

基于改进YOLOv5m的血管介入导丝检测算法

鲍芳嘉^{1,2}, 张明伟^{1,2}, 张天逸^{1,2}, 程云章^{1,2}

(1. 上海理工大学健康科学与工程学院; 2. 上海介入医疗器械工程技术研究中心, 上海 200093)

摘要: 针对目前主流目标检测算法在微创血管介入导丝检测中应用较少、检测准确率低和检测速度慢等问题, 提出了一种改进的YOLOv5m网络用于血管介入导丝检测。首先, 在YOLOv5m的主干网络中引入可变形卷积替换部分标准卷积, 并在主干网络的CSP模块中添加坐标注意力机制; 其次, 在颈部采用加权双向特征金字塔网络进行特征融合, 提高模型对不同特征层的融合能力。实验结果表明, 改进YOLOv5m算法的mAP@0.5达到87.8%, 比YOLOv5m提升了5.7%, 表明该算法在血管介入导丝检测方面具有较高应用价值。

关键词: 血管介入; YOLOv5m; 导丝检测; 可变形卷积; 坐标注意力机制; BiFPN

DOI: 10.11907/rjtk.232010

开放科学(资源服务)标识码(OSID):

中图分类号: TP391.41; R318 文献标识码: A

文章编号: 1672-7800(2024)010-0199-08



Vascular Intervention Guidewire Detection Algorithm Based on Improved YOLOv5m

BAO Fangjia^{1,2}, ZHANG Mingwei^{1,2}, ZHANG Tianyi^{1,2}, CHENG Yunzhang^{1,2}

(1. School of Health Science and Engineering, University of Shanghai for Science and Technology;

2. Shanghai Interventional Medical Device Engineering Technology Research Center, Shanghai 200093, China)

Abstract: In order to solve the problems of current mainstream target detection algorithms being less used in minimally invasive vascular interventional guidewire detection, low detection accuracy and slow detection speed, an improved YOLOv5m network is proposed for the detection of vascular interventional guidewires. First, deformable convolution is introduced into the backbone network of YOLOv5m to replace some standard convolutions, and a coordinate attention mechanism is added to the CSP module of the backbone network; BiFPN is used in the neck to perform feature fusion improving the model's ability to fuse different feature layers. Experimental results show that the mAP@0.5 of the improved YOLOv5m algorithm reaches 87.8%, which is 5.7% higher than YOLOv5m, indicating that this algorithm has relatively high application value in vascular interventional guidewire detection.

Key Words: vascular intervention; YOLOv5m; guidewire detection; deformable convolution; coordinate attention mechanism; BiFPN

0 引言

血管介入手术(Vascular Interventional Procedure)是一种通过导管和显影剂介入血管系统进行诊断和治疗的外科手术技术^[1]。该手术常用于治疗各种血管疾病,如冠心病、动脉狭窄、血栓形成和肿瘤等。血管介入手术通过在患者的血管内插入导管,经过导引将导丝等器械引导到目

标部位,然后进行疏通、栓塞、支架植入等治疗操作。相比传统的开放手术,血管介入手术具有微创性、恢复快、创伤小等优势,能够减少患者的痛苦和康复时间,同时提高手术的安全性和效果。

在血管介入手术中,准确地定位和引导导丝至治疗区域至关重要,不准确的导丝引导可能导致并发症或降低治疗效果^[2]。实时血管介入导丝检测技术在一定程度上能够确保导丝准确地到达治疗区域,医生可以更好地执行手

收稿日期: 2023-11-01

扫描二维码阅读全文:



作者简介: 鲍芳嘉(1998-),男,上海理工大学健康科学与工程学院硕士研究生,研究方向为血管介入机器人、计算机视觉、医学图像;张明伟(1997-),男,上海理工大学健康科学与工程学院博士研究生,研究方向为血管介入机器人、深度学习;张天逸(1983-),女,上海理工大学健康科学与工程学院博士研究生,研究方向为图像处理、深度学习;程云章(1964-),男,博士,上海理工大学健康科学与工程学院教授、博士生导师,研究方向为生物医学工程、流体机械与流体动力工程、血流动力学及其临床应用(FFR-CT技术)、精准医疗科学与工程(基因检测技术)、医疗器械大数据应用。本文通讯作者:程云章。

术,减少并发症风险,这对于处理复杂的心血管疾病和病变情况尤其重要。该技术研究有助于医疗领域不断改进导丝引导方法,开发更精确和更安全的治疗方式,并为未来的血管介入手术提供创新方法。同时,可改善心血管介入手术效果,减少潜在风险,更好地满足患者需求。总而言之,该技术可有效推动医疗技术发展,具有显著的医学和临床应用意义。

