

DOI:10.3969/j.issn.1005-202X.2022.09.011

医学影像物理

多感受野的轻量化YOLOv4用于检测肺结核

王佳浩,王宝珠,郭志涛,王京华

河北工业大学电子信息工程学院,天津 300401

【摘要】肺结核疾病特征错综复杂,人工筛查成本较高,缺少规范的数据集。当前基于卷积神经网络的检测模型结构复杂、参数量大且检测精度有待进一步提高,为此提出一种改进的轻量化YOLOv4的肺结核检测模型。首先选取300例实际病例,制作一套规范的数据集,用于评估模型的性能;随后通过残差通道注意力模块改进MobileNetv3的结构,并作为YOLOv4的主干提取器,进一步减少参数量并融合上下文信息;然后在主干提取器的3个有效特征层后加入多感受野模块,有效增强低特征层的信息提取能力并降低对小型肺结核病灶的漏检率;最后,将以上改进的模块与YOLOv4的多尺度结构相结合,构建一种多感受野的轻量化YOLOv4的肺结核检测模型。与原始YOLOv4相比,该模型的参数量减少了约47%,平均精度(mAP)值提升至96.60%,漏检率降低至6%,验证该模型能有效辅助影像科医师诊断肺结核。

【关键词】肺结核;YOLOv4;MobileNetv3;多感受野

【中图分类号】R318;R521

【文献标志码】A

【文章编号】1005-202X(2022)09-1119-09

Lightweight YOLOv4 with multi-receptive fields for detection of pulmonary tuberculosis

WANG Jiahao, WANG Baozhu, GUO Zhitao, WANG Jinghua

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

Abstract: The characteristics of pulmonary tuberculosis are complex, with the high cost of manual screening, lack of standardized data sets. The current detection model based on convolution neural network has intricate structure, large number of parameters and detection accuracy needs to be further ameliorated. Therefore, an improved lightweight YOLOv4 model is proposed for pulmonary tuberculosis detection. A standardized dataset is constructed using 300 actual cases for evaluating the performance of the model. MobileNetv3 improved with residual channel attention module is used as the backbone extractor of YOLOv4 for further decreasing the number of parameters and fusing context information. Then the multi-receptive field module is added after the 3 effective feature layers of the backbone extractor, which effectively enhances the information extraction ability of the low feature layer and reduces the missed etection rate of small pulmonary tuberculosis lesions. The above improved modules were combined with the multi-scale structure of YOLOv4 to construct a lightweight YOLOv4 model with multi-receptive field for pulmonary tuberculosis detection. Compared with the original YOLOv4, the proposed model reduces the number of parameters of the model by about 47%, elevates the mAP value to 96.60%, and decreases the missed detection rate to 6%. It is verified that lightweight YOLOv4 with multi-receptive fields can effectively assist radiologists in the diagnosis of pulmonary tuberculosis.

Keywords: pulmonary tuberculosis; YOLOv4; MobileNetv3; multi-receptive field

前言

肺结核已经成为全球突发公共卫生事件,是由肺结核患者感染了一种由结核分枝杆菌引起的慢性传染病^[1]。根据世界卫生组织报道,结核病是世界上

主要的传染病致死原因之一。在估计的1 000万结核病病例中,每年约有64%被发现及治疗^[2]。早发现对此病的治疗具有重要意义^[3]。计算机断层扫描(CT)技术作为目前肺结核筛查的主要诊断方法,对比X光检查有更低的漏检率和误检率。CT对于结核病患者肺部实质性病变的识别和严重程度的诊断更有效率^[4]。

随着人工智能的发展,一些学者逐渐开始将深度学习的图像处理算法与CT技术进行融合,实现精度更高的肺部疾病的诊断和检测,以深度卷积神经

【收稿日期】2021-11-23

【基金项目】国家自然科学基金(61801164)

【作者简介】王佳浩,硕士,研究方向:智能信息处理、计算机视觉、机器学习、医疗图像处理,E-mail: wangjiahao1ad777@163.com

网络(Convolution Neural Network, CNN)为基础搭建了众多的计算机辅助诊断的深度学习模型。Gao等^[5]结合CT技术,提出基于CNN与支持向量机对5种肺结核的高精度分类模型。Ma等^[6]提出一种基于U-Net^[7]的活动性肺结核的自动检测模型,能够较精确地检测病灶位置。Liu等^[8]以肺癌的肺结节CT图像分别搭建模拟和真实数据集,并通过改进的单阶段目标检测网络YOLOv3^[9]实现对病灶的自动检测。Yang等^[10]以双阶段目标检测网络Faster-RCNN为主体,提高了肺结核的检测精准度,但其参数规模较大,不利于泛化。叶子勋等^[11]通过MobileNetv3^[12]替换了YOLOv4^[13]的主干网络,实现了模型轻量化,但对图像深层次特征信息的挖掘欠佳,未能解决YOLOv4对小目标的漏检问题,检测精准度仍有待提高。

肺结核CT影像主要有空洞^[14]、树芽征^[15]等典型特征,然而公开的数据集参差不齐,格式不统一,需要设计一套规范的数据集,况且现有肺结核检测模型拥有巨大的参数量,耗费大量的计算资源,难以适用于高性能设备匮乏的医疗机构,因此实现模型的轻量化是一项重要目标。同时,普通的轻量化检测模型未能获取CT图像更深的语义信息,存在漏检率高的问题,因此保证较高的肺结核检测精准度和检测效率是本研究的主要目标。

