

基于目标检测的轻量化乳腺肿瘤检测算法

张教锋, 李 东, 刘晓静

(青海大学 计算机技术与应用学院, 青海 西宁 810016)

摘 要: 随着乳腺癌发病率不断上升, 乳腺肿瘤早期筛查准确率依然是临床诊断中的关键问题。为解决现有检测方法在准确性与计算效率之间的矛盾, 提出了一种基于改进YOLOv10的轻量级检测算法, 命名为ZLP-YOLO。在该方法中, 采用SPDConv替换了传统卷积模块以显著降低计算复杂度, 并引入DySample上采样策略进一步增强了特征表达能力, 最后采用CoordAtt注意力机制, 使得模型在捕捉关键特征的同时实现了高效信息融合。实验结果显示, 所提算法将整体模型参数量从2.7 M降低到2.5 M, 计算量由8.4 GFLOPs大幅下降到0.9 GFLOPs; 在公开乳腺超声肿瘤数据集BUSI、BUS-UCLM上的检测精度mAP@.50分别提高了6.1%、6.7%, mAP@[.5:.95]分别提高了1.8%、1.4%。

关键词: 乳腺癌肿瘤; 深度学习; 目标检测; 轻量化算法; YOLOv10

DOI: 10.11907/rjdk.251263

开放科学(资源服务)标识码(OSID):

中图分类号: TP391.41

文献标识码: A

文章编号: 1672-7800(2025)007-0169-08



Lightweight Breast Tumor Detection Algorithm Based on Object Detection

ZHANG Aofeng, LI Dong, LIU Xiaojing

(School of Computer Technology and Application, Qinghai University, Xi'ning 810016, China)

Abstract: With the incidence of breast cancer continually rising, the accuracy of early tumor screening remains a critical issue in clinical diagnosis. To resolve the trade-off between detection accuracy and computational efficiency in existing methods, this paper proposes a lightweight detection algorithm based on an improved YOLOv10, named ZLP-YOLO. In this approach, we replace the conventional convolutional modules with SPDConv to significantly reduce computational complexity, introduce the DySample upsampling strategy to further enhance feature representation, and finally employ the CoordAtt attention mechanism to achieve efficient information fusion while capturing key features. Experimental results demonstrate that the proposed algorithm reduces the total model parameters from 2.7 million to 2.5 million and slashes the computational cost from 8.4 GFLOPs to 0.9 GFLOPs, while improving detection accuracy on the public BUSI and BUS-UCLM breast ultrasound tumor datasets by 6.1% and 6.7% in mAP@.50, and by 1.8% and 1.4% in mAP@.50-95, respectively.

Key Words: breast tumor; deep learning; object detection; lightweight algorithm; YOLOv10

0 引言

乳腺癌在女性恶性肿瘤中致死率较高, 其发病率逐年上升且被证实为全球癌症第五大主要死因^[1-2]。鉴于此, 乳腺癌的早期检测与诊断对提高患者生存率至关重要。随着人工智能技术的持续进步, 其与机器学习在医学影像领域的融合日益深入, 逐渐成为常规影像学辅助诊断的重

要工具。

阶段目标检测算法命名为YOLO算法, 其为单阶段目标检测算法^[3]。与双阶段目标检测算法(如Faster R-CNN^[4])不同, 单阶段目标检测算法无需进行候选框的生成和标注, 即可直接对输入图像进行精确预测。同时, 其在检测性能、检测精度方面有着独特优势。然而, 由于实际所需部署医用仪器存在设备老旧、算力不足等问题, 故单阶段目标检测算法往往难以满足模型计算需求。

收稿日期: 2025-05-20

扫描二维码阅读全文:

基金项目: 国家自然科学基金项目(62366043)

作者简介: 张教锋(1998-), 男, 青海大学计算机技术与应用学院硕士研究生, 研究方向为深度学习与目标检测; 李东(1987-), 男, 博士, 青海大学计算机技术与应用学院副教授、硕士生导师, 研究方向为神经网络和医学图像处理; 刘晓静(1980-), 女, 硕士, 青海大学计算机技术与应用学院教授、硕士生导师, 研究方向为计算机视觉与神经网络。本文通讯作者: 刘晓静。



为此,本文基于YOLOv10提出了一种新颖的乳腺肿瘤检测方法:ZLP-YOLO。首先,将YOLOv10中原本的卷积模型替换为SPDConv以提高整体模型的检测精度,在大幅度降低整体模型计算成本并提高检测精度的同时,提高整体模型的轻量化程度^[5];其次,引入了轻量级动态上采样DySample,以提高小目标检测性能^[6];再次,通过引入轻量化注意力机制CoordAttention提升算法对乳腺肿瘤图像的空间关系捕捉能力^[7]。实验结果表明,所提算法满足轻量化要求,同时检测性能也有所提升。

本文提出的轻量化检测算法在乳腺肿瘤检测任务上有着良好表现。在公开乳腺肿瘤图像数据集BUSI上,本文算法相较于原始YOLOv10算法,参数量由2.7 M减少到2.5 M,计算量由8.4 GFLOPs下降到0.9 GFLOPs,同时mAP@.50由0.724提升到0.785,mAP@.50-95由0.459提升到0.478,为乳腺肿瘤快速检测提供了方法支持。

本文主要贡献如下:①对公开数据集BUSI进行预处理,使用Labeling标注工具生成检测标签,共创建了2类共780张图像组成的乳腺肿瘤图像数据集;②将YOLOv10的传统卷积模型替换为SPDConv,大幅度降低了整体模型计算量,并提高了模型检测精度;③引入DySample上采样策略与CoordAttention注意力机制,进一步提高了整体模型的轻量化程度,并进一步提升了检测精度,使模型泛化能力大幅度增强。

