

基于改进 Faster R-CNN 的肺结核病原体检测^①



鞠瑞文¹, 孙 振¹, 李庆党²

¹(青岛科技大学 数据科学学院, 青岛 266061)

²(青岛科技大学 中德科技学院, 青岛 266061)

通信作者: 李庆党, E-mail: lqd@qust.edu.cn

摘 要: 本文提出一种基于 Faster R-CNN 的肺结核病原体检测方法, 以更高的准确率和更低的漏检率检测肺结核。首先, 通过 Mosaic 数据增强方法对数据集进行扩充, 提高模型泛化能力, 同时引入 K-means 聚类算法, 对所用数据集重新聚类来生成对的锚点初始候选框大小。其次, 将 Faster R-CNN 中的原始特征提取网络替换为 Res2Net, 并将其卷积核全部替换为空洞卷积, 在与原卷积相比参数量不变的情况下, 增大了感受野。再者, 引入改进后的 GC-FPN 模块, 使模型在轻量化的同时更好的关注小目标信息。最后, 引入 ROI Align, 解决候选框和初始回归位置存在偏差的问题。实验结果表明, 在公开数据集上, 改进的 Faster R-CNN 模型与原本的 Faster R-CNN 算法相比, 准确率提高了 2.7%, 召回率提升了 1.4%, 该算法不仅在结核图像数据集上得到了验证, 而且具有较高的准确率。

关键词: 肺结核; 深度学习; 目标检测; Faster R-CNN

引用格式: 鞠瑞文, 孙振, 李庆党. 基于改进 Faster R-CNN 的肺结核病原体检测. 计算机系统应用, 2024, 33(11): 121-130. <http://www.c-s-a.org.cn/1003-3254/9679.html>

Tuberculosis Pathogen Detection Based on Improved Faster R-CNN

JU Rui-Wen¹, SUN Zhen¹, LI Qing-Dang²

¹(College of Data Science, Qingdao University of Science and Technology, Qingdao 266061, China)

²(Sino-German Institute of Science and Technology, Qingdao University of Science and Technology, Qingdao 266061, China)

Abstract: In this study, a detection method for tuberculosis pathogens based on Faster R-CNN is proposed to detect tuberculosis with higher accuracy and lower missed detection rate. First, the Mosaic data enhancement method is used to expand the dataset to improve the generalization ability of the model. At the same time, the K-means clustering algorithm is introduced to re-cluster the used dataset to generate the initial candidate box size of the paired anchor points. Secondly, the original feature extraction network in Faster R-CNN is replaced with Res2Net, and all its convolution kernels are replaced with empty convolution. This can bring a larger receptive field compared with the original convolution when the number of parameters remains unchanged. Furthermore, the improved GC-FPN module is introduced to make the model pay more attention to small target information while being lightweight. Finally, ROI Align is introduced to solve the problem of deviation between the candidate box and the initial regression position. The experimental results show that, compared with the original Faster R-CNN algorithm, the improved Faster R-CNN model has a 2.7% higher accuracy and an 1.4% higher recall rate on the open data set. This algorithm has been verified on the dataset of tuberculosis images and possesses high accuracy.

Key words: tuberculosis; deep learning; object detection; Faster R-CNN

① 基金项目: 山东省泰山学者项目 (tshw201502042)

收稿时间: 2024-04-22; 修改时间: 2024-05-20; 采用时间: 2024-06-04; csa 在线出版时间: 2024-09-27

CNKI 网络首发时间: 2024-09-29

结核病,一种由结核分支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]。《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,死于该病人数约为3万。

肺结核病的诊断方法主要有痰结核杆菌检查和医学影像检查,后者需要先通过计算机断层扫描(computed tomography, CT)技术进行影像拍摄,再将CT影像交由医生进行检查。虽然医学影像检查速度较快,但检测费用较高,且医学影像诊断结果大多取决于医师的主观判断与临床经验,并且这些可能受到医师诊断水平或周边环境等客观因素干扰,在一定程度上会对结核病的筛查产生不利的影响。而痰结核杆菌检测虽然检测时间相比于医学影像检测有所增加,但痰结核杆菌检测成本较低,且检测准确率比医学影像检测有所提高。

随着计算机科学的蓬勃发展,人工智能的应用无形中改变了线上线下的医疗模式,实现多学科融合、多路径诊断、多路径治疗的智能互联网医疗体系。目标检测作为计算机视觉领域的一个研究方向,在医学领域得到了广泛应用,如CT检测、血液检测等,发挥着无可替代的作用^[2]。

目前国内外有大量人员通过构建目标检测的神经网络来实现对结核病的检测。孙涵璇等人^[3]采用了Top-hot变换进行图像数据增强,将增强后的图像阈值化,并引入了目标和背景的颜色特征以及一个边界形状特征,之后使用BP网络进行目标分类,然后使用机器学习算法对结核杆菌进行识别。Nayak等人^[4]提出基于HSI颜色空间分割和区域分类应用于结核杆菌的自动化检测。以上两种是使用传统算法对结核杆菌进行识别的研究,其检测精度和效率相比于现在使用神经网络和深度学习算法的检测方法都有所不足;卞景帅等人^[5]采用深度学习处理方法,将Faster R-CNN目标检测算法与医学检测相结合,对结核杆菌进行检测,崔少国等人^[6]提出了一种将深度可分离卷积和图卷积相结合的肺结核疾病检测方法,使得将肺结核疾病检测准确率提升至89.55%,以上两种方法检测率较好,但其模型参数量较大,导致检测速度较慢。王佳浩等人^[7]提出了多感受野的轻量化YOLOv4用于检测肺结核,在YOLOv4的基础上加入轻量化的MobileNetV3网络,并构建了一个RFB(receptive field block)模块来改

