

# 基于 MSEF-YOLO 的儿童腕部骨折检测算法

官 硕, 蒋 强, 李婷雪

(沈阳理工大学 自动化与电气工程学院, 沈阳 110159)

**摘要:** 针对儿童腕部 X 光图像中细微骨折检测精度较低的问题, 提出一种基于 YOLOv11n 改进的 MSEF-YOLO (multi-scale efficient fusion network-YOLO) 目标检测算法。首先, 将空间和通道重构卷积 (SCConv) 模块与 C3k2 模块融合, 通过空间重构单元 (SRU) 和通道重构单元 (CRU) 并行处理空间与通道的冗余, 增强对小目标的感知能力; 其次, 引入多尺度扩张注意力 (MSDA) 机制提高特征提取能力, 进而提高模型检测精度与泛化性, 有效减少漏检和误检; 最后, 优化尺度序列特征融合 (SSFF) 模块并设计 SSFF-X 模块, 通过 3D 卷积增强多尺度特征融合能力, 进一步提升对细微骨折的检测效果。实验结果表明, 相较于原 YOLOv11n 算法, MSEF-YOLO 算法的精确率、召回率、mAP@0.5 和 mAP@0.5~0.95 分别提高了 3.1%、3.8%、3.0% 和 3.5%。MSEF-YOLO 算法能够有效协助放射科医生检测儿童腕部骨折, 为医学图像的诊断提供技术支持。

**关键词:** 儿童腕部骨折; 空间和通道重构卷积; 多尺度扩张注意力; 特征融合

**中图分类号:** TP391.4 **文献标志码:** A **DOI:** 10.3969/j.issn.1003-1251.2025.05.005

## Pediatric Wrist Fracture Detection Algorithm Based on MSEF-YOLO

GONG Shuo, JIANG Qiang, LI Tingxue

(Shenyang Ligong University, Shenyang 110159, China)

**Abstract:** To address the issue of low detection accuracy for subtle fractures in pediatric wrist X-ray images, an improved YOLOv11n-based object detection algorithm called MSEF-YOLO (multi-scale efficient fusion network-YOLO) is proposed. First, the spatial and channel reconstruction convolution (SCConv) module is integrated with the C3k2 module, utilizing the spatial reconstruction unit (SRU) and channel reconstruction unit (CRU) to process spatial and channel redundancy in parallel, thereby enhancing the perception of small objects. Second, the multi-scale dilated attention (MSDA) mechanism is introduced to improve feature extraction capability, thereby enhancing detection accuracy and generalization ability while effectively reducing missed and false detections. Finally, the scale sequence feature fusion (SSFF) module is optimized and the SSFF-X module is designed, leveraging 3D convolution to enhance multi-scale feature fusion, further improving the detection performance for subtle fractures. Experimental results demonstrate that, compared to the original YOLOv11n algorithm, the MSEF-YOLO algorithm improves precision, recall, mAP@0.5, and mAP@0.5~0.95 by 3.1%, 3.8%, 3.0%, and 3.5%, respectively. The MSEF-YOLO algorithm effectively assists radiologists in detecting pediatric wrist fractures, providing technical support for medical image diagnosis.

**Key words:** pediatric wrist fracture; spatial and channel reconstruction convolution; multi-scale dilated attention; feature fusion

儿童腕部骨折(临床上称为柯力氏骨折)是常见的骨科疾病之一,普遍发生在活动量比较大的孩子中,大都由于跌倒后手掌着地所导致。腕部骨折的及时诊断对于避免长期功能障碍至关重要。

尽管成像技术在骨骼异常诊断方面已经非常成熟,但一些微小或早期的骨折可能表现为隐匿性,常规成像方法难以清晰显示骨折情况,可能造成漏检或误检,导致患者治疗延误或恢复不良,尤其在紧急情况下医学影像诊断的错误率较高。