1 相关工作

YOLO目标检测在检测精度方面效果显著,但由于其本身整体网络所需计算资源量大的特点,引来了一大批学者尝试在不改变精确度甚至是提高精确度的前提下降低了整体模型的参数量和计算量以满足算力有限的设备使用需求。例如:吴百强等^[8]在引入FasterNet以替换Yolov8骨干网络的同时引入Lion优化器,并提出了草莓病害检测算法。卢美光等^[9]在Yolov8的基础上引入MobileNetv3轻量级骨干网络,同时使用SimSPPF替换原模型中的SPPF以减少模型参数量并在引入CA注意力机制与WIoU损失函数的情况下使得模型的平均精确度提升10.3%。刘印等^[10]提出改进的Yolov10算法,利用Mamba线性注意力机制和DySample动态上采样模块替换原模型上采样,同时使用空间深度转换卷积替换原模型中卷积,使得改进后的算法相较于原模型参数下降了30.5%,计算量下降了19%。周翔等^[11]提出BIMAFPN以提升模型检测能力,同时引入LSCD损失函数,使得改进后的模型整体相较于改进前在精确度mAP上提高了2.7%,参数量下降了25%。易杰等^[12]将AKConv可变核卷积与iAFF迭代注意力特征融合后,引入损失函数Shape-IoU,使得整体模型检测精度提高了2.2%。

目标检测在医学领域应用广泛,如杨岚兰等^[13]在YO-

LOv5模型上引入ASFF与CBAM注意力机制,以提高整体算法精确度,将精确度由72.1%提升至79.6%。杨涛等^[14]在YOLOv5模型上引入BiFPN双向特征金字塔和CBAM注意力机制将模型整体精确度提升至8.4%。然而,这些改进通常会增加模型规模和计算量。以YOLOv5s为例,其已有约9.1 M参数和24.0 GBFLOPs,YOLOv8s也有约11.2 M参数,而引入BiFPN、可变形卷积等结构后模型会更庞大,精度提升也往往有限。总体而言,现有基于YOLO的方法虽然借助注意力机制和特征融合等技术取得了精度上的提升,但检测模型仍然偏大,在mAP@0.5和mAP@0.5:0.95等指标上提升空间有限,且针对乳腺超声图像的适应性不足。因此,亟需一种更轻量且高精度的YOLO模型以提高乳腺肿瘤超声检测的实用性。鉴于此,本文在YOLOv10的基础上提出了一种轻量化、高精度的乳腺肿瘤检测方法,以克服现有方法的不足并提升检测性能。

2 ZLP-YOLO改进算法

鉴于现有主流目标检测模型存在参数量和计算量庞大、难以在算力不足的检测设备上部署等问题,本文提出了一种改进的轻量化乳腺肿瘤检测方法:LTP-YOLO。其网络架构如图1所示。

2.1 SPDConv模块

针对乳腺肿瘤小型肿瘤多且难以检测的特点,在YOLOv10模型中引入SPDConv模块,该模块能够提高低分辨率图像及小物体识别能力,由两部分组成:首先是一个空间到深度(Space-to-Depth, SPD)层,其次是一个非跨步(Nonstrided)卷积层。其核心思想在于通过替换传统的步长(Strided)卷积和池化操作,以保留更多0细粒度信息,从而改善在处理小尺寸对象和低分辨率图像时的性能。尺度为2时的SPDConv如图2所示。

其中,SPD层对中间特征图进行切片操作,从而生成多个子特征图,并将子特征图基于原特征图的特定区域进行划分。每个子特征图通过特定的缩放因子对原特征图进行降采样,从而降低空间维度,同时增加通道维度。随后,采用一个非跨步卷积层对经过SPD层处理的特征作进一步处理,该层通过可学习参数减少通道数量,以避免信息冗余,同时保留关键的特征信息。

其原理如下:当输入图像尺寸为 $S \times S \times C_1$ 的中间特征图 X 时,首先将其按照隔点采样的方式划分为4个尺寸 $(S/2 \times S/2 \times C_1)$ 减半的特征图;再经拼接操作后得到尺寸为 $S/2 \times S/2 \times 4C_1$ 的特征图,这样能将特征图中空间维度的信息转换为通道维度信息;最后经过 1×1 卷积生成尺寸为 $S/2 \times S/2 \times C_1/2$ 的特征图。

