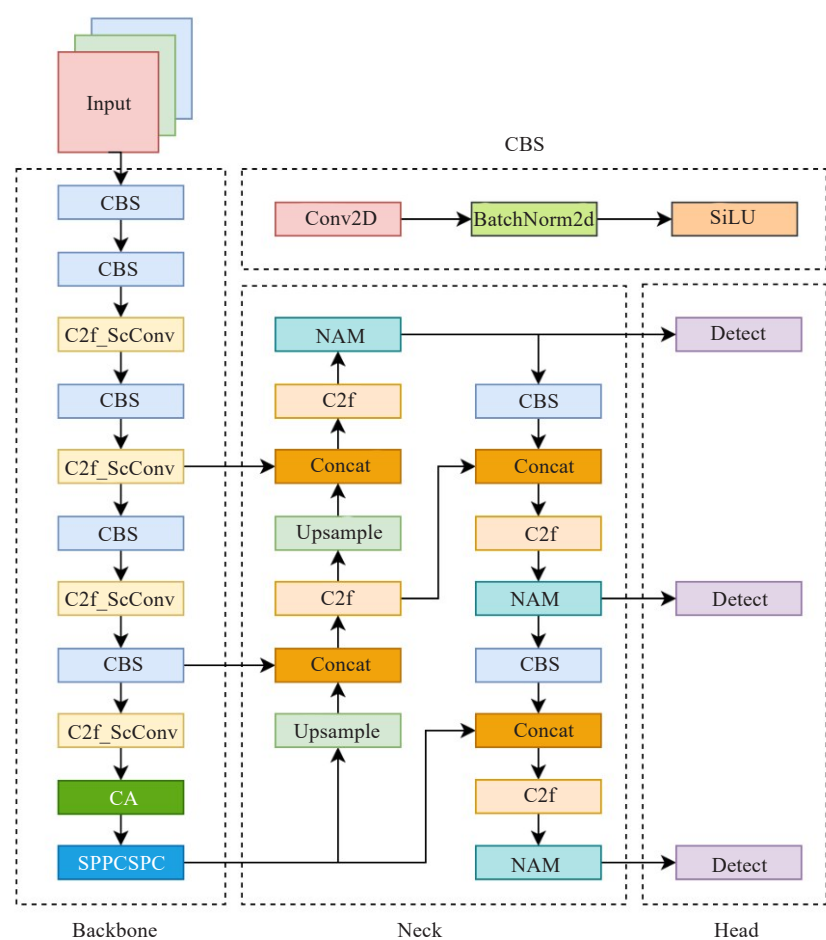


图8 改进的YOLOv8s检测算法结构

Figure 8 Improved YOLOv8s detection algorithm structure

地提升了模型的检测性能,说明该模块很好地去掉了冗余信息,提升了特征表示的有效性。CA的引入使模型关注到了全局信息,扩大了感受野。将SPPCSPC替换SPPF,使得模型在精确率和召回率上

都有提升,说明该模块学习到了不同尺度层的相互信息。NAM注意力机制的添加使得模型效果进一步提升,说明该模块通过权重控制,很好地抑制了不需要的信息。

表1 各改进模块对肺结核检测模型的影响

Table 1 Effect of each improved module on the tuberculosis detection model

模型	精确率	召回率	mAP0.5
YOLOv8s	0.770	0.774	0.842
YOLOv8s+C2f_ScConv	0.782	0.778	0.851
YOLOv8s+C2f_ScConv+CA	0.785	0.784	0.854
YOLOv8s+C2f_ScConv+CA+SPPCSPC	0.794	0.788	0.861
YOLOv8s+C2f_ScConv+CA+SPPCSPC+NAM	0.797	0.789	0.865

3.5 与其他目标检测算法的对比实验

为了验证本文提出的算法模型对结核病病原体检测的有效性,将该模型与Faster-RCNN、SSD和其他YOLO系列模型进行对比实验,实验结果如表2所示。

从实验结果可知,在相同的实验环境下,本文方法在测试集上均取得更好的结果,并且画面每秒帧数(Frames Per Second, FPS)已超过30,能达到实时检测的需求。本文模型在该数据集上的检测效果如图9~13所示。