1 相关研究

随着深度学习和GPU计算能力的快速发展,目标检测技术已广泛应用于众多领域,包括智慧农业、工业检测、医疗、自动驾驶等。在医疗领域,将目标检测技术应用于血管介入导丝检测中,可以提升血管介入导丝检测准确率和速度,从而提高血管介入手术安全性和成功率^[3]。目前,许多学者正在研究计算机辅助血管介入导丝的实时跟踪检测,Zhu等^[4]提出一种混合多阶段算法用于导丝跟踪和分割,该算法包括3个关键阶段:预测(用于检测所有类似导丝的特征)、锁定(用于消除非导丝特征)、跟踪(用于在锁定阶段之后跟踪识别真实导丝),该算法在多个机构脑血管介入数据集上表现出了卓越效果。Zhou等^[5]提出快速递归残差网络(FRR-Net)和金字塔注意力递归网络(PAR-Net)^[6],用于实时导丝分割和跟踪,该网络包含注意力模块、递归残差模块和预训练MobileNetV2编码器,并提出了增强焦损和骰子损失的混合损失函数,以更好地解决类别不平衡和误分类样本的问题。Li等^[7]提出的两阶段框架采用YOLOv3在第一阶段进行导丝检测,然后经过后处理算法精细化检测框。在第二阶段,引入了一种监督注意力模块的SA-Hourglass Network,用于预测每个检测框中可能包含的导丝端点的密集热图。

针对血管介入导丝检测面临复杂的血管结构、低对比度、噪声干扰、尺寸和形状变化及实时性要求等问题,本文方法采用YOLOv5m作为网络基础框架,在网络的Backbone部分引入了DCNv2替代部分标准卷积,从而增强了网络对特征的提取能力,聚焦全局重要信息^[8];在Backbone的C3模块中引入了坐标注意力机制(Coordinate Attention, CA)^[9]。这有助于提高模型对重要信息的关注度,进一步提升了模型检测性能;在Neck部分引入了BiFPN,用于对不同尺度的特征图进行加权融合,与以往的特征融合方法相比,这种新的融合方式更加简单和高效^[10]。通过这些改进和创新,该方法有效地应对血管介入导丝检测中的挑战,提高了检测准确性和实时性。

2 YOLOV5算法

2.1 目标检测算法

当前主流的目标检测算法主要分为One-Stage和

Two-Stage两种类型。One-Stage算法将整个图像划分为网格,每个网格负责预测目标的类别和位置,无需生成候选区域。代表性算法包括YOLO系列^[11-12]、SSD^[13]、RetinaNet^[14]、FCOS^[15]和EfficientDet^[10]。Two-Stage算法通过Region Proposal Network(RPN)等方法生成候选框,然后利用卷积神经网络提取相关特征,用于目标分类和定位,包括R-CNN系列(R-CNN^[16]、Fast R-CNN^[17]、Faster R-CNN^[18]、Mask R-CNN^[19]等)、TridentNet^[20]和Cascade R-CNN^[21]。One-Stage算法适合于实时应用和对速度要求较高的场景,而Two-Stage算法适应于准确性要求较高的任务。

2.2 YOLOv5m算法

YOLOv5算法通过控制深度参数和宽度参数形成4种网络结构,分别为YOLOv5s、YOLOv5m、YOLOv5l、YOLOv5x。本文将YOLOv5m作为改进的基准模型,其深度参数和宽度参数分别为0.67和0.75。YOLOv5m算法由输入、主干网络(Backbone)、颈部(Neck)和预测输出(Head)组成,其框架如图1所示。

输入端主要进行数据处理增强,包括Mosaic数据增强、自适应锚框计算、自适应图片缩放等。Backbone对图像进行特征提取,主要包括CBS模块(卷积模块、批量归一化模块和激活函数SiLU)、CSPLayer(C3)模块和快速空间金字塔池(SPPF)模块。C3模块包括CSP1_X和CSP2_X,它们的区别在于CSP1_X中存在残差模块,可以增加层间反向传播的梯度值,避免网络加深带来的梯度损失,从而可以提取更细粒度的特征,主要应用于主干网络,而CSP2_X主要应用于颈部网络。Neck结合了FPN(Feature Pyramid Network)和PAN(Perceptual Adversarial Network)的特征金字塔融合技术,从而更有效地合并从骨干网络提取的多尺度特征,网络最终输出80×80、40×40、20×20尺寸的特征图,然后将其传输到检测层分别检测小、中、大目标,经过非极大值抑制等后处理操作,消除了大量冗余预测帧。Head输出置信度得分最高的预测类别,并返回目标位置的坐标。

3 YOLOv5m算法优化

3.1 坐标注意力机制

不同于传统的注意力机制,CA主要关注输入数据的空间坐标信息,而不是特征通道之间的关系。CA的核心思想是在输入数据的不同位置应用不同的权重,以突出空间上的重要信息,有助于模型更好地理解输入数据的空间结构。引入CA注意力机制可有效缓解网络深层特征分辨率较低引起的导丝模糊及信息丢失问题,使网络更加关注导丝区域,提高导丝区域的权重,提高对细长目标或复杂场景的处理性能。CA坐标注意力机制如图2所示。