2.2 DySample上采样算子

YOLOv10目标检测算法中的特征聚合网络仅仅基于像素的空间位置以确定上采样核,不能充分利用特征信

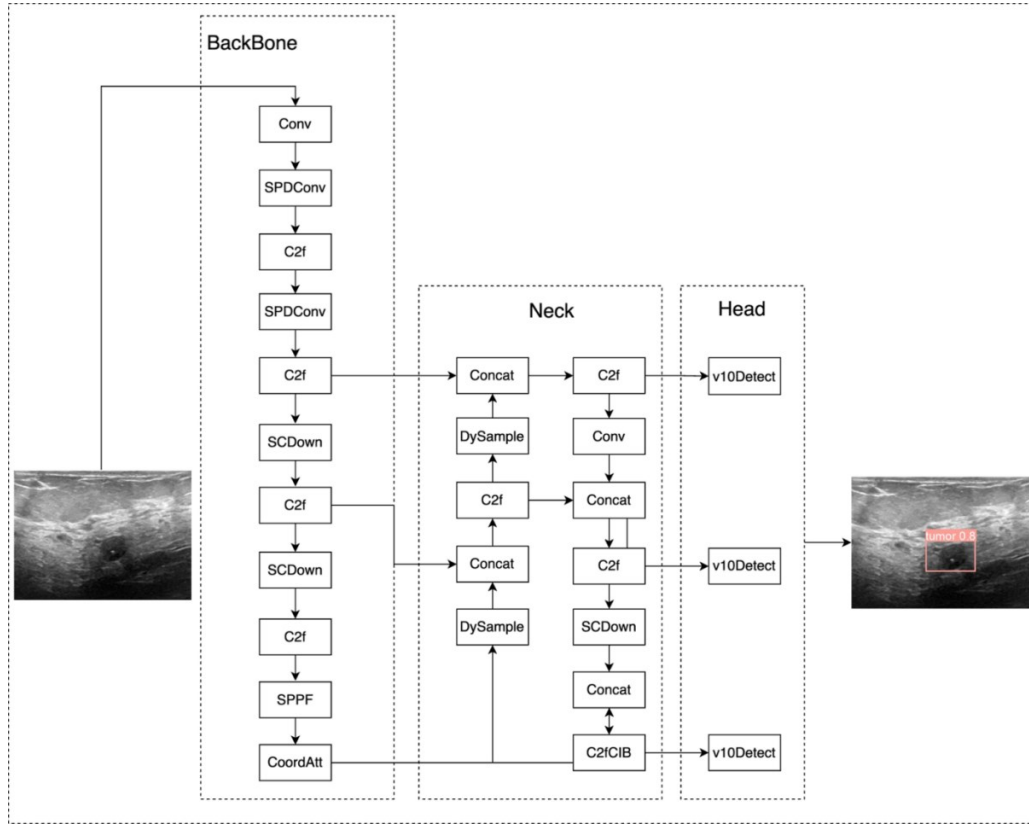


Fig. 1 Overall network structure of LTP-YOLO algorithm

图 1 LTP-YOLO 算法整体网络结构

SPD-Conv for Low-Resolution Images and Small Objects

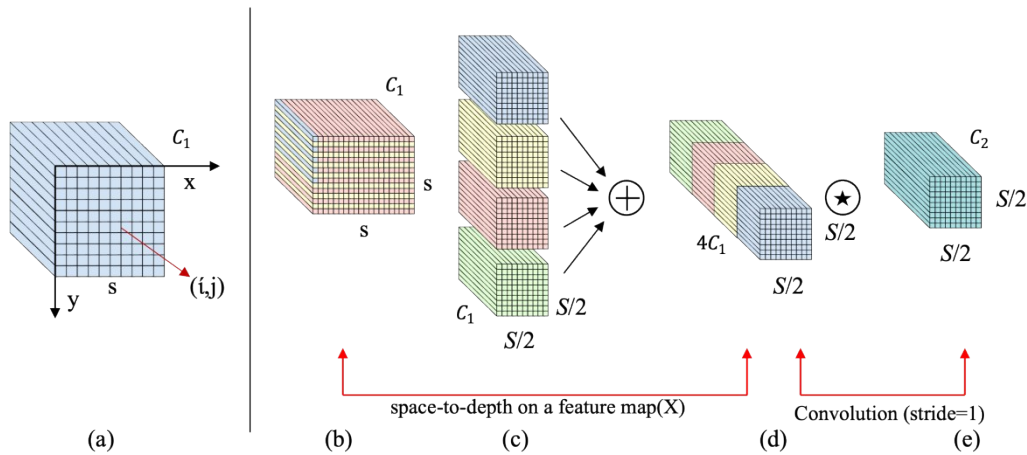


Fig. 2 Illustration of SPDConv when scale = 2

图 2 尺度为 2 时的 SPDConv

息。因此, Wang 等^[15]提出了内容感知特征重组 CARAFE 上采样算子,该上采样算子具有较大的感受野,能够更好地利用周围的特征信息,从而有效提升算法识别精度,但其引入了大量工作负载。为了减少算力开支,可从点采样的角度提出 DySample^[10]。图 3 展示了 DySample 的核心思想,即通过动态采样点重新采样输入特征图,从而实现上采样。具体描述为:首先,在生成采样点之后,对于输入的特征图 X 利用生成后的采样点,通过双线性插值法从 X 中提取信息,进而生成新的特征图 X' 。其过程定义如式(1)

所示。

$$X' = \text{grid_sample}(X, S) \quad (1)$$

图 4 详细展示了 DySample 中采样点生成器模块的结构和运作方式。具体而言,模块接收一个上采样尺度因子和一个输入特征图,通过一个线性处理单元,生成一组偏移量。接着,通过像素重组操作,将偏移量调整为适合的形状。最后,采样集通过将偏移量与原始采样网格相结合而得到。这一过程实现了动态采样点的生成,使得模型能够根据输入特征图的内容自适应地调整采样策略,从而在