1 数据集与轻量化方法

1.1 数据集的制作

本研究的实验数据是由北京胸科医院的影像科提供的实际数据集,总共收集了300例肺结核病人的CT构成实验数据集,每张CT切片按照1.25 mm厚度标准进行分割,其中肺结核空洞与树芽征的CT切片各占50%。数据集中空洞的直径范围10.0~126.4 mm,且选取密度均匀的树芽征作为病灶样本,并规定CT切片上大小在16×16、32×32和64×64的区域分别为一簇树芽征的病灶。以上数据集标准是由北京胸科医院3位具有5年以上从业经验的影像科医师经过讨论后规定的,并使用工具LabelImg^[16]对肺结核病灶进行勾画和校准,为数据集的一致性和有效性提供保障。每张CT做了统一的预处理,其格式由DICOM换成了更便于处理的PNG格式,分辨率大小调整为512×512,编码位数为24。此外,为了减少对检测模型的干扰,滤除了肺部轮廓外的床板、衣物等背景信息。数据集共包含2764张空洞与树芽征的CT切片,从中随机抽取70%作为训练集,其余的前10%作为测试集1,后20%作为测试集2。测试集1负责用于后续的消融实验,测试集2则负责

评估肺结核检测模型的诊断水平。

数据增强能够丰富数据的多样性,因此对训练集的每张CT图像都有40%的几率进行放缩和水平翻转,并以在0.8~2.0范围内的宽高比对输入图像进行随机扭曲,将训练集的样本数量扩充至原来的2倍。本研究已经与北京胸科医院影像科签订患者信息保密协议,通过技术手段滤除患者的敏感信息,且该数据集的使用权仅在本研究中起效。

1.2 轻量化方法

目前基于CNN的检测模型具有参数量大和模型训练时间长的问題,对硬件的计算能力要求较高,因此本文将通过轻量化方法减少模型的参数量,使模型也适用于一般的硬件设备。

MobileNetv3是一种轻量化的神经网络,常用于图像处理领域,其主要基于深度可分离卷积原理实现轻量化^[17]。与传统卷积相比,深度可分离卷积有更少的参数量。深度可分离卷积是只与输入通道一一对应地进行遍历,再通过1×1的逐点卷积进行输出通道数的扩充。深度可分离卷积的原理如图1上部分所示。传统卷积和深度可分离卷积所需参数量分别设为 N_1 与 N_2 ,其计算公式如式(1)和式(2)所示:

$$N_1 = k^2 \cdot I_h \cdot I_w \cdot C_{in} \cdot C_{out} \quad (1)$$

$$N_2 = (k^2 + C_{out}) \cdot C_{in} \cdot I_h \cdot I_w \quad (2)$$

其中, k 为所选取的卷积核的尺寸, I_h 与 I_w 是输入张量的高和宽, C_{in} 与 C_{out} 分别代表输入和输出通道数的大小。

更进一步地如式(3)所示:

$$\frac{N_1}{N_2} = \frac{C_{out}}{1 + \frac{C_{out}}{k^2}} \quad (3)$$

在实际应用中, N_2 远远小于 N_1 ,即深度可分离卷积可以更大程度上减少计算开销。

文献[11]将MobileNetv3替换了YOLOv4的主干提取器CSPDarkNet53^[18],在很大程度上降低了参数量,但若用于检测肺结核,则难以实现高精度的目标。若以牺牲检测精度为代价提高模型的轻量化程度,这将与本研究的主要目标背道而驰。因此,本文将对MobileNetv3以及YOLOv4进行改进,减少肺结核检测模型的参数量,提高检测效率并改善精准度。

2 肺结核检测模型

开发的肺结核检测模型主要由3部分构成,即主干提取器、加强特征提取器和预测框生成器。本文将对主干提取器和加强特征提取器进行改进。

2.1 改进的主干提取器

主干提取器的作用是对肺结核CT图像进行初步解析,提取肺结核病灶的抽象特点。根据MobileNetv3的线性瓶颈的逆残差结构,本研究采用通道注意力机制,构建出新的IRCA (Inverted

Residuals Channel Attention)模块,该模块凭借残差结构能够搭建较深层的网络且不易出现梯度爆炸,引入的通道注意力可以在每个IRCA模块上捕获相应的上下文信息,提升网络的特征提取能力。IRCA模块的结构如图1下部分所示。

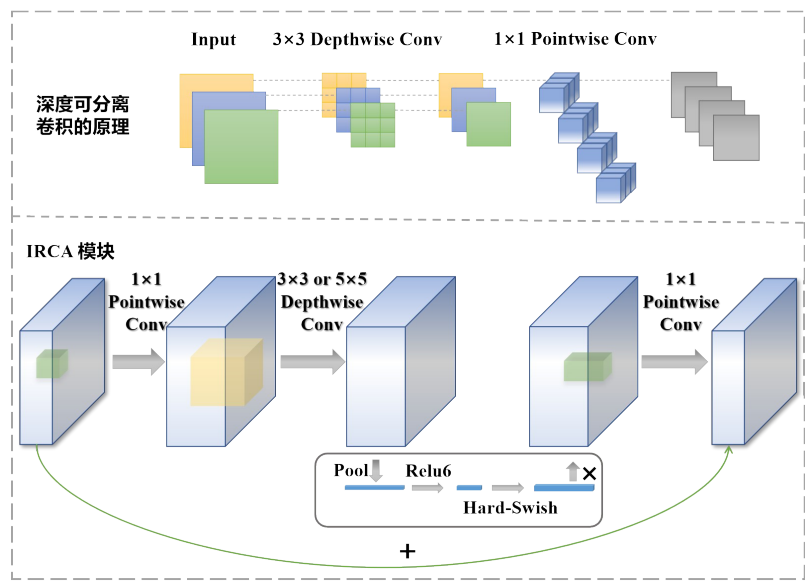


图1 深度可分离卷积的原理与IRCA模块
Figure 1 Principle of the depthwise separable convolution and IRCA module