表 2 不同检测模型实验结果对比

Table 2 Experimental result comparison of different detection models

模型	精确率	召回率	mAP0.5	FPS
Faster-RCNN	0.742	0.770	0.819	28.02
SSD	0.695	0.681	0.730	44.5
YOLOv3	0.770	0.746	0.816	55.04
YOLOv5	0.773	0.751	0.837	62.20
YOLOv7	0.758	0.778	0.825	38.94
本文模型	0.797	0.789	0.865	48.78



图 9 本文模型检测示例 1

Figure 9 Example 1 of the detection using the proposed model



图 10 本文模型检测示例 2

Figure 10 Example 2 of the detection using the proposed model

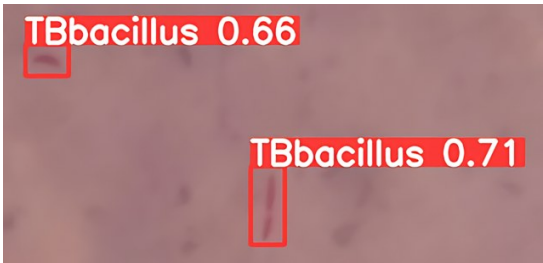


图 11 本文模型检测示例 3

Figure 11 Example 3 of the detection using the proposed model

4 结 论

本文在网络YOLOv8s的基础上进行改进使其更
适合结核病原体目标检测。首先提出 C2f_ScConv
结构替换原主干中的组件限制特征冗余；引入CA注



图 12 本文模型检测示例 4

Figure 12 Example 4 of the detection using the proposed model

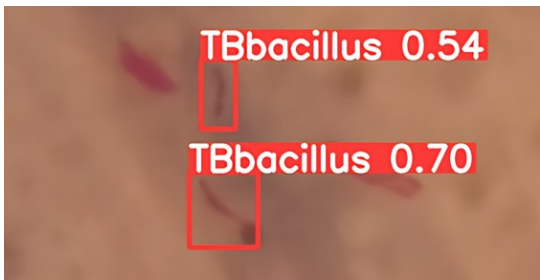


图 13 本文模型检测示例 5

Figure 13 Example 5 of the detection using the proposed model

意力机制提升感受野；使用SPPCSPC捕获不同尺度
上的信息；嵌入NAM注意力机制抑制冗余信息。实
验结果表明，本文模型相比原始YOLOv8s网络
mAP0.5提升了2.3%。性能也优于其他深度学习检
测方法。这表明本文方法在结核病原体检测上具
有可行性。但模型检测性能还有进一步提高的空
间，可以尝试引入其他结构例如Transformer来提升
效果。

【参考文献】

[1] 胡建平, 韩俊垒, 边红芝. 结核分枝杆菌 Rv2991、Rv3428c 重组蛋白对菌阴肺结核的辅助诊断价值[J]. 精准医学杂志, 2020, 35(6): 481-484.
Hu JP, Han JL, Bian HZ. Value of recombinant proteins Rv2991 AND Rv3428c of Mycobacterium tuberculosis in assisting the diagnosis of smear-negative pulmonary tuberculosis [J]. Journal of Precision Medicine, 2020, 35(6): 481-484.

[2] 舒薇, 刘宇红. 世界卫生组织《2023 年全球结核病报告》解读[J]. 结核与肺部疾病杂志, 2024, 5(1): 15-19.
Shu W, Liu YH. Interpretation of WHO global tuberculosis report 2023 [J]. Journal of Tuberculosis and Lung Disease, 2024, 5(1): 15-19.

[3] 高静楠, 刘宇红. 2019 年世界卫生组织全球结核病报告要点解读[J]. 河北医科大学学报, 2022, 43(7): 745-749.
Gao JT, Liu YH. Interpretation of the key points of the 2019 World Health Organization global tuberculosis report [J]. International Journal of Respiration, 2022, 43(7): 745-749.

[4] 刘宁, 付洪义, 孙金昊, 等. 四种病原学方法在检测痰液标本中结核杆菌的诊断价值对比[J]. 公共卫生与预防医学, 2021, 32(6): 107-110.
Liu N, Fu HY, Sun JH, et al. Comparative study on the diagnostic value of four pathogenic methods in the detection of Mycobacterium tuberculosis in sputum specimens [J]. Journal of Public Health and Preventive Medicine, 2021, 32(6): 107-110.

[5] 王新宇, 赵静文, 刘翔, 等. 人工智能在肺结节筛查和肺癌诊断中的应用[J]. 中国医学物理学杂志, 2023, 40(9): 1182-1188.