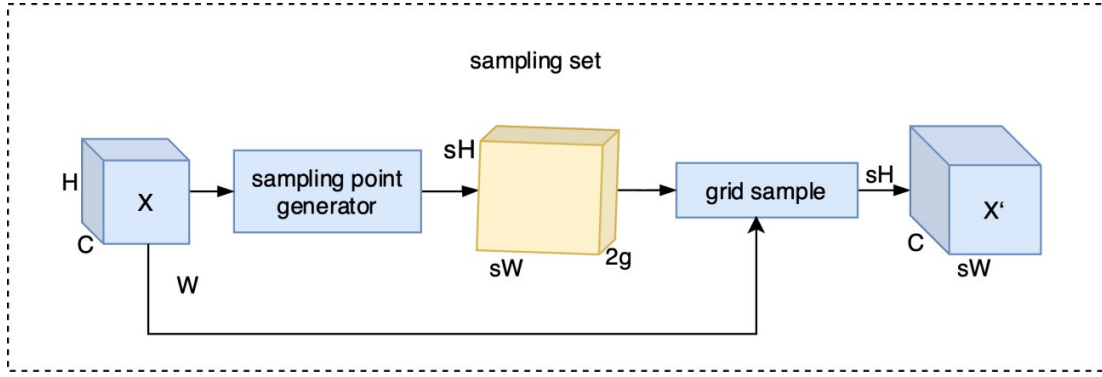


Fig. 3 Sampling based dynamic upsampling

图3 基于采样的动态上采样

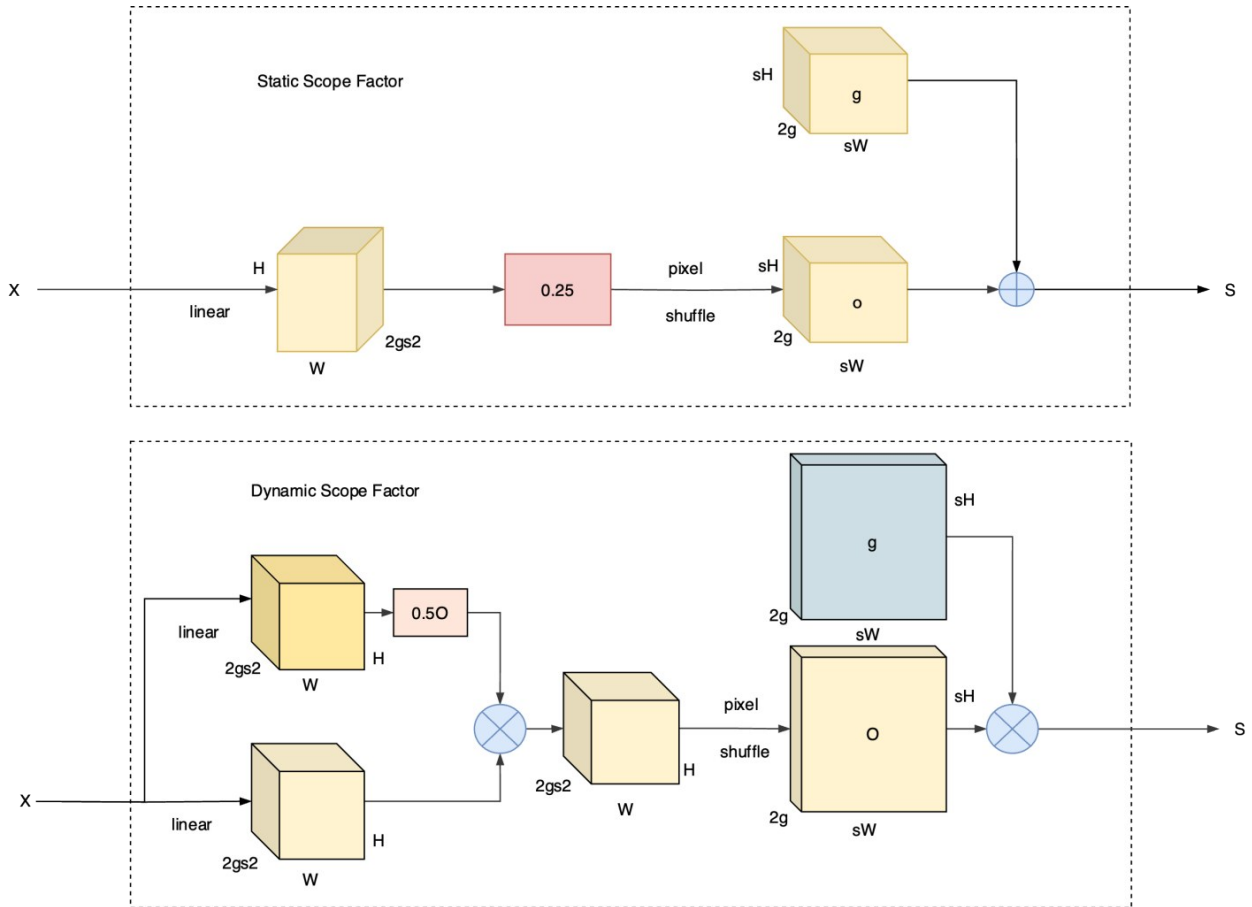


Fig. 4 Sampling point generator in DySample

图4 DySample中的采样点生成器

保持计算效率的同时增强特征表达能力。

2.3 CoorAtt 注意力机制

CoorAtt (CoordAttention) 作为一种在计算机视觉领域备受青睐的注意力机制,具备捕捉特定位置空间关系的能力,其结构如图5所示。与传统的注意力机制相比,CoordAttention 在计算过程中不仅深入考虑输入的特征信息,还兼顾每个像素点的位置信息,从而更精准地捕捉空间上的局部与全局关系。

CoordAttention 将通道注意力分为“X Avg Pool”和“Y Avg Pool”,也即将特征图转换为在水平方向和垂直方向进

行两组一维向量。在两个方向分别聚合信息特征后可相互补充信息,并最终生成能准确感知方向与位置的注意力特征图。这样能够更聚焦于需要关注的对象,同时也能增强特征图。

使用为 $(H, 1)$ 和 $(1, W)$ 的池化核沿水平与竖直方向对每个通道进行池化。在产生的两个向量中,水平方向中宽度为 w 的第 c 个通道的输出 $Z_c^w(h)$ 可表示为如式(2)所示。