CA输入特征图的通道和大小为 $[C, H, W]$,使用两个

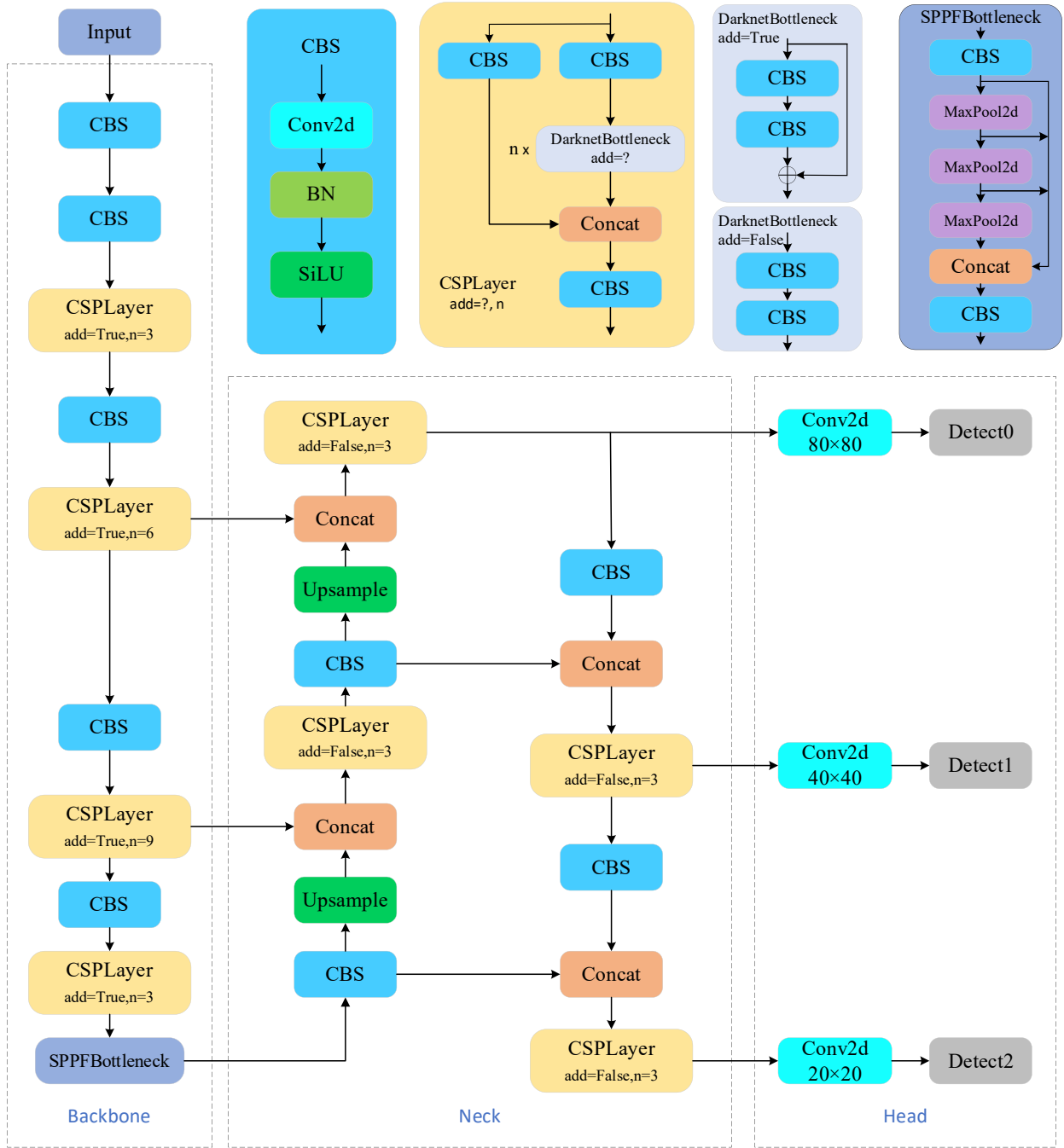


Fig. 1 YOLOv5m model structure

图1 YOLOv5m模型结构

空间范围的池化内核(H,1)或(1,W)分别沿水平X坐标和垂直Y坐标对每个通道进行一维特征编码。第c通道在高度h和宽度w的输出如式(1)、式(2)所示。其中, x_c 为输入特征, $Z_c^h(h)$ 为高度为h的第c通道的输出,表示对高度为h的输入进行平均池化。 $Z_c^w(w)$ 为宽度为w的第c通道的输出,表示对宽度为w的输入进行平均池化。

$$Z_c^h(h) = \frac{1}{W} \sum_{0 \leq i < W} x_c(h, i) \quad (1)$$

$$Z_c^w(w) = \frac{1}{H} \sum_{0 \leq j < H} x_c(j, w) \quad (2)$$

上述两个变换分别沿着两个空间方向聚合特征,产生

一对方向感知的特征图,然后将平均池化后两个X、Y方向的特征图进行拼接,再进行1x1卷积、批归一化和非线性激活操作,最后将拼接的特征图分离编码成两个注意力图。对于输入特征图 Z_c^h 和 Z_c^w ,输出如式(3)、式(4)所示。其中, δ 为非线性激活函数, σ 为Sigmoid激活函数, C_h 和 C_w 表示垂直方向和水平方向上的1x1卷积, g^h 和 g^w 表示在垂直方向和水平方向上输出的注意力特征图。对输出 g^h 和 g^w 进行扩展,分别作为注意力权重,得到坐标注意力机制的输出,如式(5)所示。