进YOLOv4的加强特征提取器,其检测速度较快,但模型平均精度还有待提高。

通过对上述研究者贡献的总结,可以发现当前对于结核病病原体的检测方法研究仍然存在一定的发展空间,因此,本文提出一种基于Faster R-CNN的肺结核病原体检测算法。通过实验对比发现,本文所提出的检测算法能够提高网络的检测精度,并且漏检率较低。

1 Faster R-CNN 目标检测算法

Faster R-CNN^[8]是一种端到端的深度学习算法,适用于目标检测任务,它由Girshick等人于2015年提出。相比于传统的目标检测方法,Faster R-CNN通过引入区域提议网络(region proposal network, RPN)实现了在单一网络中进行目标定位和分类,无需额外的预处理步骤,极大地提高了检测速度和准确性。首先,RPN在输入图像上滑动窗口来生成候选目标框,同时用分类和回归分支对这些框进行评分和调整,从而提取出潜在的目标区域。然后,经过ROI(region of interest)池化,将不同大小的候选框映射到固定大小的特征图上,保持了特征的空间位置信息。最后,利用全连接层对ROI中的特征进行目标分类和边界框回归。整个网络采用端到端的训练策略,通过共享的特征提取网络、RPN网络和Fast R-CNN网络共同训练,避免了传统目标检测方法中多阶段的训练流程,使其在速度和准确性上都取得了显著提升。Faster R-CNN在物体检测、行人检测、交通场景分析等领域得到广泛应用,成为目标检测领域的主流方法之一,为实时目标检测提供了强大的技术支持。其模型架构图如图1所示。

2 基于改进Faster R-CNN的肺结核病原体检测算法

通过实验,本文提出了一种基于Faster R-CNN的改进模型,首先,该模型在原始Faster R-CNN的基础上,通过Mosaic数据增强方法对数据集进行扩充,增加训练数据,提高模型泛化能力,同时可以使模型在更小的范围内识别目标;然后引入K-means聚类算法,对所用数据集重新聚类来生成对的锚点初始候选框大小。对于模型的主干部分,将Faster R-CNN中的原始特征提取网络(backbone)替换为Res2Net,并将其卷积核全部替换为空洞卷积(atrous convolution),在与原卷积相比参数量不变的情况下,增大了感受野。同时加入改进后

的 GC-FPN 模块, 在 FPN 的基础上引入了自底向上的路径, 即 PANet 使得底层信息更容易传递到高层顶部, 同时引入幻影卷积 (GhostConv) 和 CA 注意力机制, 使模型在轻量化的同时更好的关注小目标信息, 从

而提高整个模型对小目标的检测效果. 最后, 引入 ROI Align, 解决 ROI pooling 经过量化, 候选框已经和最开始回归出来的位置有一定的偏差所产生的不匹配问题 (misalignment).

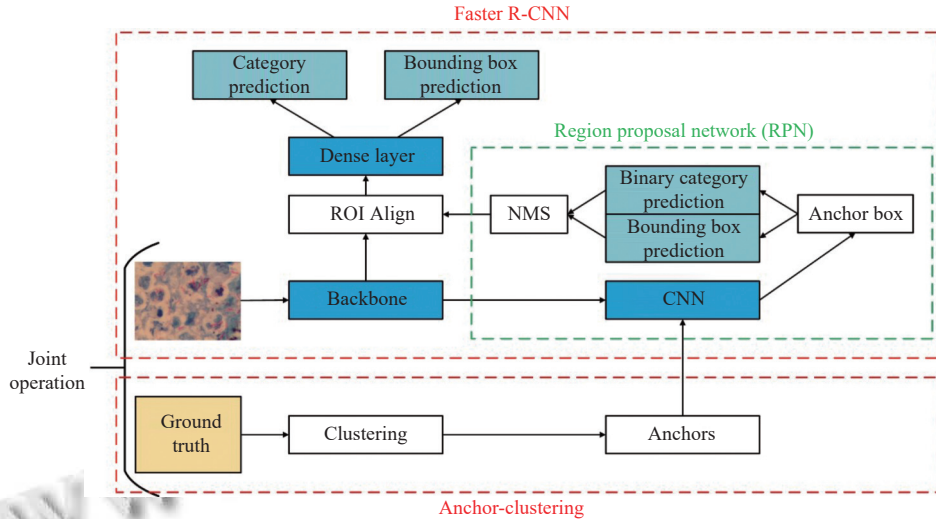


图1 Faster R-CNN 模型架构图

2.1 Mosaic 数据增强

Mosaic 数据增强^[9]算法是一种用于增强训练数据集的方法, 该算法的核心思想是将多张图片按照一定比例拼接组合成一张新图片, 可以得到大量的样本图像, 增加数据集的数量, 使得数据集更加丰富. 而且该方法拥有随机缩放的功能, 经过缩放后的样本图像中存在大量小目标图像, 使得模型对于小目标的检测效果更好, 这正是我们在检测肺结核病原体时所需要的. 并且在经过 Mosaic 数据增强后, 引入更多的随机性和多样性, 可以有效抑制模型对于训练数据的过拟合问题, 提高了模型的鲁棒性和泛华能力.