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

同理,在垂直方向高度为 h 的第 c 通道处的输出可表示为如式(3)所示。

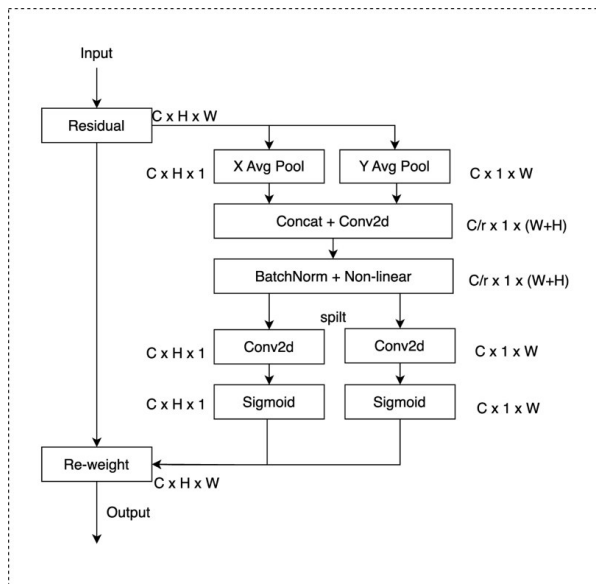


Fig. 5 CoordAttention architecture

图5 CoordAttention 结构

$$Z_c^h(h) = \frac{1}{W} \sum_{0 \leq j \leq w} x_c(i, j) \quad (3)$$

其中, $x_c(j, w)$ 、 $x_c(i, j)$ 分别表示第 c 个通道上位置为 (j, w) 和 (i, j) 位置的值。

借助 CoordAtt 模块实现坐标信息的嵌入与坐标注意力的生成,不仅能够捕获跨通道的丰富信息,同时还能精确掌握特征图内部物体的方向与位置细节。

3 实验与结果分析

3.1 实验平台与评价指标

采用 Linux 操作系统, CPU 为 Intel Xeon CPU E5-2603 v4@1.70 GHz, 16.0 GB 内存, 使用 Nvidia GTX1080Ti 显卡训练。实验的批量大小设置为 4, 训练周期为 500 个 epoch。训练是图像尺寸为 640×640。

为实现轻量化的同时兼顾精确度的提升, 选用 mAP (平均检测精度)、FPS、Params (参数量)、GFLOPs (计算量) 以衡量算法检测能力。mAP 的计算方式如式 (4) 和式 (5) 所示。

$$AP = \int_0^1 P(r) dr, \quad (4)$$

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

其中, mAP50 为 IOU 阈值为 0.5 时预测精度, mAP50-95 在 IoU=0.50、0.55、...、0.95 (共 10 个阈值, 步长 0.05) 下分别计算 mAP, 再对这 10 个值取平均, 反映模型在从宽松到严格重叠要求下的整体性能。GFLOPs 表示模型在一次前向推理过程中所需执行的十亿级浮点运算次数, 用以衡量计算复杂度。Params (参数量) 主要表示模型大小, FPS 指模型在推理阶段每秒钟能够处理的图像帧数, 也即单位时间内模型能完成多少张图片的前向计算和检测任务。

3.2 数据集搭建

本文实验基于 Scholar 平台提供的乳腺超声图像数据集 BUSI 而开展, 其共计 780 张超声图像。使用 LabelImg 标注工具对乳腺超声图像进行标注, 使其分为两类, 即有肿瘤、无肿瘤。训练集、验证集按照 8:2 的比例随机划分, 分别为 624 张、156 张图像。同时, 为证明改进后模型的泛化能力, 在 BUS-UCLM 数据集的基础上同样使用 LabelImg 标注工具获得可以用来进行目标检测的数据集, 并按照 8:2 的比例随机将其划分为训练集和验证集, 分别为 546、137。

3.3 实验结果分析

3.3.1 不同卷积模块对比实验

卷积模块在乳腺肿瘤目标检测任务上的检测能力各有优略, 本文以 SPDConv、DSConv、AKConv 和 SCConv 等卷积模型进行对比实验, 其结果如表 1 所示。

Table 1 Comparative study of different convolutional models (based on the BUSI dataset)

表 1 不同卷积模型对比实验 (基于 BUSI 数据集)

卷积模型	mAP@.5 /%	mAP@ [.5:.95] /%	Parameter (10 ⁶)	GFLOPs	FPS
DSConv ^[16]	73.1	47.1	2.70	7.8	285
AKConv	74.7	46.7	2.67	7.9	56
SCConv ^[17]	71.2	45.2	2.51	7.8	147
Shift-ConvNets ^[18]	72.3	45.0	2.87	11.1	212
LSKA ^[19]	76.0	46.7	2.30	7.2	263
RFB ^[20]	68.0	42.2	3.22	9.5	144
DynamicConv ^[21]	74.2	46.3	2.88	7.8	277
DySnakeConv ^[22]	75.6	47.4	3.15	9.3	102
EVCBlock ^[23]	73.7	46.7	6.99	11.7	166
RFAConv ^[24]	73.2	47.6	2.70	8.4	161
SPDConv	76.8	44.1	2.77	1.0	416

注: 加粗数据为选用模型、最优值。

由表 1 可知, 虽然 LSKA 与 SPDConv 都能够提高整体模型精确度, 但 LSKA 计算量减动幅度不大, 且在 FPS 指标上相较于 SPDConv 有明显差距, 故本文综合考虑上述性能指标后, 选择 SPDConv 替换 YOLOv10 中的卷积模块。SPDConv 能够在参数量提高 0.07 的情况下将计算量下降到基准模型的 11.9%, 且其 mAP@.5 上升 4.4% 的同时, mAP@[.5:.95] 虽下降 1.8%, 但 FPS 增加 83 帧/s。SPDConv 虽然在 mAP@[.5:.95] 上低于基准模型, 但考虑其参数量与计算量下降较大, 因而能满足移动端的时效性、精确度要求。

3.3.2 注意力机制对比实验

不同注意力机制对模型检测精度影响不同, 因而开展了多组对比实验, 结果如表 2 所示。可以看出, 卷积注意力模块 (Convolutional Block Attention Module) 能够将 mAP@.5 提高 1.2%, 但由于其 mAP@[.5:.95] 下降 0.7%, 故难以满足精确度提高的目标, 同时 SK、PPA 计算量难以满足轻量化的要求, 因此注意力机制最终考虑使用 CoordAtt。

为更加详细地解释算法的适用性, 本文使用 Grad-CAM 方法生成热力图, 如图 7 所示 (彩图扫 OSID 码可见),

Table 2 Comparative experiments of attention mechanisms (based on the BUSI dataset)**表 2** 注意力机制对比实验 (基于 BUSI 数据集)