IRCA 模块在完成1×1的卷积操作前先进行Average-Pooling获取尺寸更小的特征图,并通过Relu6和Hard-Swish激活函数得到特征图的权重。最后再进行加权相乘。Average-Pooling的计算公式如式(4)所示:

$$Y_{ap} = \frac{\sum_{i=0}^n \sum_{j=0}^n x_{i,j}}{n^2}$$
(4)

式中, $x_{i,j}$ 是特征图上的像素点, Y_{ap} 是Average-Pooling后的像素均值。

Relu6激活函数的计算公式如式(5)所示:
$$F_1(Y_{ap}) = \text{Min}\left(6, \text{Max}\left(0, Y_{ap}\right)\right)$$
(5)

式中, $F_1(Y_{ap})$ 为Relu6的激活结果,限制该函数取值上限为6。

Hard-Swish激活函数的计算公式如式(6)所示:
$$F_2(Y_{ap}) = \frac{Y_{ap}}{6} \cdot F_1(Y_{ap} + 3)$$
(6)

式中, $F_2(Y_{ap})$ 是Hard-Swish的激活结果。实验发现,在较深的卷积层中使用Hard-Swish激活函数可提高神经网络的精度。

$$F_3(X) = X \cdot F_2\left(F_1(Y_{ap})\right)$$
(7)

式(7)中深度可分离卷积的输出X是由n²个像素构成的

特征图, F_3 即整个IRCA模块的加权计算结果。
为了让主干提取器能获取不同尺度的病灶信息,本设计借鉴了PSPNet (Pyramid Scene Parsing Network)^[19]的结构,设计了1个金字塔池化模块作为主干提取器的池化层,其融合了4个不同尺度(32×32, 16×16, 8×8, 4×4)的池化核,进一步分析病灶的位置、大小等信息。
综上所述,本设计利用IRCA模块和金字塔池化模块改进了MobileNetv3,组合成新的主干提取器MIP。MIP的结构如图2所示。

2.2 改进的加强特征提取器

主干提取器会输出3个分辨率分别为原始图像的1/8、1/16和1/32的特征图,为后续加强特征提取器提供了不同大小肺结核病灶的像素信息。

YOLOv4初始的加强特征提取器主要由SPP (Spatial Pyramid Pooling)^[20]模块和PAN (Path Aggregation Network)^[21]构成。SPP会将主干提取器的最后一个特征层的结果作为输入,将特征图划分成3种不同大小的子区域(8×8、4×4和2×2)分别进行最大池化,汇集了多尺度的特征值。PAN是一种往复式的结构,负责收集主干提取器3个有效特征层的输出以构建特征金字塔,每一个有效特征层都融入了另外两层的采样信息,完成多尺度的特征融合。

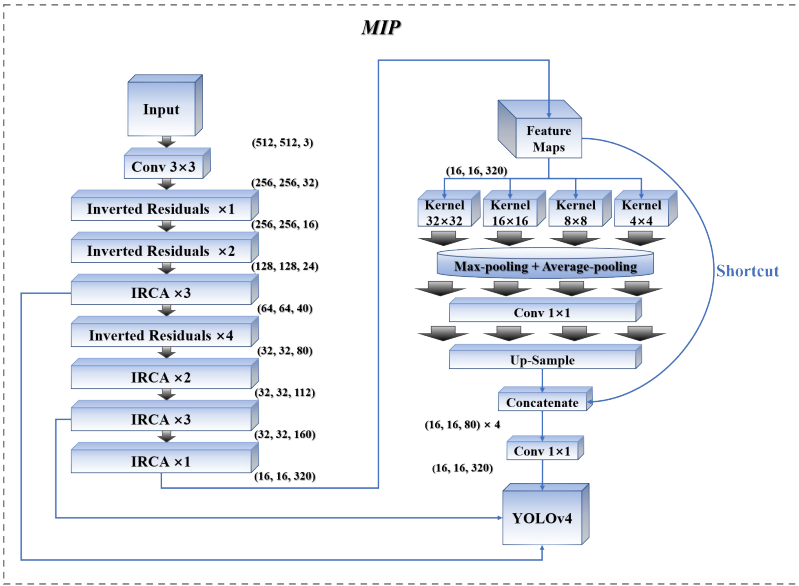


图 2 MIP 的结构
Figure 2 Structure of MIP

初始的加强特征提取器虽然利用多尺度的特征融合方法采集图像信息,但对多感受野信息的利用远远不够,容易忽略小目标的特征表达。因此,本文构建了一个RFB模块(Receptive Field Block)^[22]来改进YOLOv4的加强特征提取器。在原理上,RFB模块采用了并行的扩张卷积,其根据不同的扩张率与

卷积核对特征图进行感受野的扩增和特征采样,并加入残差短接防止特征信息丢失,最后拼接各卷积的采样结果以融合多感受野的信息。在结构上,RFB模块作为PAN的头部,与主干提取器MIP连接,弥补了PAN特征图信息收集不足的缺陷。RFB的结构如图3所示。