$$g^h = \sigma(C_h(\delta(Z_c^h))) \quad (3)$$

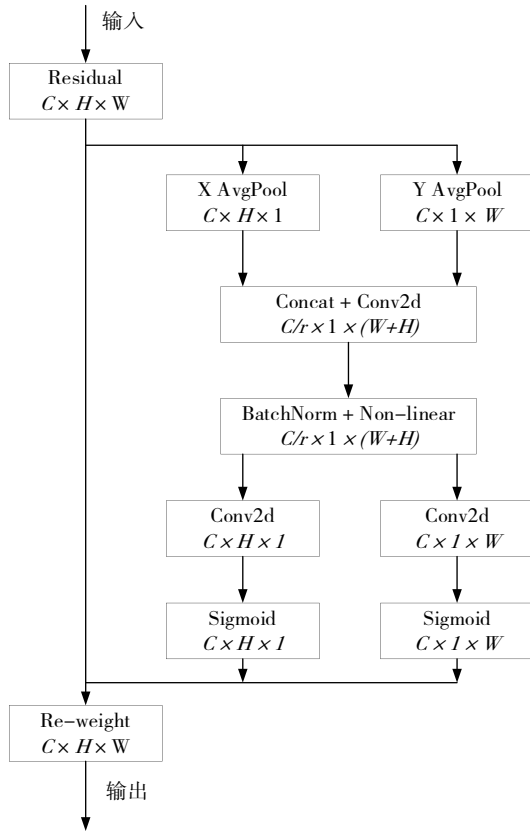


Fig. 2 CoordAtt coordinate attention mechanism structure
图2 CoordAtt坐标注意力机制结构

$$g^w = \sigma(C_w(\delta(Z_c^w))) \quad (4)$$

$$y_c(i, j) = x_c(i, j) \times g_c^h(i) \times g_c^w(j) \quad (5)$$

3.2 可变形卷积DCNv2

可变形卷积v2(Deformable Convolution v2)在标准卷积的基础上加入可学习偏移量和调制标量,使卷积核针对目标区域的对象进行动态调整,不会受到固定卷积核大小限制,扩大了感受野,能自适应地提取目标物体的不规则形状变化,从而提取细长导丝的更多特征信息。可变形卷积v2如式(1)所示,给定 3×3 个采样位置的卷积核,依次在 K 为9的网格位置采样。 $x(p)$ 和 $y(p)$ 分别表示位置 p 经输入特征映射 x 和输出特征映射 y 后的特征, w_k 和 p_k 分别表示第 k 个位置的权重和预先指定的偏移量, $p_k \in \{(-1, -1), (-1, 0), (-1, 1) \dots (1, 1)\}$, Δp_k 和 Δm_k 分别表示第 k 个位置的可学习偏移量和调制标量。可变形卷积如图3所示,通过在相同的输入特征图上应用两个单独的卷积层获得通道数为 $2K$ 的偏移量特征图和通道数为 K 的调制量特征图,输出可学习偏移量和权重调制标量。可学习偏移量是范围不受约束的实数,初始值设为0;调制标量 Δm_k 取值在 $[0, 1]$ 区间,初始值设为0.5。

$$y(p) = \sum_{k=1}^K w_k \cdot x(p_0 + p_k + \Delta p_k) \cdot \Delta m_k \quad (6)$$

3.3 特征融合模块改进

YOLOv5m算法在颈部采用路径聚合网络PAN对特征

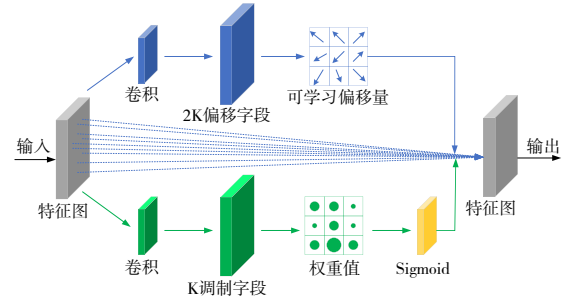


Fig. 3 3x3 DCNv2 module
图3 3x3可变形卷积模块

图先进行自顶向下的特征融合,再进行自底向上的特征融合,将深层特征和浅层特征相互融合,使预测层能够获取深浅层的整体信息。这种特征融合方式存在一些问题:①PAN将来自不同特征图的信息按同等重要性融合在一起,忽略了不同分辨率特征图的重要性;②随着网络下采样所导致的大量特征信息丢失,PAN缺乏对浅层特征信息的再利用;③PAN是基于金字塔结构,对输入分辨率和目标大小比较敏感,不够灵活。当输入分辨率和目标大小发生变化时,PAN需要重新训练,从而降低了模型适用性。因此,本文引入BiFPN对YOLOv5m进行改进。

BiFPN对不同分辨率的特征图进行多层次融合,如图4所示。通过自适应的权重调整机制,自动学习到每个特征图对最终检测结果的重要性,从而更加精确地融合特征图,增强模型对特征细节的感知能力,从而提高检测算法的准确率。