2.2 K-means 聚类算法

在 Faster R-CNN 模型中, 通过计算候选框与标记框的交并比 IoU (intersection over union), 来修正目标候选框与标记候选框之间的误差, 从而确定最终的目标检测区域, 如式 (1) 所示. 所以不同的 IoU 对于 Faster R-CNN 来说, 收敛速度与回归精度都会有所不同. 在原始的 Faster R-CNN 模型中, 初始的候选框尺寸是 128、256 和 512 像素, 其对应的高宽比分别为 0.5、1.0 和 2.0. 但对于肺结核病原体这样的小目标来说, 这样的数值显然是不利于我们检测的. 因此, 我们通过 K-均值 (K-means)^[10] 算法, 对我们的数据集进行重新聚类, 得到更适用于检测肺结核病原体的锚点初始候选框大小, 聚类结果如

图 2 所示, x 轴为标注的宽度, y 轴为标注的高度.

$$IoU = \frac{H \cap A}{H \cup A} \quad (1)$$

其中, IoU 为交并比, H 为候选框, A 为标记框

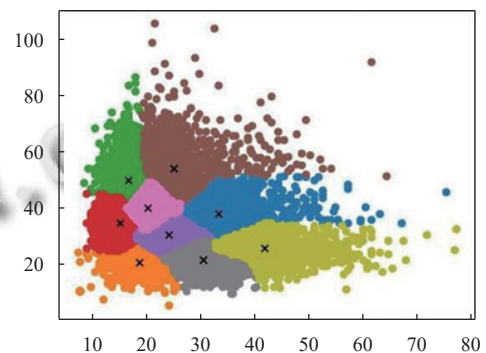


图2 K-means 聚类结果图

2.3 Res2Net-Atrous

由于肺结核病原体属于小目标, 特征有限, 原始的在 Faster R-CNN 模型对于肺结核病原体的特征提取效果有限, 容易出现错检或者漏检的情况, 因此, 本文将 Faster R-CNN 的特征提取模块替换为 Res2Net, 并将其中的卷积核全部替换为空洞卷积以达到更好的检测效果.

2.3.1 空洞卷积

空洞卷积 (atrous convolution)^[11]是一种在深度学

习领域中常用的卷积操作, 其传统的卷积操作相比具有一定的特殊性. 传统的卷积操作通过滑动卷积核在输入特征图上进行操作, 每个位置的输出仅由相邻位置的输入影响. 而空洞卷积则通过在卷积核中引入额外的间隔 (或称为“膨胀率”或“扩张率”), 使得卷积核在感受野内部具有间隔. 这意味着, 在保持感受野范围的同时, 空洞卷积可以利用更大的卷积核进行操作, 从而扩大了感受野的范围, 捕获了更广泛的上下文信息.

空洞卷积中的空洞率参数决定了卷积核内部的间隔大小, 它控制了感受野内部相邻位置之间的距离. 具体来说, 如果将空洞率设置为 1, 那么就是普通的卷积操作; 而将空洞率设置为 2, 表示在卷积核内部间隔 1 个像素进行操作, 空洞率为 3 则表示间隔 2 个像素, 以此类推. 通过调整空洞率, 可以在不增加参数数量的情况下增加感受野的范围, 从而捕获更多的上下文信息.

另一个空洞卷积的优点是它的计算效率. 由于空洞卷积的参数共享特性, 其计算量并不随着空洞率的增加而增加太多, 而且在与原卷积相比参数数量不变的情况下, 增大了感受野. 因此, 我们将 Faster R-CNN 特征提取中的卷积核全部替换为空洞卷积.

2.3.2 Res2Net

Res2Net^[12]是一种创新性的深度神经网络架构, 旨在克服传统深度学习模型在处理多尺度信息时的限制.

它的名称“Res2Net”是“residual relation network”的简称, 强调了其基于残差学习的机制以及对多尺度关系的建模能力.

传统的深度学习模型, 如 ResNet^[13], 虽然在处理图像任务中取得了巨大成功, 但仍存在着一些局限性. 其中一个主要问题是对于不同尺度的信息提取不足. 在现实世界的图像中, 物体的尺度会因其距离、角度、大小等因素而变化, 因此模型需要能够有效地处理这种多尺度信息.

Res2Net 通过引入并行的子网络结构, 使得网络能够同时学习不同尺度的特征表示. 这种并行结构不同于传统的堆叠结构, 而是将多个子网络置于同一个残差块内, 每个子网络负责学习不同尺度的特征信息. 这些子网络可以有不同感受野大小, 从而能够捕获不同尺度的语义信息.

在 Res2Net 中, 每个残差块被设计为具有多个分支, 每个分支都有自己的权重和特征映射, 然后这些分支的特征被逐元素相加, 形成最终的输出. 这种分支结构使得网络可以更有效地学习到多尺度的特征信息, 从而提高了网络的表达能力和泛化能力.

除了处理多尺度信息外, Res2Net 还具有计算效率高和易于训练的优点. 由于其并行的设计, Res2Net 相对于传统的堆叠结构具有更少的参数量和计算成本. 改进后的特征提取模块 Res2Net-Atrous 如图 3 所示.