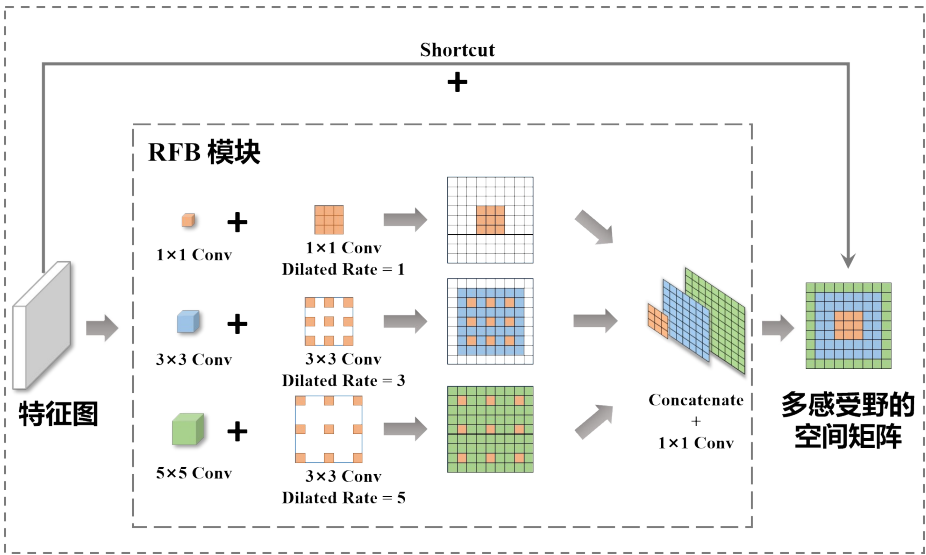


图 3 RFB 模块的结构
Figure 3 Structure of RFB module

2.3 预测框生成器

预测框生成器的作用是实现肺结核病灶信息(病灶大小、类别和位置)的回归,其由 3 个 Head 构成,每个 Head 在 PAN 采样得到的多尺度的特征图上生成 3 个不同尺寸的预测框来包围肺结核病灶,预测

框的参数包括中心点坐标、尺寸、病灶类别和置信度分数,但是这些参数信息并不能直接反映出最终的预测框在图片上的位置,还需要进一步对预测框的信息解码。

2.3.1 预测框的解码预测框 由 Anchor 解码得到,

Anchor是输入图片每个特征点上预定义边界框,其解码所参考的激活函数如式(8)所示:

$$\text{Sigmoid}(v) = \frac{1}{1 + \exp(-v)} \quad (8)$$

其中, v 可代表每个预测框的置信度分数或类别概率,通过 $\text{Sigmoid}(v)$ 函数将这两类参数映射到 $[0, 1]$ 范围。预测框中心点坐标的解码定义如式(9)和式(10)所示:

$$C_x = a_x + \text{Sigmoid}(x_{os}) \quad (9)$$

$$C_y = a_y + \text{Sigmoid}(y_{os}) \quad (10)$$

其中, x_{os} 与 y_{os} 是相对于预测框中心点的坐标偏移量。 (a_x, a_y) 为Anchor的中心点坐标,解码后得到预测框的中心点坐标 (C_x, C_y) 。预测框的高和宽的解码公式如式(11)和式(12)所示:

$$H = e^h \cdot a_h \quad (11)$$

$$W = e^w \cdot a_w \quad (12)$$

其中, h 和 w 是解码前预测框的高和宽, a_h 与 a_w 为Anchor的高和宽,分别计算得到解码后的预测框的高 H 和宽 W 。

解码后图片上同一位置的肺结核病灶会被大量预测框包围,因此需要根据设定好的交并比(Intersection of Union, IOU)阈值(通常设置为0.5)过滤冗余的预测框,再通过非极大抑制算法^[23]筛选出置信度分数最高的预测框。IOU的数学定义如式(13)所示:

$$\text{IOU} = \frac{|B \cap B^*|}{|B \cup B^*|} \quad (13)$$

其中,分母表示预测框 B 与真实框 B^* 面积的并集,分子表示交集。IOU反映了检测网络的预测结果与真实标签的相似程度。

2.3.2 组合损失函数肺结核检测 网络的组合损失函数主要由回归损失 L_{CIoU} 、类别损失 L_{Class} 和置信度损失 L_{Conf} 构成。检测网络的总体损失函数如式(14)所示:

$$L = L_{\text{CIoU}} + L_{\text{Class}} + L_{\text{Conf}} \quad (14)$$

式中,回归损失 L_{CIoU} 的数学表达式如式(15)所示:

$$\begin{cases} u_{ij} = \frac{4}{\pi^2} \left(\arctan \frac{W_{ij}^*}{H_{ij}^*} - \arctan \frac{W_{ij}}{H_{ij}} \right)^2 \\ \alpha_{ij} = \frac{u_{ij}}{1 - \text{IOU}_{ij} + u_{ij}} \\ L_{\text{CIoU}} = \sum_i \sum_j \left[1 - \text{IOU}_{ij} + \frac{\rho(B_{ij}, B_{ij}^*)}{D_{ij}^2} + \alpha_{ij} u_{ij} \right] \end{cases} \quad (15)$$

式中,下角标 (i, j) 代表特征图中第 i 个像素点上第 j 个预测框的序号, R^2 为特征图分辨率大小, N 是每个特

征图上预测框的数量。 W_{ij}^* 与 H_{ij}^* 为真实框的宽和高, $\alpha_{ij} u_{ij}$ 作为 L_{CIoU} 的惩罚因子, $\rho(B_{ij}, B_{ij}^*)$ 是预测框与真实框中心点的欧氏距离, D_{ij} 代表的是能够同时包含预测框与真实框的最小封闭区域的对角线距离。类别损失 L_{Class} 定义如式(16)所示:

$$L_{\text{Class}} = \sum_i \sum_j \sum_{t=1}^N \left[P_i(t) \lg(P_i^*(t)) + (1 - P_i(t)) \lg(1 - P_i^*(t)) \right] \quad (16)$$