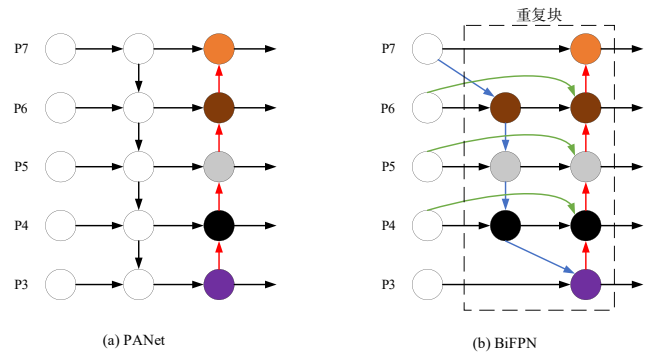


Fig. 4 Comparison of PANet and BiFPN structures
图4 PANet与BiFPN结构比较

3.4 改进后的YOLOv5m网络模型

本文以YOLOv5m为基础模型进行改进,其算法结构如图5所示(彩图扫OSID码可见)。其中,橙色、紫色、红色模块为本文添加的模块,红色线条为新增的网络传递方向。首先,将特征提取网络中的第2、3、4个卷积模块替换成DCNv2模块,并在CSP模块的残差结构和SPPF模块前添加CA注意力模块,增强网络特征提取能力,为后续算法网络提供高质量的特征图;其次,将Neck中的PANet替换成BiFPN模块,降低模型参数量,增强模型的多尺度特征融合能力,提高检测性能。

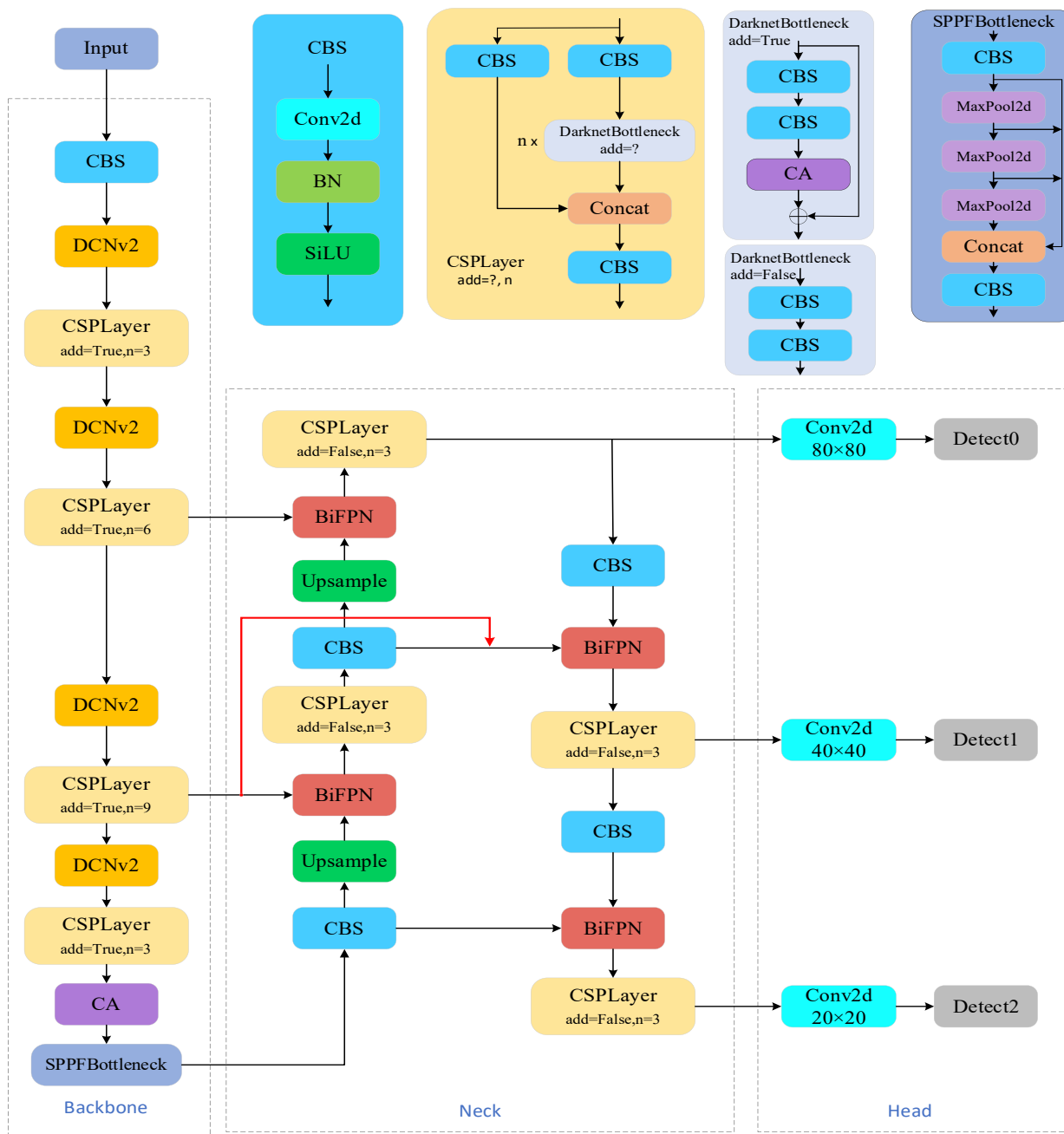


Fig. 5 Improved YOLOv5m network structure

图 5 改进的 YOLOv5m 网络结构

4 实验与结果分析

4.1 实验环境及参数

本次实验的环境配置如表 1 所示。

Table 1 Experimental environment configuration

表 1 实验环境配置

| 名称 | 环境配置 |
|--------|---|
| 操作系统 | Ubuntu20.04 |
| GPU | RTX3060 |
| CPU | Intel(R) Core(TM) i9-10920X CPU @ 3.50GHz |
| 集成开发环境 | Anaconda, CUDA11.6 |
| 深度学习框架 | PyTorch 1.13.1, Python 3.9 |