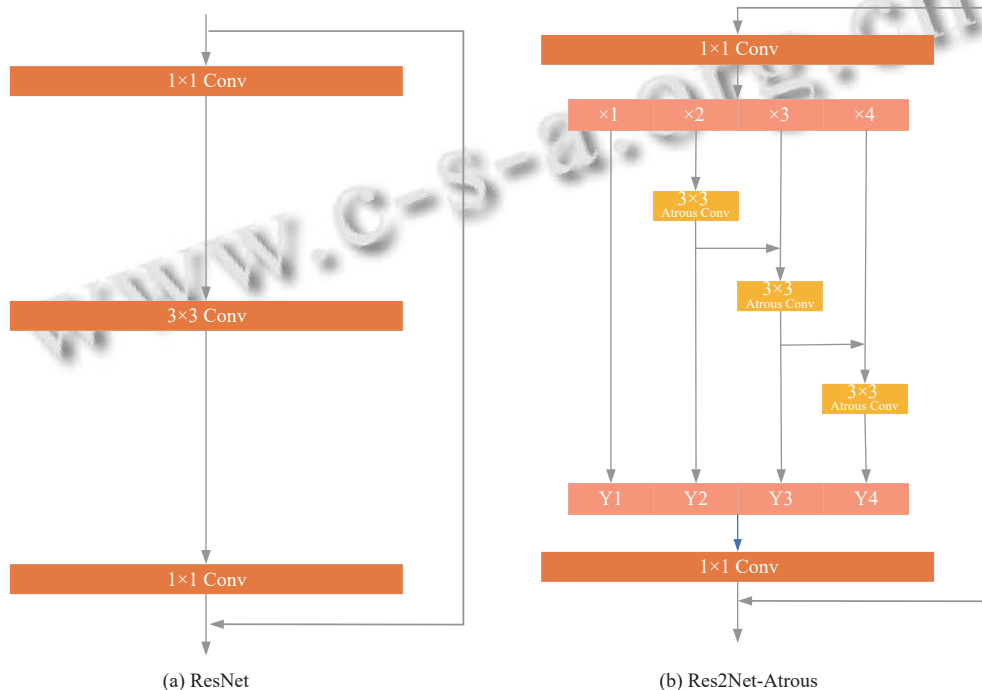


图 3 改进的特征提取模块

2.4 GC-FPN 模块

为了增强对肺结核病原体的特征提取效果,引入改进后的 GC-FPN 模块,在传统 FPN^[14]的基础上增加一条自底向上的路径增强,缩短了低层与顶层之间的路径,即路径聚合网络 (PANet),并引入幻影卷积 (GhostConv) 和 CA 注意力机制,对最后输出的特征图进行处理,使模型轻量化的同时更好的关注小目标信息,从而提高整个模型对于小目标的检测效果,GC-FPN 模块架构如图 4 所示。

2.4.1 PANet

PANet (pyramid attention network) 是一种用于目

标检测任务的深度学习架构,它在 2018 年由华中科技大学提出。PANet^[15]的核心创新在于引入了金字塔形式的注意力机制,以有效地利用多尺度的特征信息,从而提升目标检测的性能。传统的目标检测方法通常使用特征金字塔来处理不同尺度的目标,但在特征融合和信息传递方面存在一定的局限性。PANet 通过在特征金字塔的每个级别上应用自注意力模块,实现了跨尺度、跨层次的特征交互和整合,从而在保持特征分辨率的同时充分利用了不同尺度的特征信息。这种金字塔形式的自注意力机制使得网络能够更好地捕获目标的上下文信息和全局语义,从而提高了目标检测的准确性和鲁棒性。

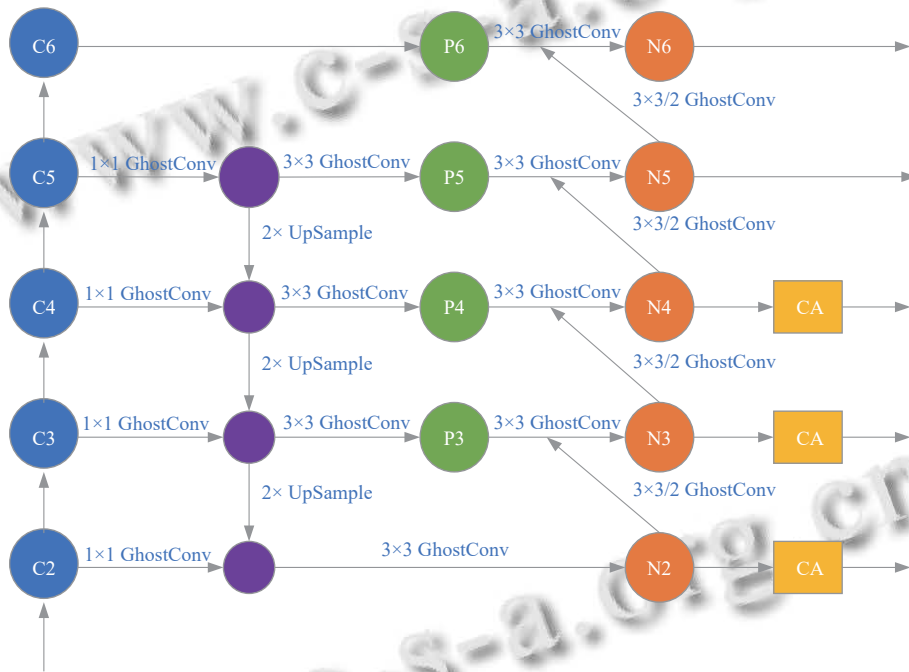


图 4 GC-FPN 模块架构图

具体而言, PANet 将特征金字塔分为自上而下的顶级和自下而上的底层,然后通过横向连接和自注意力机制在各个层级之间进行信息传递和特征融合。在横向连接过程中,高层特征被传递给低层特征,以提供更丰富的上下文信息;而自注意力机制则通过学习每个特征位置的重要性权重,有针对性地加强或减弱不同位置的特征响应,以进一步提升特征表达能力。通过这种精心设计的特征交互和注意力机制, PANet 能够更好地适应各种尺度和分辨率的目标,提高了对小目标和遮挡目标的检测能力,显著改善了目标检测的性能和鲁棒性。PANet 架构如图 5 所示。