式中, t 是类别序号, $P_i(t)$ 为网路判定的类别分数, $P_i^*(t)$ 则代表真实类别分数。

置信度损失 L_{Conf} 的数学模型如式(17)所示:

$$L_{\text{Conf}} = \sum_i \sum_j \left[O_{ij} \left[-\lg(Z_{ij}) \right] + \gamma \sum_i \sum_j (1 - O_{ij}) \left[-\lg(1 - Z_{ij}) \right] \right] \quad (17)$$

式中 O_{ij} 是判断预测框内是否包含肺结核病灶的二进制数, Z_{ij} 代表肺结核病灶的置信度,当预测框与真实框的IOU大于阈值时 γ 为1,否则取0。

经过上述对肺结核检测模型的结构改进和损失函数的选取,本文开发的肺结核检测模型命名为MIP-RY,该模型在YOLOv4基础上创建了轻量化的主干提取器MIP,并使用RFB模块替代了加强特征提取器部分冗余的卷积层,仅占用较少的内存,增强了多感受野信息的获取能力,最后通过组合损失函数在模型训练中完成肺结核病灶信息的回归与分类。

肺结核检测模型MIP-RY的整体结构见图4。

2.4 评价方法

肺结核检测模型的评价主要由参数量、单张CT的检测时间、查准率(Precision)、查全率(Recall)、误检率和平均精准度(mean Average Precision, mAP)决定^[24]。其中查准率 E_{pr} 如式(18)所示:

$$E_{\text{pr}} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (18)$$

式中,TP为样本的真阳性,FP为假阳性,二者之和为模型的预测结果。

查全率 E_{re} 如式(19)所示:

$$E_{\text{re}} = \frac{\text{TP}}{\text{FP} + \text{FN}} \quad (19)$$

其中,FN为假阴性,FP与FN之和为真实框总量。

本研究中误检率 E_{m} 是将验证集2中肺结核CT的平均误检率映射到对数空间,其数学定义如式(20)所示:

$$\begin{cases} \lambda_i = \frac{\text{FP}_i}{\text{TP}_i + \text{FP}_i} \\ E_{\text{m}} = \exp \left[\frac{1}{S} \sum_{i=1}^S \ln(\max(\lambda_i, \mu)) \right] \end{cases} \quad (20)$$

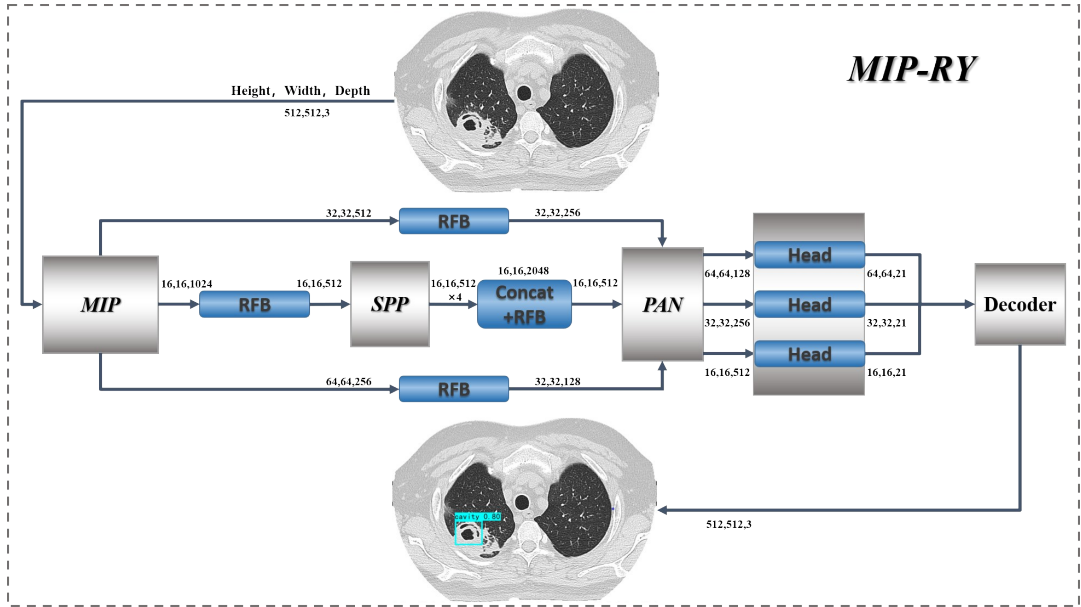


图 4 MIP-RY 的整体结构

Figure 4 Overall structure of MIP-RY

其中, λ_i 代表单张 CT 切片的误检率, S 为 CT 切片总数。考虑到对数函数可能会出现无数学定义的情况, 故加入较小值 μ 防止对数函数的自变量为 0。

设置阈值为 0.5 的 IOU 以判定样本的 TP 与 FP, 分别以查全率和查准率为横纵坐标构建检测网络的 P-R 曲线, mAP 等于所有类别的 P-R 曲线下面积的均值, mAP 的数学定义如式 (21) 所示:

$$\text{mAP} = \frac{\sum_0^n \int E_{\text{pr}} \cdot E_{\text{re}} d(E_{\text{re}})}{n} \tag{21}$$

其中, n 为类别总数。

3 实验设置与结果分析

本实验的运行系统为 Windows 10, 处理器为 Intel Core i7 和 RTX 2060。以 PyCharm 为集成开发环境, 使用基于 Python3.8 的 Keras 深度学习框架。为适应模型的训练, 初始学习率大小设为 1.0×10^{-4} , 选用 Adam^[25] 优化器和 500 次迭代训练。最后采用余弦退火衰减策略 (Cosine Annealing Attenuation, CAA)^[26] 使肺结核检测模型能够根据迭代次数自适应地调整学习率。