模型训练过程中, 设置随机梯度下降(SGD)函数优化参数, 初始学习率为 0.01, 动量为 0.937, 权重衰减为 0.0005, epoch 为 200 次, WarmUp 为 5, Batch size=8, 输入图片大小为 640×640, 并采用余弦退火的学习率衰减方式。

4.2 损失函数

损失函数可以衡量网络输出预测结果和真实结果匹配程度。本文损失函数由置信度损失 L_{obj} 、分类概率损失 L_{cls} 和边界框损失 L_{box} 组成, 其中边界框损失函数采用 C-IoU 损失函数, 该函数添加了两框间的中心距离和长宽比损失, 进一步提升了目标框回归的准确性, C-IoU 公式如式 (7) 所示。

$$\begin{cases} Loss_{IoU} = 1 - IoU + \frac{\rho^2(b, b^{gt})}{c^2} + \alpha v \\ v = \frac{4}{\pi^2} (\arctan \frac{w^{gt}}{h^{gt}} - \arctan \frac{w}{h})^2 \\ \alpha = \frac{v}{(1 - IoU) + v} \end{cases} \quad (7)$$

其中, b 和 b^{gt} 分别为预测框和标注框的中心点, $\rho(\cdot)$ 为两中心点的欧式距离, c 为预测框和标注框最小包围外矩形的对角线距离, w 和 h 为预测框的宽和高, w^{gt} 和 h^{gt} 为标注框的宽和高。

4.3 模型性能评价指标

本文实验采用平均精确度均值 (mean Average Precision, mAP)、参数量 (Parameters)、浮点运算数 (Giga Floating-point Operations Per Second, GFLOPs) 和帧率 (FPS) 分析评价模型性能。准确度 (P) 是测试集中正确识别目标的百分比, 该百分比由真阳性 (TP) 和假阳性 (FP) 定义, TP 表示正样本被预测为正样本的个数, FP 表示正负样本被预测为正样本的个数。召回率 (R) 是测试集中所有正样本被正确识别的概率, 由真阳性 (TP) 和假阴性 (FN) 定义, FN 表示负样本被预测为负样本的个数。 AP 表示不同召回率下精确率的平均值。 mAP 表示不同类别的 AP 值的平均值, N 表示总类别数。其公式如式(8)一式(11)所示。

$$P = \frac{TP}{TP + FP} \quad (8)$$

$$R = \frac{TP}{TP + FN} \quad (9)$$

$$AP = \int_0^1 P(R) dR \quad (10)$$

$$mAP = \frac{1}{N} \sum_{i=1}^N AP_i \quad (11)$$

4.4 数据集与预处理

本文使用的数据集由上海长海医院提供, 包括5位病人的微创血管介入手术影像视频数据。使用Python脚本程序将手术视频格式转换成为JPG图片, 经裁剪、矫正操作等操作后分辨率大小为640×640, 使用Labelimg软件标注导丝 (Guidewire) 和导管鞘管 (Conduit) 的位置, 标注结果由上海长海医院专业医生进行校验再标注, 以确保准确性和可靠性。表2展示了增强后的数据集分布情况, 图6展示了数据集中的部分图片数据。

Table 2 Distribution of the dataset

表2 数据集分布表

| Dataset | Image type | | | Total |
|---------|------------|---------|-------------------|--------|
| | Guidewire | Conduit | Guidewire&Conduit | |
| Train | 6 740 | 270 | 3 430 | 10 440 |
| Val | 560 | 130 | 320 | 1 010 |
| Test | 740 | 130 | 130 | 1 000 |
| Total | 8 040 | 530 | 3 880 | 12 450 |

4.5 实验结果分析

4.5.1 不同模型对比实验分析

将本文改进的YOLOv5算法与其他算法进行比较, 以



Fig. 6 Vascular interventional guidewire raw data

图6 血管介入导丝原始数据

相同比例划分数据集为基准, 采用现有的主流目标检测模型 Faster-RCNN、SSD、RetinaNet、YOLOv3、YOLOX 和 YOLOv5 在介入导丝数据集上开展对比实验, 结果如表3所示。实验结果表明, 本文改进的YOLOv5算法相较于其他检测算法在mAP@0.5指标上都有较大提升, 并且在模型计算量和参数量上也有明显降低。

Table 3 Results of comparative experiment

表3 对比实验结果

| Model | Backbone | GFLOPs | Params/M | mAP@0.5/% |
|-------------------------|---------------------------|--------|----------|-----------|
| SSD | VGG16 | 87.6 | 24.53 | 75.8 |
| Faster-RCNN | Resnet50+FPN | 134.5 | 41.36 | 79.2 |
| RetinaNet | Resnet101 | 122.5 | 55.36 | 77.6 |
| YOLOv3 ^[22] | Darknet53 | 155.3 | 61.5 | 80.8 |
| YOLOv5-m | CSPDarkNet53 | 48.2 | 20.8 | 82.1 |
| YOLOX-m ^[23] | CSPDarkNet53 | 75.7 | 25.28 | 82.5 |
| YOLOv5m-DCB | CSPDarkNet53 +DCNv2+CA | 40.5 | 15.2 | 87.8 |