2.4.2 幻影卷积

幻影卷积 (ghost convolution) 是一种用于深度学习中的卷积操作,其旨在提高模型的特征表达能力和泛化性能。幻影卷积^[16]的核心思想是在卷积核中引入额外的“幻影通道”,使得每个通道可以学习到主通道的信息的线性组合,从而增加了特征的多样性和丰富度。具体而言,幻影卷积通过将输入特征图分成主通道和幻影通道两部分,并对主通道进行普通卷积操作,同时对幻影通道进行线性变换得到新的特征图,然后将两部分特征图相加得到最终的输出。这种设计使得幻影卷积能够利用额外的信息来丰富特征表示,从而提

高了网络对于复杂数据的处理能力。

2.4.3 CA 注意力机制

在目标检测任务中, 加入注意力机制可以提高模型对关键目标特征的关注程度, 从而增强目标检测的准确性和鲁棒性. 通过动态调整不同特征的权重, 注意力机制可以有效地提取和利用图像中与目标相关的信息, 减少背景干扰, 提高目标定位和分类的性能, 为模型在复杂场景下的应用提供更强大的表征能力. 因此, 我们在 Faster R-CNN 的主干网络中加入 CA 注意力机制 (channel attention mechanism)^[17]. CA 注意力机制的核心思想是通过建模通道之间的关联性, 动态地调整不同通道的权重, 使得模型能够更加有效地捕捉通道之间的相关性, 从而提高了特征的表达能力. 具体而言, CA 注意力机制通过在特征图上的每个通道上引入全局池化操作, 然后通过全连接层和激活函数计算出每个通道的权重, 最后将这些权重乘以原始特征图得到加权后的特征表示. 这种设计使得 CA 注意力机制能够根据不同通道的重要性动态地调整特征的权重, 从而提高了网络的特征表达能力和泛化性能, CA 注意力机制结构如图 6 所示.

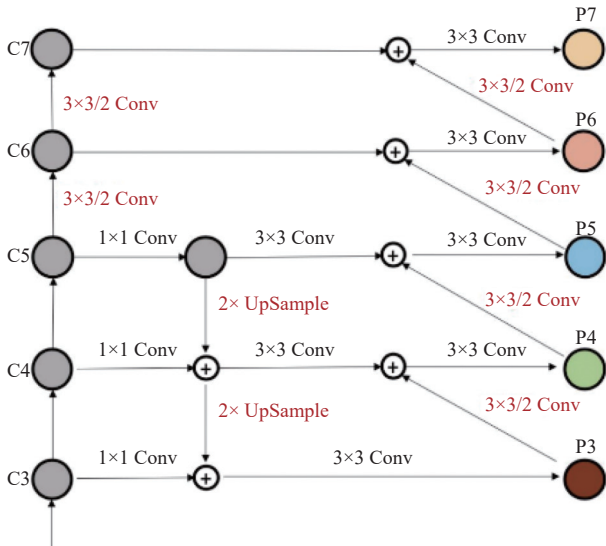


图 5 PANet 模块架构图

2.5 ROI Align

在原始的 Faster R-CNN 网络中, 采用 ROI pooling^[18] 从特征图中提取感兴趣区域对应的特征, 以便后续的目标分类和定位. 但 ROI pooling 存在信息损失和空间失真的问题, 因其使用量化操作和最近邻插值, 可能导致对 ROI 内部特征的不准确采样, 尤其在小目标或边

界细节的处理上表现不佳. 所以我们采用了 ROI Align, 其相比于 ROI pooling 的主要优势在于更精确地对感兴趣区域 (ROI) 内的特征图像素, 避免了信息丢失和量化误差, 从而提高了目标检测任务中对小目标和边界细节的处理能力, 进而提升了检测精度和定位准确度. ROI Align^[19] 的更加准确的插值方式使得特征图上的像素与原始图像上的像素一一对应, 有效地解决了 ROI pooling 中可能出现的空间信息损失问题, 使得模型能够更好地适应各种尺度和位置的目标, 提高了目标检测的性能和鲁棒性. 整体改进后的 Faster R-CNN 模型如图 7 所示.

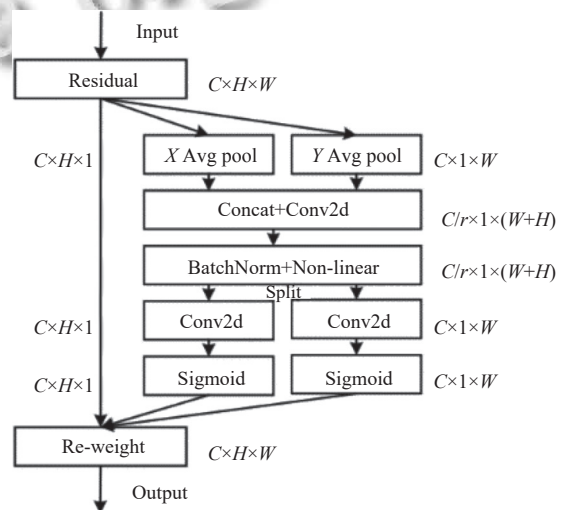


图 6 CA 注意力机制结构图

3 实验设置与结果分析

3.1 实验环境

本文所有的实验所采用的实验环境配置如表 1 所示. 在模型训练时所采用的训练参数如下: 初始学习率 0.001, 训练批次为 8, 训练轮次为 200 轮, 其余训练参数使用 Faster R-CNN 模型的默认设置.