注意力机制	mAP@.5 /%	mAP@ [.5:.95]/%	Parameter(10 ⁶)	GFLOPs	FPS
—	72.4	45.9	2.70	8.4	333
CBAM ^[25]	75.6	45.2	2.51	8.0	303
CoordAtt ^[26]	73.3	48.4	2.45	8.0	303
SK ^[27]	75.2	49.5	7.99	12.4	243
DAT ^[28]	70.1	43.1	3.18	9.2	238
ECA ^[29]	75.2	47.9	2.69	8.2	294
Triplet ^[30]	73.8	47.2	2.69	8.2	285
GAM ^[31]	75.3	45.4	4.33	9.5	270
PPA ^[32]	75.1	46.0	5.56	13.2	200
ACmix ^[33]	74.2	46.7	2.91	8.2	172
MSDA ^[34]	71.3	44.3	2.95	8.4	263
EMA ^[35]	72.0	46.1	2.70	8.4	285

注:加粗数据为选用模型、最优值。

该方法可以观察到算法在运行时的注意力分布^[36]。可以得知,本文算法在运行时能够更高地关注于目标区域,因而可以更加精准地识别并肿瘤位置。相对而言,诸如 YOLOv10 以及 CBAM、ECA、SK 等注意力机制,在关注点的集中度和精确性方面则显得不足,这进一步凸显了本文算法在理解和捕捉目标语义特征方面的显著优势。

3.3.3 不同模型对比实验

为了验证不同的目标检测方法在乳腺肿瘤超声图像中的目标检测能力,本文选取了不同模型进行对比。其中, YOLOv5+BiFPN+CBAM 与 YOLOv5+ASFF+CBAM 为杨涛等^[14]提出的用于乳腺超声图像检测的 YOLO 模型,具体实验结果如表 3 所示。

可以得知,传统 YOLOv10 相较于 RetinaNet、Faster-RCNN 等两阶段或高复杂度单阶段检测器,其计算量 8.4 GFLOPs 和参数量 2.70 M 均处于较低水平,且在 mAP@.5 (72.4%) 和 mAP@[.5:.95] (45.9%) 上的表现与这些重量级模型的差距并不显著。这说明原始 YOLOv10 在保持实时性和轻量化的前提下,已能兼顾基本的检测精度。进一步地, YOLOv10n 改进模型在将计算量压缩至 0.9 GFLOPs、参数量减少至 2.54 M 的同时, mAP@.5 提升至 78.5%, mAP@[.5:.95] 提升至 47.8%, 实现了更高精度和更低算力需求的兼得。同时, YOLOv10n 改进模型在 mAP@[.5:.95] 上已略

Table 3 Comparative experiments (based on the BUSI dataset)**表 3** 对比实验 (基于 BUSI 数据集)

方法	mAP@.5 /%	mAP@ [.5:.95]/%	Parameter (10 ⁶)	GFLOPs	FPS
YOLOv8n ^[37]	77.9	48.3	3.01	8.2	212
YOLOv10 ^[38]	72.4	45.9	2.70	8.4	333
Faster-RCNN	79.9	45.5	41.75	175	4.3
RetinaNet ^[39]	85.5	47.9	56.96	259	5.7
YOLOv5s+BiFPN+CBAM	76.9	46.9	7.4	16.6	238
YOLOv5s+ASFF+CBAM	76.6	48.5	12.71	24.49	245
YOLOv10n 改进	78.5	47.8	2.54	0.9	357

注:加粗数据为改进模型、最优值。

超 RetinaNet (47.9%) 和 Faster-RCNN (45.5%)。虽然其 mAP@.5 (78.5%) 仍略低于 RetinaNet (85.5%) 和 Faster-RCNN (79.9%), 这很可能与后者模型复杂度更高、检测头更多、对高 IoU 阈值下的候选框回归更为精准有关。因此, 在需要兼顾超高精度定位的场景中, RetinaNet 和 Faster-RCNN 依然具有一定优势; 而在极度受限的算力与存储环境下, YOLOv10n 改进模型通过大幅度轻量化带来的 357 FPS 推理性能 (高于原始 YOLOv10 的 333 FPS) 和接近甚至略超传统主流算法的平均精度, 展现了极佳的实用价值。

3.3.4 消融实验

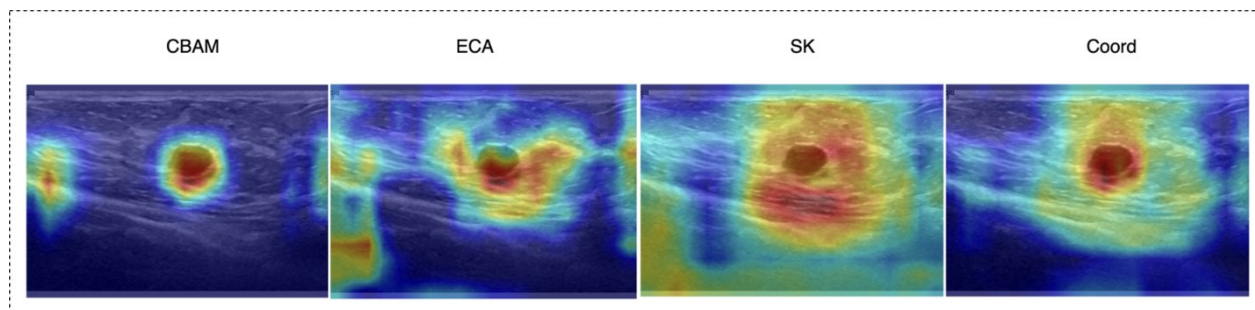
为验证本文算法各改进模块的有效性, 选择 YOLOv10 为基准模型开展消融实验, 并展示不同模块添加后的效果, 结果如表 4 所示。

Table 4 Module ablation experiment (based on the BUSI dataset)**表 4** 模块消融实验 (基于 BUSI 数据集)

SPDConv	DySample	CoordAtt	mAP@.5 /%	mAP@ [.5:.95] /%	Parameter (10 ⁶)	GFLOPs	FPS
—	—	—	72.4	45.9	2.70	8.4	333
√	—	—	76.8	44.1	2.77	1.0	416
—	√	—	70.2	46.2	2.70	8.2	357
—	—	√	73.3	48.4	2.45	8.0	303
√	—	√	75.4	45.7	2.53	0.9	400
—	√	√	72.3	47.6	2.46	8.0	294
√	√	√	78.5	47.8	2.54	0.9	357