3.1 消融实验

为验证本文改进点的有效性, 利用测试集 1 对各个改进模块 (MIP 和 RFB) 进行消融实验, 以观察不同模块对肺结核检测模型性能的影响。为了便于区分, 将文献 [11] 的轻量化模型 MobileNetv3-YOLOv4 命名为模型 1, 仅使用 MIP 的模型命名为模型 2, 仅使用 RFB 的模型命名为模型 3, 以及融合了所有改进模块的 MIP-RY。消融实验结果如表 1 所示。由表 1 中

数据可以推断出, 较模型 1, MIP 模块使模型 2 仅增加了 0.93M 的参数量, 平均精准度却提高了 2.64%, 误检率下降了 3%。模型 3 仅使用 RFB, 使其平均精准度提升至 91.92%, 但由于其未使用轻量化模块 MIP, 导致其参数量仍然很高, 检测效率低下。本文提出的 MIP-RY 结合了两个改进模块的优点, 保持轻量化结构的同时又扩大了对特征图的感受野, 使其平均精准度跃升至 95.59%, 比模型 1 高出 9.73%, 参数量下降了 4.73M, 误检率下降了 10%, 同时缩短了 1.07 s 的检测时间。

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 pulmonary tuberculosis detection model

模型	平均精准度/%	参数量/M	误检率/%	检测时间/s
模型 1	85.86	38.64	16	6.95
模型 2	88.50	39.57	13	7.87
模型 3	91.92	58.29	10	8.86
MIP-RY	95.59	33.91	6	5.88

3.2 主流模型对比实验

通常高效率、高精度、低内存占用的肺结核检测模型才易适用于临床诊断。为进一步验证 MIP-RY 模型的可靠性, 使用测试集 2, 与目前几个主流的目标检测模型进行对比实验。测试集 2 中包含 359 个空洞样本与 546 个树芽征样本, 这些 CT 样本被用于评估不同检测模型的性能。不同模型检测肺结核空洞和树芽征的 P-R 曲线如图 5 所示。由图 5 可知, 相比其他主流检测

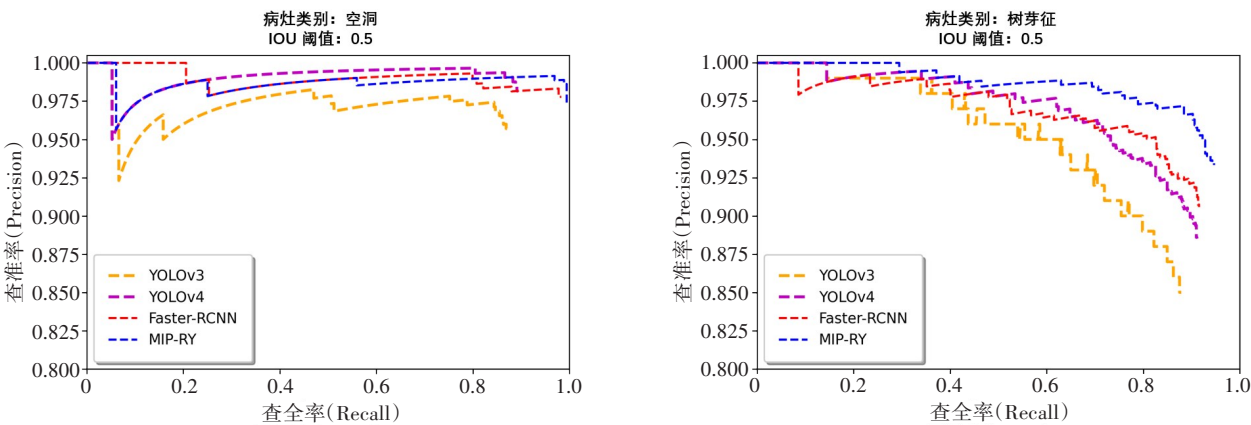


图5 不同模型检测空洞和树芽征的P-R曲线

Figure 5 P-R curves of cavity and tree-in-bud pattern detected by different models

模型,MIP-RY的P-R曲线包围的范围最大。另外,MIP-RY对于检测树芽征的P-R曲线的优化效果最为明显,说明该模型更容易捕捉到成簇的树芽征。

不同肺结核检测模型在测试集2上的评估结果如表2所示。经过定量分析,文献[9]的YOLOv3和文献[13]的YOLOv4单阶段检测模型性能欠佳,这是由于在测试集2中有较多的小型空洞和小簇的树芽征,容易与肺区内正常的支气管混淆,这也导致这两个模型有过高的漏检率。相比之下,文献[10]采用的双阶段检测模型有良好的mAP和较低的漏检

率,但其模型参数量巨大,不适用于算力一般的设备,且检测时间略长,难以满足高效率的要求。本研究提出的MIP-RY模型在测试集2上表现优越,对比YOLOv4模型,查准率提升了7.49%,查全率提升了4.47%,mAP提升了9.06%,参数量减少了约47%,误检率降低了8%。此外MIP-RY也拥有较高的检测效率,与Faster-RCNN相比,对单张肺结核CT的检测时间缩短了6.53 s。从评估结果的数值上分析,MIP-RY能够满足高检测精度的要求,其模型参数占用空间较小,进一步完成了轻量化任务。