3.2 数据集

本文所有实验采用的数据集为 Kaggle 上公开的一个结核病图像数据集, 来自马凯雷雷大学的自动化实验室诊断数据库 (Makerere Automated Lab diagnostics database), 取自于穆拉戈国家转诊医院 (Mulago National Referral Hospital), 该数据集设计了特殊的软件和硬件来从痰液中提取样本. 我们将数据集分为训练集、验证集和测试集, 具体比例如图 8 所示. 数据集以 PASCAL VOC 格式进行标注.

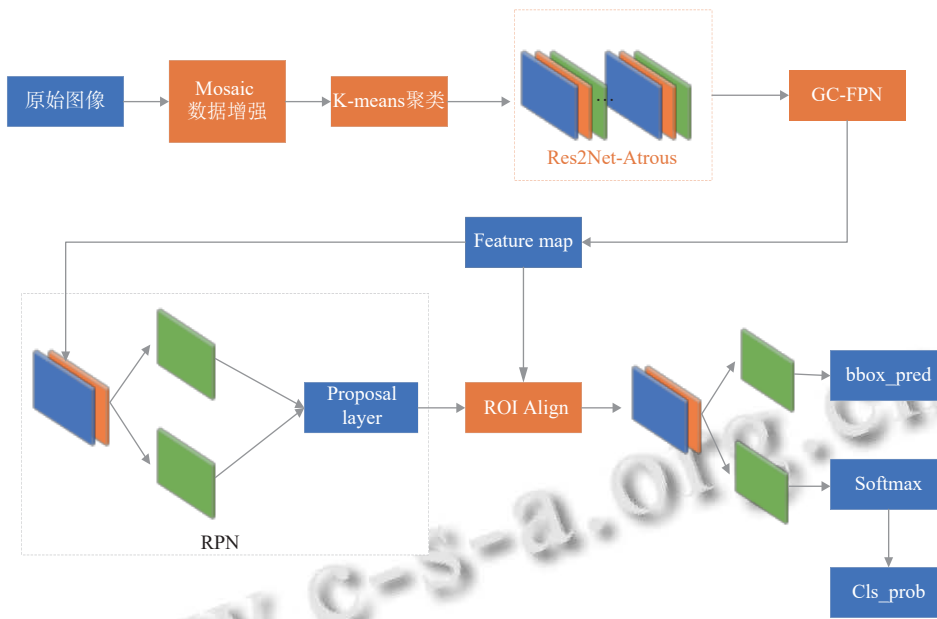


图7 改进后的 Faster R-CNN 模型架构图

表1 实验环境配置

配置	参数
操作系统	Windows 10 64 bit
GPU	NVIDIA GeForce RTX 3090
CPU	Intel(R) Core(TM) i9-10900X CPU@3.70 GHz
内存	64 GB
深度学习框架	PyTorch 1.8.1, CUDA 11.1

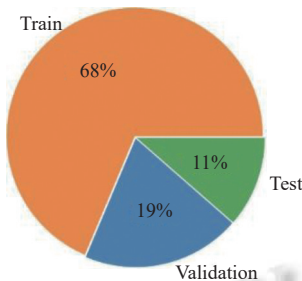


图8 数据集比例饼状图

3.3 评价指标

为了对模型的训练效果进行对比和量化,本次采用的评价指标有准确率 (Precision, P)、召回率 (Recall, R)、平均准确率 (AP) 置信度为 0.5 时平均精度均值 ($mAP_{0.5}$)。具体计算公式如下:

$$P = \frac{TP}{TP + FP} \quad (2)$$

$$R = \frac{TP}{TP + FN} \quad (3)$$

$$AP = \int_0^1 P(R) dR \quad (4)$$

$$mAP = \frac{1}{n} \sum_{i=1}^n AP_i \quad (5)$$

其中, TP 为识别正确的正样本数量, FP 为识别错误的正样本数量, FN 为识别错误的负样本数量, AP 表示某一类别样本的 PR (Precision-Recall) 曲线下的面积, $mAP_{0.5}$ 是 IoU (intersection over unit) 阈值等于 0.5 时所有类别样本的 AP 的均值,用于衡量模型对于所有类别样本的平均精准度。

3.4 损失函数

本文所提出的算法并未对 Faster R-CNN 的损失函数进行改进,故采用原始 Faster R-CNN 中的损失函数作为本文算法的损失函数,具体公式如式 (6) 所示。

$$L(\{p_i\}, \{t_i\}) = \frac{1}{N_{cls}} \sum_i L_{cls}(p_i, p_i^*) + \lambda \frac{1}{N_{reg}} \sum_i p_i^* L_{reg}(t_i, t_i^*) \quad (6)$$

其中, p_i 为 anchor 预测为目标的概率; p_i^* 如式 (7) 所示,用来代表是否有物体存在,存在则为 1,反之则为 0。

$$p_i^* = \begin{cases} 0, & \text{negative label} \\ 1, & \text{positive label} \end{cases} \quad (7)$$

其中, t_i 是一个向量,代表了预测时的 bounding box 的坐标向量,代表预测目标; t_i^* 是与 positive label 相对应的 bounding box 的 4 个坐标,代表真实目标; L_{cls} 是目标和非目标的对数损失,具体如式 (8) 所示。

$$L_{cls}(p_i, p_i^*) = -\log[p_i^* p_i + (1 - p_i^*)(1 - p_i)] \quad (8)$$

L_{reg} 是回归损失函数, 如式 (9) 所示. 在训练 RPN 的过程中, 通常会选择 256 个 anchor, 所以 N_{cls} 的值就是 256.

$$L_{reg}(t, t_i^*) = R(t_i - t_i^*) \quad (9)$$

为了验证本文所提出的改进 Faster R-CNN 模型在肺结核病原体检测上具有较好的收敛性和学习能力, 将改进后的 Faster R-CNN 模型与原始 Faster R-CNN 模型在同一数据集下进行训练, 得到损失函数曲线如图 9、图 10 所示. 可以看到, 随着训练轮数的增加, 改进后的 Faster R-CNN 模型与原始模型相比收敛速度更快, 损失波动较小, 对肺结核病原体的检测精度不断提高, 证明改进后的 Faster R-CNN 模型检测效果更好, 具有更好的鲁棒性.