注:加粗数据为最优值。

可以得知, 使用 SPDConv 模块替换 YOLOv10 中的卷积模块后, 在参数量增加 0.07 的情况下, mAP@.5 提升 4.4%, mAP@[.5:.95] 相对降低 1.8%, 但其计算量从 8.4 大幅下降

**Fig. 7** Heatmaps of different attention mechanisms (based on the BUSI dataset)**图 7** 不同注意力机制下的热力图 (基于 BUSI 数据集)

至0.9。其中,mAP@[.5:.95]的降低是因为SPDConv在低分辨率和小目标检测中表现出色,但在高IoU阈值下的性能可能会受到影响。这是因为SPDConv的设计主要是针对细粒度信息的保留,而在某些情况下可能导致特征表示不够全面,从而影响了在多个IoU阈值下的平均精度。

DySample的动态采样机制能够根据输入内容自适应地调整采样策略,使得模型在多个IoU阈值下的整体检测性能有所提升。但是,在同时使用3个模块结合的情况下能够取得良好效果,在计算量降低7.5 GFLOPs的同时,mAP@.5提升6.1%,mAP@[.5:.95]提升1.8%,同时FPS相较于基准模型YOLOv10增加24,表明了改进算法的有效性。

4 结语

本文算法相较于基准模型在实现了轻量化的同时还提升了整体模型的检测精度。通过对比实验(基于BUSI数据集)可知,虽然传统的YOLOv10在实时性(333 FPS)和轻量化(2.7 M参数、8.4 GFLOPs)方面具有显著优势,但其基础检测精度与重量级模型相比仍有差距,相比之下,RetinaNet和Faster-RCNN在高精度(mAP@.5分别为85.5%和79.9%)方面依然保持领先,但是其推理速度较慢(分别为5.7 FPS与4.3 FPS),不利于实时应用。本文所提出的YOLOv10n改进模型在保持极低算力消耗和高帧率的同时,实现了mAP@.5达78.5%、mAP@[.5:.95]达47.8%,其综合性能优于原始YOLOv10,并在多阈值平均精度上略超RetinaNet和Faster-RCNN,凸显了其在算力受限场景下的实用价值。

同时,在验证模型跨数据集泛化能力时,YOLOv10n改进模型和基准模型YOLOv10同时在BUS-UCLM数据集上进行评估,结果显示mAP@.50提高6.7%,mAP@[.5:.95]提高1.4%。该结果充分证明了本文方法不仅能在同源数据上取得优异结果,也能稳健适应不同来源的乳腺超声图像。

综上所述,本文提出的轻量级YOLOv10n改进模型不仅在两个公开乳腺超声数据集上取得了显著的精确度提升,而且在帧率与算力消耗方面依然保持领先水平,能够满足临床实时检测与资源受限部署的双重需求。

参考文献:

- [1] MERAJVER S D, MILLIRON K. Breast cancer risk assessment: a guide for clinicians using the NCCN breast cancer risk reduction guidelines [J]. Journal of the National Comprehensive Cancer Network, 2003, 1 (2): 297-301.
- [2] SUNG H, FERLAY J, SIEGEL R L, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries [J]. CA: a Cancer Journal for Clinicians, 2021, 71 (3): 209-249.
- [3] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection [C]//IEEE Conference on Computer Vision and Pattern Recognition, 2016: 779-788.
- [4] REN S, HE K M, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 39 (6): 1137-1149.
- [5] SUNKARA R, LUO T. No more strided convolutions or pooling: a new CNN building block for low-resolution images and small objects [DB/OL]. <https://arxiv.org/abs/2208.03641>, 2022.
- [6] LIU W, LU H, FU H, et al. Learning to upsample by learning to sample [DB/OL]. <https://arxiv.org/abs/2308.15085>, 2023.
- [7] HOU Q B, ZHOU D Q, FENG J S. Coordinate attention for efficient mobile network design [C]//IEEE Conference on Computer Vision and Pattern Recognition, 2021: 13708-13717.
- [8] WU B Q, LIU W K, LI Y. Design of strawberry disease detection and segmentation system based on improved YOLOv8 [J/OL]. China Stereology & Image Analysis, 2025-06-06. <http://kns.cnki.net/kcms/detail/11.3739.R.20250325.2358.002.html>.
吴百强, 刘文康, 李阳. 基于改进YOLOv8的草莓病害检测与分割系统设计 [J/OL]. 中国电视学与图像分析, 2025-06-06. <http://kns.cnki.net/kcms/detail/11.3739.R.20250325.2358.002.html>.
- [9] LU M G, WANG X C, ZHANG X L, et al. Lightweight recognition method for chrysanthemum based on improved YOLOv8 [J/OL]. Journal of Nanjing Agricultural University, 2025-06-06. <http://kns.cnki.net/kcms/detail/32.1148.s.20250327.1103.002.html>.
卢美光, 汪小畧, 张晓蕾, 等. 基于改进YOLOv8的金丝皇菊轻量化识别方法 [J/OL]. 南京农业大学学报, 2025-06-06. <http://kns.cnki.net/kcms/detail/32.1148.s.20250327.1103.002.html>.
- [10] LIU Y, GONG C Y, XU G D. A lightweight object detection algorithm based on improved YOLOv10 [J]. Automation & Information Engineering, 2025, 46 (1): 29-35.
刘印, 龚长友, 徐国栋. 基于改进YOLOv10的轻量化目标检测算法 [J]. 自动化与信息工程, 2025, 46 (1): 29-35.
- [11] ZHOU X, WANG K Q, ZHOU X X, et al. Helmet wearing detection algorithm for electric vehicles based on improved YOLOv10n [J]. Electronic Measurement Technology, 2025, 48 (5): 40-49.
周翔, 王可庆, 周新翔, 等. 基于改进YOLOv10n的电动车头盔佩戴检测算法 [J]. 电子测量技术, 2025, 48 (5): 40-49.
- [12] YI J, SHAN Y C. Strawberry maturity detection model at various developmental stages based on YOLOv10-IAS [J/OL]. Journal of Nanjing Agricultural University, 2025-06-06. <http://kns.cnki.net/kcms/detail/32.1148.S.20241223.1025.002.html>.
易杰, 单以才. 基于YOLOv10-IAS的草莓各个发育阶段成熟度检测模型 [J/OL]. 南京农业大学学报, 2025-06-06. <http://kns.cnki.net/kcms/detail/32.1148.S.20241223.1025.002.html>.
- [13] YANG L L, YANG T, ZHAO H J. Application of improved YOLOv5 model in ultrasound detection of breast tumors [J]. Automation Applications, 2025, 66 (4): 37-41.
杨岚兰, 杨涛, 赵红佳. 改进型YOLOv5模型在乳腺肿瘤超声图像检测中的应用 [J]. 自动化应用, 2025, 66 (4): 37-41.
- [14] YANG T, YANG L L, YANG M Y, et al. Application of improved YOLO network object detection algorithm in ultrasound detection of breast tumors [J]. China Medical Equipment, 2024, 21 (9): 23-27.
杨涛, 杨岚兰, 杨米扬, 等. 基于改进型YOLO网络目标检测算法在乳腺肿瘤超声图像检测中的应用研究 [J]. 中国医学装备, 2024, 21 (9): 23-27.
- [15] WANG J Q, CHEN K, XU R, et al. Carafe: content-aware reassembly of-