4.5.2 不同模块消融实验分析

为验证DCNv2、坐标注意力机制和BiFPN对模型改进的有效性, 以YOLOv5m算法为基准, 将改进分为A、B、C、D共4组进行消融实验, 结果如表4所示。

Table 4 Results of ablation experiment

表4 消融实验结果

| Model | Improved Module | | | GFLOPs | Params/ M | mAP @ 0.5/% | FPS (frame/ s) |
|-------------|-----------------|----|-------|--------|--------------|-------------------|----------------------|
| | DCNv2 | CA | BiFPN | | | | |
| YOLOv5m | / | / | / | 48.2 | 20.8 | 82.1 | 91.7 |
| A | √ | / | / | 40.9 | 21.0 | 82.9 | 78.5 |
| B | / | √ | / | 48.4 | 20.9 | 82.3 | 75.4 |
| C | / | / | √ | 26.4 | 17.2 | 84.2 | 85.6 |
| D | √ | √ | / | 41.3 | 21.4 | 85.1 | 67.1 |
| YOLOv5m-DCB | √ | √ | √ | 40.5 | 15.2 | 87.8 | 62.6 |

模型A表示加入DCNv2模块的YOLOv5m模型, 模型B表示加入CA坐标注意力模块的YOLOv5m模型, 模型C表示加入BiFPN结构的YOLOv5m模型, “√”表示加入相应模块。通过消融实验, 本文改进后的YOLOv5m-DCB算法相较于原始YOLOv5-m算法在导丝数据集上mAP@0.5值提升5.7%, 并且模型参数量减少了27%, 计算量减少了16%。综合模型参数量、模型计算量和mAP@0.5等实验数据, 改进的YOLOv5m-DCB表现出更好的性能。

改进前YOLOv5m和改进后YOLOv5m-DCB的P-R曲

线分别如图 7、图 8 所示,改进前 YOLOv5m 算法对导丝 (Guidewire) 和导管鞘管 (Conduit) 预测的 AP 值分别为 90.2% 和 74.0%,两类检测的 mAP@0.5 为 82.1%。改进后的 YOLOv5m-DCB 算法对导丝 (Guidewire) 和导管鞘管 (Conduit) 预测的 AP 值分别为 94.7% 和 80.9%,分别提升了 4.5% 和 6.9%,两类检测的 mAP@0.5 为 87.8%,提升了 5.7%。

实验结果表明,引入 DCNv2、CA 坐标注意力机制和 BiFPN 对 YOLOv5m 算法作出改进,在血管介入导丝数据集上提升了导丝和导管鞘管的检测精度。

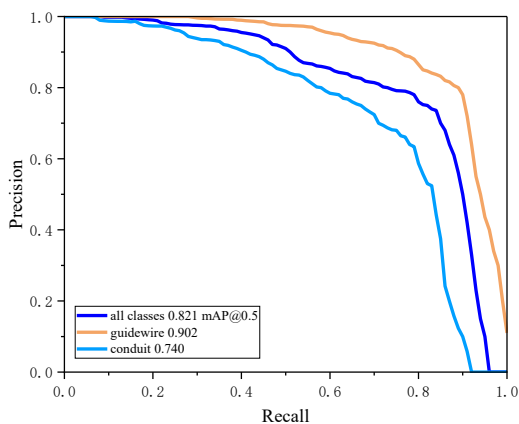


Fig. 7 YOLOv5m P-R curve
图 7 YOLOv5m P-R 曲线

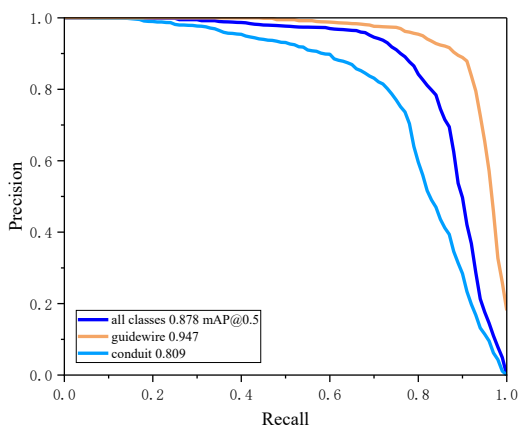


Fig. 8 YOLOv5m-DCB P-R curve
图 8 YOLOv5m-DCB P-R 曲线

4.6 算法检测结果可视化

为验证 YOLOv5m-DCB 模型的检测效果,使用 YOLOv5m 算法和改进后的 YOLOv5m-DCB 算法在介入导丝测试集上进行了测试。图 9 为原始数据标注,图 10 为 YOLOv5m 算法检测效果,图 11 为改进算法 YOLOv5m-DCB 的检测效果。

从图示结果看,YOLOv5m 模型对于导丝导管目标存

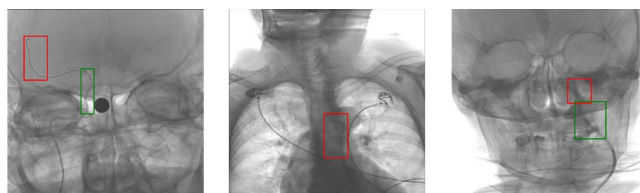


Fig. 9 Raw data labeling
图 9 原始数据标注

在漏检、错检情况,YOLOv5m-DCB 模型检测性能更好,检测效果有明显提升。



Fig. 10 YOLOv5m detection effect
图 10 YOLOv5m 检测效果



Fig. 11 YOLOv5m-DCB detection effect
图 11 YOLOv5m-DCB 检测效果

5 结语

本文在血管介入导丝检测的医学应用背景下,针对导丝检测中漏检、错检、检测精度低等问题,提出了一种血管介入导丝检测算法 YOLOv5m-DCB。首先,在 YOLOv5m 主干网络中融入 DCNv2 和 CA 坐标注意力机制,能够更加关注导丝的特征信息,获取更多的细节信息;其次,在 Neck 网络采用 BiFPN 结构进行多尺度特征融合,提高网络特征融合能力。