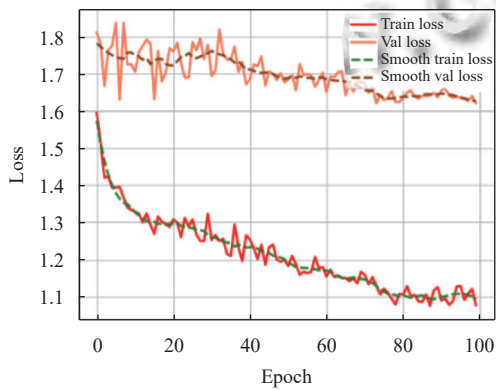


图 9 原始模型损失函数曲线图

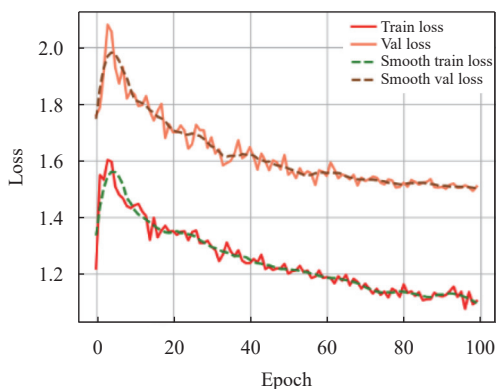


图 10 改进模型损失函数曲线图

3.5 消融实验

为了验证本文所提出的改进 Faster R-CNN 模型中所添加的每个模块都是有效果的, 本文以原始 Faster R-CNN 模型为基础, 逐步添加每一个改进的模块, 进行消融实验, 以观察其是否对肺结核病原体检测有提升效果, 实验结果如表 2 所示.

表 2 消融实验

方法	Mosaic	K-means	Res2Net-Atrous	GC-FPN	Precision	Recall	mAP0.5
原始	—	—	—	—	0.747	0.773	0.831
改进1	√	—	—	—	0.753	0.775	0.833
改进2	—	√	—	—	0.757	0.774	0.835
改进3	—	—	√	—	0.758	0.779	0.841
改进4	—	—	—	√	0.759	0.781	0.839
改进5	√	√	—	—	0.758	0.778	0.837
改进6	√	√	√	—	0.769	0.783	0.851
改进7	√	√	√	√	0.774	0.787	0.857

由表 2 中实验数据可以看出, Mosaic 数据增强和 K-means 聚类对数据集进行处理后, 对于肺结核病原体的识别准确度是有提升的. 而改进后的特征提取模块 Res2Net-Atrous 对于小目标特征的提取也有较好的帮助, GC-FPN 的加入在引入注意力机制后, 使得模型关注到了痰样本全局的信息, 扩大了感受野. 具体检测效果如图 11、图 12 所示, 可以看到, 改进后的模型提高了模型精度, 降低了误检率.

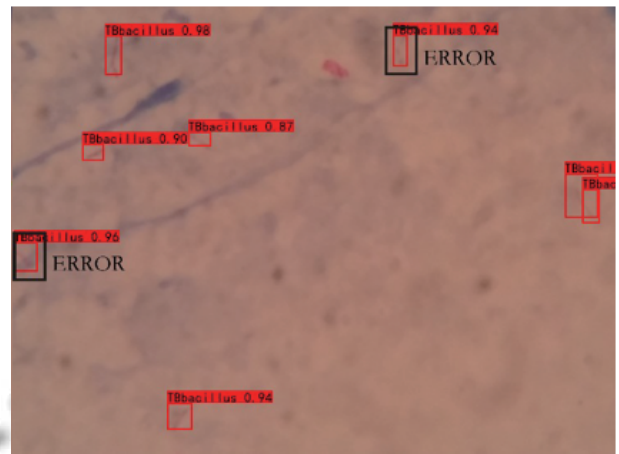


图 11 原始模型检测图



图 12 改进模型检测图

3.6 与其他目标检测算法的对比实验

为了对本文所提出的算法模型的有效性进行验证,将该模型与 SSD 和其他 YOLO 系列模型进行对比实验,实验结果如表 3 所示.通过实验结果我们可以看到,在相同的实验环境下,本文所提出的研究方法在精确度、召回率以及 mAP 这 3 个指标上均取得更好的结果,说明本文所提出的改进后的 Faster R-CNN 模型对于肺结核痰样本的检测是有提升的.具体检测效果如图 13 所示.

表 3 不同检测模型实验结果对比

模型	精确率	召回率	mAP	FPS
SSD	0.689	0.691	0.730	34.51
YOLOv3	0.761	0.732	0.805	54.27
YOLOv5	0.767	0.757	0.839	51.54
YOLOv7	0.753	0.768	0.823	41.73
YOLOv8	0.771	0.773	0.847	46.15
本实验方法	0.774	0.787	0.857	37.59



图 13 检测示例图

在检测效率方面, YOLO 系列模型本身就检测速度快而受到广泛应用,而 Faster R-CNN 模型以检测精度高而在目标检测领域受到青睐,在综合考虑后,我们认为在医学检测领域,应该以检测准确度作为首要条件,这也是在本课题选择 Faster R-CNN 模型来进行改进的一个重要原因.虽然本文所提出的模型的检测速度并不如 YOLO 系列模型,但 FPS 已超过 30,能够满足实时检测的需求,故说明本文所提出的模型是有实际应用价值的.