- features[C]//Proceedings of the IEEE International Conference on Computer Vision, 2019: 3007–3016.
- [16] NASCIMENTO M G, FAWCETT R, PRISACARIU V A. Dsconv: efficient convolution operator [C]//Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019: 5148–5157.
- [17] LI J F, WEN Y, HEL H. SCConv: spatial and channel reconstruction convolution for feature redundancy [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 6153–6162.
- [18] LI D, LI L, CHEN Z, et al. Shift-ConvNets: small convolutional kernel with large kernel effects [DB/OL]. <https://arxiv.org/abs/2401.12736>, 2024.
- [19] LAU K W, PO L M, REHMAN Y A U. Large separable kernel attention: rethinking the large kernel attention design in CNN [DB/OL]. <https://arxiv.org/abs/2309.01439>, 2023.
- [20] LIU S T, HLUANG D, WANG Y H. Receptive field block net for accurate and fast object detection [C]//European Conference on Computer Vision, 2018: 404–419.
- [21] HAN K, WANG Y, GUO J, et al. ParameterNet: parameters are all you need for large-scale visual pretraining of mobile networks [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024: 15751–15761.
- [22] QI Y L, HE Y T, QI X M, et al. Dynamic snake convolution based on topological geometric constraints for tubular structure segmentation [C]//2023 IEEE/CVF International Conference on Computer Vision, 2023: 6047–6056.
- [23] QUAN Y, ZHANG D, ZHANG L Y, et al. Centralized feature pyramid for object detection [J]. IEEE Transactions on Image Processing, 2023, 32: 4341–4354.
- [24] ZHANG X, LIU C, YANG D, et al. RFConv: innovating spatial attention and standard convolutional operation [DB/OL]. <https://arxiv.org/abs/2304.03198>, 2023.
- [25] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention-module [C]//Proceedings of the European Conference on Computer Vision, 2018: 3–19.
- [26] HOU Q, ZHOU D, FENG J. Coordinate attention for efficient mobile network design [C]//Nashville: 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021.
- [27] LI X, WANG W, HU X, et al. Selective kernel networks [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019: 510–519.
- [28] XIA Z, PAN X, SONG S, et al. Vision transformer with deformable attention [C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022: 4784–4793.
- [29] WANG Q, WU B, ZHU P, et al. ECA-Net: efficient channel attention for deep convolutional neural networks [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 11531–11539.
- [30] MISRA D, NALAMADA T, ARASANIPALAI A U, et al. Rotate to attend: convolutional triplet attention module [C]//Proceedings of the IEEE Winter Conference on Applications of Computer Vision, 2021: 3139–3148.
- [31] LIU Y C, SHAO Z R, HOFFMANN N. Global attention mechanism: retain information to enhance channel-spatial interactions [DB/OL]. <https://arxiv.org/abs/2112.05561>, 2021.
- [32] XU S, ZHENG S C, XU W, et al. HCF-Net: hierarchical context fusion network for infrared small object detection [DB/OL]. <https://arxiv.org/abs/2403.10778>, 2024.
- [33] PAN X R, GE C J, LU R, et al. On the integration of self-attention and convolution [C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022: 815–825.
- [34] JIAO J Y, TANG Y M, LIN K Y, et al. DilateFormer: multi-scale dilated transformer for visual recognition [J]. IEEE Transactions on Multimedia, 2023, 25: 8906–8919.
- [35] OUYANG D L, HE S, ZHANG G Z, et al. Efficient multi-scale attention module with cross-spatial learning [C]//Rhodes Island: 2023 IEEE International Conference on Acoustics, Speech and Signal Processing, 2023.
- [36] SELVARAJU R R, COGSWELL M, DAS A, et al. Grad-CAM: visual explanations from deep networks via gradient-based localization [J]. International Journal of Computer Vision, 2020, 128: 336–359.
- [37] VARGHESE R, SAMBATH M. YOLOv8: a novel object detection algorithm with enhanced performance and robustness [C]//Chennai: The International Conference on Advances in Data Engineering and Intelligent Computing Systems, 2024.
- [38] WANG A, CHEN H, LIU L, et al. YOLOv10: real-time end-to-end object detection [DB/OL]. <https://arxiv.org/abs/2405.14458v2>, 2024.
- [39] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(2): 318–327.

(责任编辑: 孙娟)