表2 各检测模型在测试集2上的评价指标

Table 2 Evaluation results of each detection model on test set 2

模型	查准率/%	查全率/%	平均精准度/%	误检率/%	参数量/M	检测时间/s
YOLOv3 ^[9]	86.47	75.32	82.86	18	61.53	9.75
Faster-RCNN ^[10]	96.96	85.54	93.91	8	136.65	12.18
YOLOv4 ^[13]	90.93	83.03	87.54	14	63.94	8.93
MIP-RY	98.42	87.50	96.60	6	33.91	5.65

本研究提出的模型MIP-RY的肺结核检测效果需要更为直观的展示。各模型对空洞的检测示例如图6和图7所示,树芽征的检测示例如图8和图9所示。Ground Truth为影像科医师规定的真实标签,肺结核空洞被包围框标注为“cavity”,树芽征被包围框标注为“tree-in-bud”,相应的置信度分数附加在包围框上,置信度分数反映了病灶身份的可信性。由检测示例分析出YOLOv3检测空洞和树芽征的置信度分数较低,造成此现象可能的原因是YOLOv3对多尺度特征信息利用不足。YOLOv4虽然有较高的置信度分数,但它在尺寸较小的病灶上出现了漏检,原因是主干提取器的下采样尺度过大,在逐层的特征提取中忽略了一部分小型病灶的空间和像素信息,

但若减小下采样尺度,则不能保证对其他目标的检测精度。图7表明Faster-RCNN能够精准捕捉到各种大小的空洞,但在图9中也出现了对树芽征的漏检。RFB模块能够有效解决病灶的漏检问题,可以观察到MIP-RY捕捉到了难以识别的小型空洞和小簇树芽征,说明RFB模块的多感受野信息收集能力发挥了作用,避免了小目标被漏检的风险。

4 结 论

本研究开发了一种多感受野的轻量化检测模型MIP-RY,用于自动检测肺结核。MIP-RY结合深度可分离卷积原理和注意力机制,以较小的参数量获取肺结核病灶的特征信息。引入的RFB模块则充分

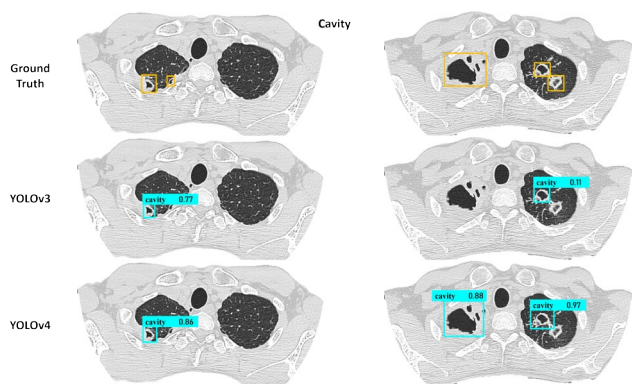


图6 空洞检测示例1

Figure 6 Example 1 of cavity detection

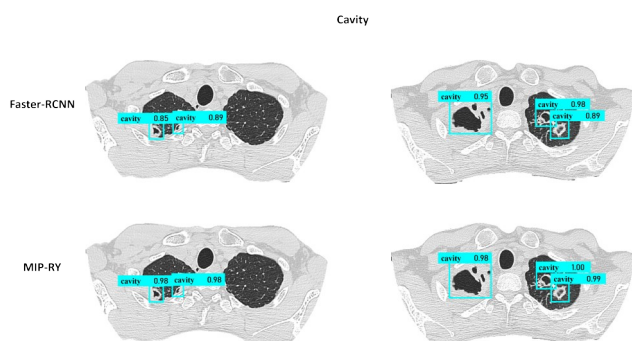


图7 空洞检测示例2

Figure 7 Example 2 of cavity detection

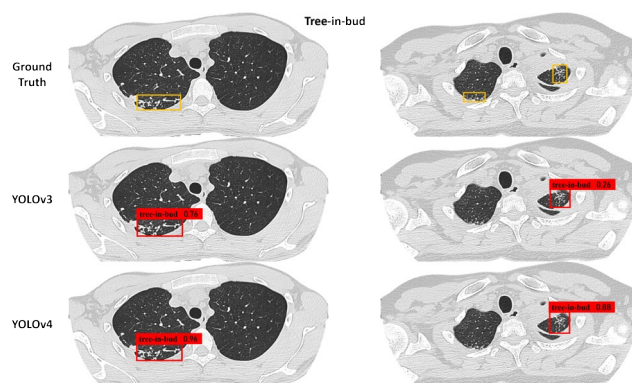


图8 树芽征检测示例1

Figure 8 Example 1 of tree-in-bud pattern detection

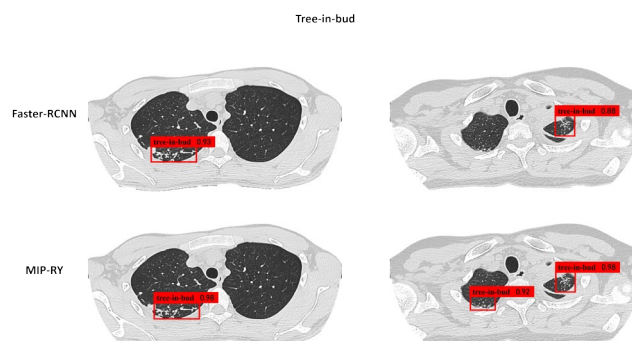


图9 树芽征检测示例2

Figure 9 Example 2 of tree-in-bud pattern detection

利用了多感受野的空间信息,使预测包围框能自适应地建立不同大小区域的相关性,提升对小病灶的识别能力。实验结果表明,改进的肺结核检测模型达到了较高检测精度,以更低的计算成本实现肺结核空洞与树芽征的自动检测,具有良好的肺结核影像诊断潜力。由于数据获取的限制,更多的肺结核影像表现,如小叶中央结节和肺实变,尚未能纳入本研究。后续将引入其他类型的肺结核病变以扩充数据集,并且尝试将肺癌等疾病的检测迁移学习到肺结核的研究中,这对今后各类肺部疾病的辅助诊断有一定的借鉴价值。