4 结论与展望

本文针对肺结核病原体,在 Faster R-CNN 模型上

进行了改进,首先引入了 Mosaic 数据增强和 K-means 聚类算法,丰富了训练数据集的同时也提高了模型对于目标的检测能力;加入改进后的特征提取模块 Res2Net-Atrous,保持了模型轻量化的同时也增大了感受野;最后加入改进后的特征金字塔模块 GC-FPN,引入 CA 注意力机制,抑制冗余信息,加强特征提取.通过实验结果表明,本文所提出的改进后的 Faster R-CNN 模型相比于原始模型, mAP 提升了 2.6%,相比于其他深度学习算法也具有一定优势.通过实验数据可知,本文所提出的检测方法在对肺结核病原体的检测上具有一定的实用价值.但模型的检测速度还有进一步提升的空间,可以尝试加入其他模块尝试效果.

参考文献

- 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. 精准医学杂志, 2020, 35(6): 481-484, 489.
- Jiao LC, Zhang F, Liu F, *et al.* A survey of deep learning-based object detection. IEEE Access, 2019, 7: 128837-128868. [doi: 10.1109/ACCESS.2019.2939201]
- 孙涵璇, 谢凤英, 姜志国, 等. 基于 BP 神经网络的结核杆菌目标识别. 中国体视学与图像分析, 2010, 15(1): 13-17.
- Nayak R, Shenoy VP, Galigeke RR. A new algorithm for automatic assessment of the degree of TB-infection using images of ZN-stained sputum smear. Proceedings of the 2010 International Conference on Systems in Medicine and Biology. Kharagpur: IEEE, 2010. 294-299.
- 卞景帅, 卢家品, 罗月童, 等. 基于 Faster R-CNN 的结核杆菌自动检测方法研究与应用. 图学学报, 2019, 40(3): 608-615.
- 崔少国, 解笋, 宋豪杰. 多卷积神经网络的胸部 X 光片肺结核检测方法. 计算机工程与应用, 2024, 60(13): 246-254.
- 王佳浩, 王宝珠, 郭志涛, 等. 多感受野的轻量化 YOLOv4 用于检测肺结核. 中国医学物理学杂志, 2022, 39(9): 1119-1127. [doi: 10.3969/j.issn.1005-202X.2022.09.011]
- Ren SQ, He KM, Girshick R, *et al.* Faster R-CNN: Towards real-time object detection with region proposal networks. Proceedings of the 28th International Conference on Neural Information Processing Systems. Montreal: MIT Press, 2015. 91-99.
- Wang H, Song ZL. Improved mosaic: Algorithms for more complex images. Journal of Physics: Conference Series, 2020, 1684: 012094. [doi: 10.1088/1742-6596/1684/1/012094]

- 10 Sinaga KP, Yang MS. Unsupervised K-means clustering algorithm. *IEEE Access*, 2020, 8: 80716–80727. [doi: [10.1109/ACCESS.2020.2988796](https://doi.org/10.1109/ACCESS.2020.2988796)]
- 11 Chen LC, Papandreou G, Schroff F, *et al.* Rethinking atrous convolution for semantic image segmentation. *arXiv:1706.05587*, 2017.
- 12 Gao SH, Cheng MM, Zhao K, *et al.* Res2Net: A new multi-scale backbone architecture. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2021, 43(2): 652–662. [doi: [10.1109/TPAMI.2019.2938758](https://doi.org/10.1109/TPAMI.2019.2938758)]
- 13 Targ S, Almeida D, Lyman K. ResNet in ResNet: Generalizing residual architectures. *arXiv:1603.08029*, 2016.
- 14 Ghiasi G, Lin TY, Le QV. NAS-FPN: Learning scalable feature pyramid architecture for object detection. *Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Long Beach: IEEE, 2019. 7029–7038.
- 15 Wang KX, Liew JH, Zou YT, *et al.* PANet: Few-shot image semantic segmentation with prototype alignment. *Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision*. Seoul: IEEE, 2019. 9196–9205.
- 16 Lu XY, Yang R, Zhou J, *et al.* A hybrid model of ghost-convolution enlightened Transformer for effective diagnosis of grape leaf disease and pest. *Journal of King Saud University —Computer and Information Sciences*, 2022, 34(5): 1755–1767. [doi: [10.1016/j.jksuci.2022.03.006](https://doi.org/10.1016/j.jksuci.2022.03.006)]
- 17 Huang YJ, Liao AH, Hu DY, *et al.* Multi-scale convolutional network with channel attention mechanism for rolling bearing fault diagnosis. *Measurement*, 2022, 203: 111935. [doi: [10.1016/j.measurement.2022.111935](https://doi.org/10.1016/j.measurement.2022.111935)]
- 18 Sun YX, Sun C, Wang D, *et al.* ROI pooled correlation filters for visual tracking. *Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Long Beach: IEEE, 2019. 5776–5784.
- 19 Bai T, Pang Y, Wang JC, *et al.* An optimized faster R-CNN method based on DRNet and ROI Align for building detection in remote sensing images. *Remote Sensing*, 2020, 12(5): 762. [doi: [10.3390/rs12050762](https://doi.org/10.3390/rs12050762)]

(校对责编: 张重毅)