- Wang XY, Zhao JW, Liu X, et al. Applications of artificial intelligence in lung nodule detection and lung cancer diagnosis [J]. Chinese Journal of Medical Physics, 2023, 40(9): 1182-1188.
- [6] 王旭. 基于颜色和形态特征的痰涂片图像结核菌识别方法研究[D]. 厦门: 厦门大学, 2014.
- Wang X. Research of tuberculosis detection in sputum smear images based on color and shape [D]. Xiamen: Xiamen University, 2014.
- [7] Alfarhli FH, Mand AA, Sayeed MS, et al. Classification of tuberculosis with SURF spatial pyramid features [C]//2017 International Conference on Robotics, Automation and Sciences (ICORAS). Piscataway, NJ, USA: IEEE, 2017: 1-5.
- [8] 张璇. 基于改进SSD网络的结核杆菌目标检测算法研究[D]. 重庆: 重庆师范大学, 2020.
- Zhang X. The research on detection algorithm for targets of tuberculosis bacilli based on improved SSD network [D]. Chongqing: Chongqing Normal University, 2020.
- [9] Liu W, Anguelov D, Erhan D, et al. SSD: single shot MultiBox detector [C]//Computer Vision-ECCV 2016. Cham: Springer International Publishing, 2016: 21-37.
- [10] Yang A, Jin XY, Li LJ. CT images recognition of pulmonary tuberculosis based on improved faster RCNN and U-net [C]//2019 10th International Conference on Information Technology in Medicine and Education (ITME). Piscataway, NJ, USA: IEEE, 2019: 93-97.
- [11] Ren SQ, He KM, 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.
- [12] Ronneberger O, Fischer P, Brox T. U-net: convolutional networks for biomedical image segmentation [C]//Medical Image Computing and Computer-Assisted Intervention-MICCAI 2015. Cham: Springer International Publishing, 2015: 234-241.
- [13] 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 (CVPR). Piscataway, NJ, USA: IEEE, 2016: 779-788.
- [14] Wang CY, Bochkovskiy A, Liao HY. YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors [C]//2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2023: 7464-7475.
- [15] Zheng ZH, Wang P, Liu W, et al. Distance-IoU loss: faster and better learning for bounding box regression [C]//Proceedings of the AAAI Conference on Artificial Intelligence. Palo Alto, CA, USA: AAAI Press, 2020: 12993-13000.
- [16] Hou QB, Zhou DQ, Feng JS. Coordinate attention for efficient mobile network design [C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2021: 13708-13717.
- [17] Liu YC, Shao ZR, Teng YY, et al. NAM: normalization-based attention module [EB/OL]. (2021-11-24). <https://arxiv.org/abs/2111.12419>.
- [18] Li JF, Wen Y, He LH. SCConv: spatial and channel reconstruction convolution for feature redundancy [C]//2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2023: 6153-6162.
- [19] Wu YX, He KM. Group normalization [C]//Computer Vision-ECCV 2018. Cham: Springer International Publishing, 2018: 3-19.
- [20] Han J, Moraga C. The influence of the sigmoid function parameters on the speed of backpropagation learning [C]//From Natural to Artificial Neural Computation. Berlin, Heidelberg: Springer Berlin Heidelberg, 1995: 195-201.
- [21] He KM, Zhang XY, Ren SQ, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition [J]. IEEE Trans Pattern Anal Mach Intell, 2015, 37(9): 1904-1916.
- [22] Wang CY, Liao HY, 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). Piscataway, NJ, USA: IEEE, 2020: 1571-1580.
- [23] Woo S, Park J, Lee JY, et al. CBAM: convolutional block attention module [C]//Computer Vision-ECCV 2018. Cham: Springer International Publishing, 2018: 3-19.
- [24] Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift [C]//Proceedings of the 32nd International Conference on International Conference on Machine Learning. Chia Laguna Resort, Sardinia, Italy: PMLR, 2015: 448-456.
- [25] Kingma DP, Ba J. Adam: a method for stochastic optimization [EB/OL]. (2017-01-30). <https://arxiv.org/abs/1412.6980>.
- [26] Meschede C, Gros D, Habermann T, et al. Anaphora resolution: analysing the impact on mean average precision and detecting limitations of automated approaches [J]. Int J Inf Retr Res, 2018, 8(3): 33-45.

(编辑: 薛泽玲)