实验结果表明,与现有检测模型 Faster R-CNN、SSD、RetinaNet、YOLOv5m、YOLOX 相比,改进后的 YOLOv5m 算法检测精度达到了 87.8%,提升了 5.7%,FPS 达到 62.6,在一定程度上满足了检测精确度和检测速度要求。但该工作仍有很大改进空间,比如轻量化网络结构,以及如何在保证检测精度的基础上降低计算量和参数量,以更好地应用于医学场景,这都有待进一步研究。

参考文献:

[1] GUO T Z. Interventional therapy of cerebrovascular diseases [M]. Kunming: Yunnan Science and Technology Publishers, 2020.
郭铁柱. 脑血管疾病介入治疗学[M]. 昆明: 云南科技出版社, 2020.

[2] ZHOU G J, LIANG S C, LI Z F, et al. Instrument tracking methods of vascular intervention based on image recognition[J]. Beijing Biomedical Engineering, 2022, 41(1): 90-96.

- 周国敬,梁世超,李振锋,等. 基于图像识别的血管介入器械追踪方法[J]. 北京生物医学工程,2022,41(1):90-96.
- [3] LI N, HE J A, CHEN Y, et al. A survey on the progress of computer-assisted vascular intervention[J]. Journal of Computer-Aided Design & Computer Graphics, 2022, 34(7):985-1010.
- 李娜,贺建安,陈阳,等. 计算机辅助血管介入技术进展综述[J]. 计算机辅助设计与图形学学报,2022, 34(7): 985-1010.
- [4] ZHU J, WANG C, ZHANG G, et al. A multi-stage guidewire tip tracking framework for cardiovascular robotic interventions[C]//2021 14th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics, 2021: 1-6.
- [5] ZHOU Y J, XIE X L, HOU Z G, et al. FRR-net: fast recurrent residual networks for real-time catheter segmentation and tracking in endovascular aneurysm repair[C]//2020 IEEE 17th International Symposium on Biomedical Imaging, 2020: 961-964.
- [6] ZHOU Y J, XIE X L, ZHOU X H, et al. Pyramid attention recurrent networks for real-time guidewire segmentation and tracking in intraoperative X-ray fluoroscopy[J]. Computerized Medical Imaging and Graphics, 2020, 83: 101734.
- [7] LI R Q, BIAN G, ZHOU X, et al. A two-stage framework for real-time guidewire endpoint localization[C]//Shenzhen: 22nd International Conference on Medical Image Computing and Computer Assisted Intervention, 2019.
- [8] ZHU X, HU H, LIN S, et al. Deformable convnets v2: more deformable, better results[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019:9308-9316.
- [9] HOU Q, ZHOU D, FENG J. Coordinate attention for efficient mobile network design[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021:13713-13722.
- [10] TAN M, PANG R, LE Q V. Efficientdet: scalable and efficient object detection[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020:10781-10790.
- [11] JIANG P, ERGU D, LIU F, et al. A review of yolo algorithm developments [J]. Procedia Computer Science, 2022, 199: 1066-1073.
- [12] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. Yolov4: optimal speed and accuracy of object detection [DB/OL]. <https://arxiv.org/abs/2004.10934>, 2020.
- [13] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multibox detector[C]//Proceedings of 14th European Conference on Computer Vision, 2016:21-37.
- [14] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[C]//Proceedings of the IEEE International Conference on Computer Vision, 2017:2980-2988.
- [15] TIAN Z, SHEN C, CHEN H, et al. FCOS: fully convolutional one-stage object detection[C]//Seoul: 2019 IEEE/CVF International Conference on Computer Vision, 2019.
- [16] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014:580-587.
- [17] GIRSHICK R. Fast R-CNN[C]//Proceedings of the IEEE International Conference on Computer Vision, 2015: 1440-1448.
- [18] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017, 39(6):1137-1149.
- [19] HE K, GKIOXARI G, DOLLÁR P, et al. Mask R-CNN[C]//Proceedings of the IEEE International Conference on Computer Vision, 2017: 2961-2969.
- [20] LI Y, CHEN Y, WANG N, et al. Scale-aware trident networks for object detection[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019:6054-6063.
- [21] CAI Z, VASCONCELOS N. Cascade R-CNN: high quality object detection and instance segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2019, 43(5): 1483-1498.
- [22] REDMON J, FARHADI A. Yolov3: an incremental improvement [DB/OL]. <https://arxiv.org/abs/1804.02767>, 2018.
- [23] GE Z, LIU S, WANG F, et al. Yolox: exceeding yolo series in 2021 [DB/OL]. <https://arxiv.org/abs/2107.08430>, 2021.

(责任编辑:孙娟)