【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.
- [2] Hu JP, Han JL, Bian HZ. The value of recombinant protein Rv2991 and Rv3428c of *Mycobacterium tuberculosis* in the diagnosis of negative pulmonary tuberculosis[J]. Journal of Precision Medicine, 2020, 35(6): 481-484.
- [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]. 国际呼吸杂志, 2020, 40(3): 161-166.
- [4] Gao JT, Liu YH. Interpretation of the key points of the 2019 World Health Organization global tuberculosis report [J]. International Journal of Respiration, 2020, 40(3): 161-166.
- [5] 杜莹, 张彬, 王大宽, 等. 非活动性肺结核人群的流行病学特征和预防干预[J]. 中国防痨杂志, 2020, 42(11): 1232-1236.
- [6] Du Y, Zhang B, Wang DK, et al. Epidemiological characteristics and preventive intervention of inactive pulmonary tuberculosis [J]. Chinese Journal of Antituberculosis, 2020, 42(11): 1232-1236.
- [7] Bhalla AS, Goyal A, Guleria R, et al. Chest tuberculosis: radiological review and imaging recommendations[J]. Indian J Radiol Imaging, 2015, 25(3): 213-215.
- [8] Gao X, Qian Y. Prediction of multi-drug resistant TB from CT pulmonary images based on deep learning techniques[J]. Mol Pharm, 2018, 15(10): 4326-4335.
- [9] Ma L, Wang Y, Guo L, et al. Developing and verifying automatic detection of active pulmonary tuberculosis from multi-slice spiral CT images based on deep learning[J]. J Xray Sci Technol, 2020, 28(5): 939-951.
- [10] Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation[M]. Cham: Springer, 2015.
- [11] Liu C, Hu SC, Wang C, et al. Automatic detection of pulmonary nodules on CT images with YOLOv3: development and evaluation using simulated and patient data[J]. Quant Imaging Med Surg, 2020, 10(10): 1917-1929.
- [12] Redmon J, Farhadi A. YOLOv3: an incremental improvement[J]. arXiv e-prints, 2018. DOI: 10.48550/arXiv.1804.02767.
- [13] Yang A, Jin X, Li L. 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). 2019.
- [14] 叶子勋, 张红英. YOLOv4口罩检测算法的轻量化改进[J]. 计算机工程与应用, 2021, 57(17): 157-168.
- [15] Ye ZX, Zhang HY. Lightweight improvement of YOLOv4 mask detection algorithm[J]. Computer Engineering and Applications, 2021, 57(17): 157-168.
- [16] Howard A, Sandler M, Chen B, et al. Searching for Mo-bileNetV3[C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV). IEEE, 2020.
- [17] Bochkovskiy A, Wang CY, Liao H. YOLOv4: optimal speed and accuracy of object detection[J]. arXiv: 2004.10934.
- [18] Dheda K, Lenders L, Srivastava S, et al. Spatial network mapping of

- pulmonary multidrug-resistant tuberclosiscavities using RNA sequencing[J]. Am J Respir Crit Care Med, 2019, 200(3): 270-280.
- [15] Sharma A, Narayanan B, Pal N, et al. An unusual cause of 'tree-in-bud' appearance in CT-chest during COVID-19 pandemic[J]. Monaldi Arch Chest Dis, 2020. DOI: 10. 4081/monaldi. 2020. 1466.
- [16] Hidayat AA, Cenggoro TW, Pardamean B. Convolutional neural networks for scopsowl sound classification[J]. Proc Comput Sci, 2021, 179(4): 81-87.
- [17] Tu CH, Lee JH, Chan YM, et al. Pruning depthwiseseparable convolutions for mobileNetcompression[C]//2020 International Joint Conference on Neural Networks (IJCNN). 2020.
- [18] Cai YF, Luan TY, Gao HB, et al. YOLOv4-5D: an effective and efficient object detector for autonomous driving [J]. IEEE Trans Instrum Meas, 2021. DOI: 10. 1109/TIM. 2021. 3065438.
- [19] Zhao H, Shi J, Qi X, et al. Pyramid scene parsing network [C]// Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 2881-2890.
- [20] Purkait P, Zhao C, Zach C. SPP-Net: deep absolute pose regression with synthetic views[C]// British Machine Vision Conference (BMVC 2018). 2017.
- [21] 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.
- [22] Liu S, Huang D, Wang Y. Receptive field block net for accurate and fast object detection[C]//European Conference on Computer Vision. Cham: Springer, 2018.
- [23] Nayerreh 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.
- [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.
- [25] 唐思豪, 滕召胜, 孙彪, 等. ADAM改进BP神经网络与动态称重应用[J]. 电子测量与仪器学报, 2021, 35(4): 127-135.
Tang SH, Teng ZS, Sun B, et al. ADAM improved BP neural network and application of dynamic weighing [J]. Journal of Electronic Measurement and Instrument, 2021, 35(4): 127-135.
- [26] 王萌, 曾艳, 彭飞, 等. 基于余弦下降学习率的卷积网络轴承故障诊断[J]. 工业技术与职业教育, 2020, 18(3): 4-8.
Wang M, Zeng Y, Peng F, et al. Convolution network bearing fault diagnosis based on cosine descent learning rate [J]. Industrial Technology and Vocational Education, 2020, 18(3): 4-8.

(编辑:黄开颜)