近年来,深度学习技术的迅速发展推动了医学影像分析的进步,研究者可借助深度学习方法对人体骨折部位进行识别与分析。深度学习模型具有强大的数据处理能力,可显著提高医学图像诊断准确性。Thaiyalnayaki 等<sup>[1]</sup>采用 CrackNet 网络优化 Faster R-CNN,通过在两阶段检测框架中引入离散小波变换(DWT)提升了检测准确率,同时降低了计算复杂度,但其仅使用 100 个样本,远低于医学影像检测任务的常规需求,易引发过拟合。Thian 等<sup>[2]</sup>提出了 Inception-ResNet Faster R-CNN 架构,用于腕部桡骨及尺骨骨折 X 光片检测,在正位和侧位 X 光片上取得了较好的检测效果,但该网络对一些正常的尺骨茎突进行了假阳性标记。Zou 等<sup>[3]</sup>提出了 YOLOv7-ATT 模型,引入注意力机制且采用 EIoU 损失函数,提升了全身骨折检测的性能,但对极端微小目标的敏感度表现不佳。Wu 等<sup>[4]</sup>针对脊柱骨折检测提出了改进 YOLOv7 模型,通过整合 ELAN 与 MPCConv 模块增强小目标特征,引入 Wasserstein 损失函数解决类别不平衡问题,采用动态归一化提升泛化能力,但由于样本量较少,模型在泛化性上仍存在不足。周涛等<sup>[5]</sup>基于 YOLOv5 设计了主辅网络架构,通过多尺度特征优化和引入注意力机制,显著提升了下颌骨骨折检测的精度,但同时也增加了计算复杂度。Zhou 等<sup>[6]</sup>提出了一种 M<sup>3</sup>YOLOv5 模型,引入全局-局部特征提取模块、深浅特征交互模块和多尺度多感受野特征混合模块,提升了下颌骨骨折区域的检测效果,但其样本数量有限,对于复杂或不常见的骨折类型检测效果不佳。

YOLOv11n 是 YOLOv11 系列中最轻量级的模型,在检测精度与速度之间实现了良好的平衡,易部署于临床辅助诊断系统,本文对其进行优化改进,提出 MSEF-YOLO(multi-scale efficient fusion network-

YOLO)目标检测算法。具体改进内容如下:

1)将空间和通道重构卷积(SCConv)<sup>[7]</sup>模块与 C3k2 模块融合,减少特征图中的冗余信息,解决对细微骨折检测精度较差的问题;

2)引入多尺度扩张注意力(MSDA)机制<sup>[8]</sup>,通过使用不同的扩张率有效提取多尺度特征,提升特征表示的多样性,解决在定位不同尺寸目标时存在的漏检与误检问题;

3)对尺度序列特征融合(SSFF)<sup>[9]</sup>模块进行优化,设计 SSFF-X 模块,通过 3D 卷积和尺度调整、维度转换、特征融合等操作,解决多尺度特征融合不充分的问题,使网络能够更有效地提取图像信息。

## 1 MSEF-YOLO 算法

MSEF-YOLO 算法的网络结构如图 1 所示。图中:Conv 表示标准卷积层;C3k2 表示特征提取模块;C3k2-SCConv 表示引入 SCConv 的特征提取模块;SPPF 表示快速空间金字塔池化模块;C2PSA-MSDA 表示含有 MSDA 的特征提取模块;Zoom-cat 表示对不同尺度的特征图进行拼接;Concat 表示拼接操作;Add 表示加法操作;SSFF-X 表示基于 SSFF 模块改进的特征融合模块;Detect 表示检测头。

### 1.1 C3k2-SCConv 模块

使用 SCConv 模块替换 C3k2 模块中 Bottleneck 的第二个卷积层,构成 C3k2-SCConv 模块。SCConv 由空间重构单元(SRU)和通道重构单元(CRU)构成<sup>[10]</sup>,通过并行分支处理空间与通道冗余,再对各结果进行融合,使模型更轻量化<sup>[11]</sup>,其网络结构如图 2 所示。

#### 1) SRU 模块

SRU 采用分离和重构的方法抑制特征图中的空间冗余并增强特征的空间相关性。

在分离阶段通过组归一化(GN)计算特征图的标准输出  $X_{out}$ ,其表达式为

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

式中: $X$  为输入特征; $\mu$  和  $\sigma$  分别为特征值的均值和标准差; $I$  为单位矩阵; $\varepsilon$  是一个很小的正数,用于保证除法稳定性; $\gamma$  和  $\beta$  是可训练参数。

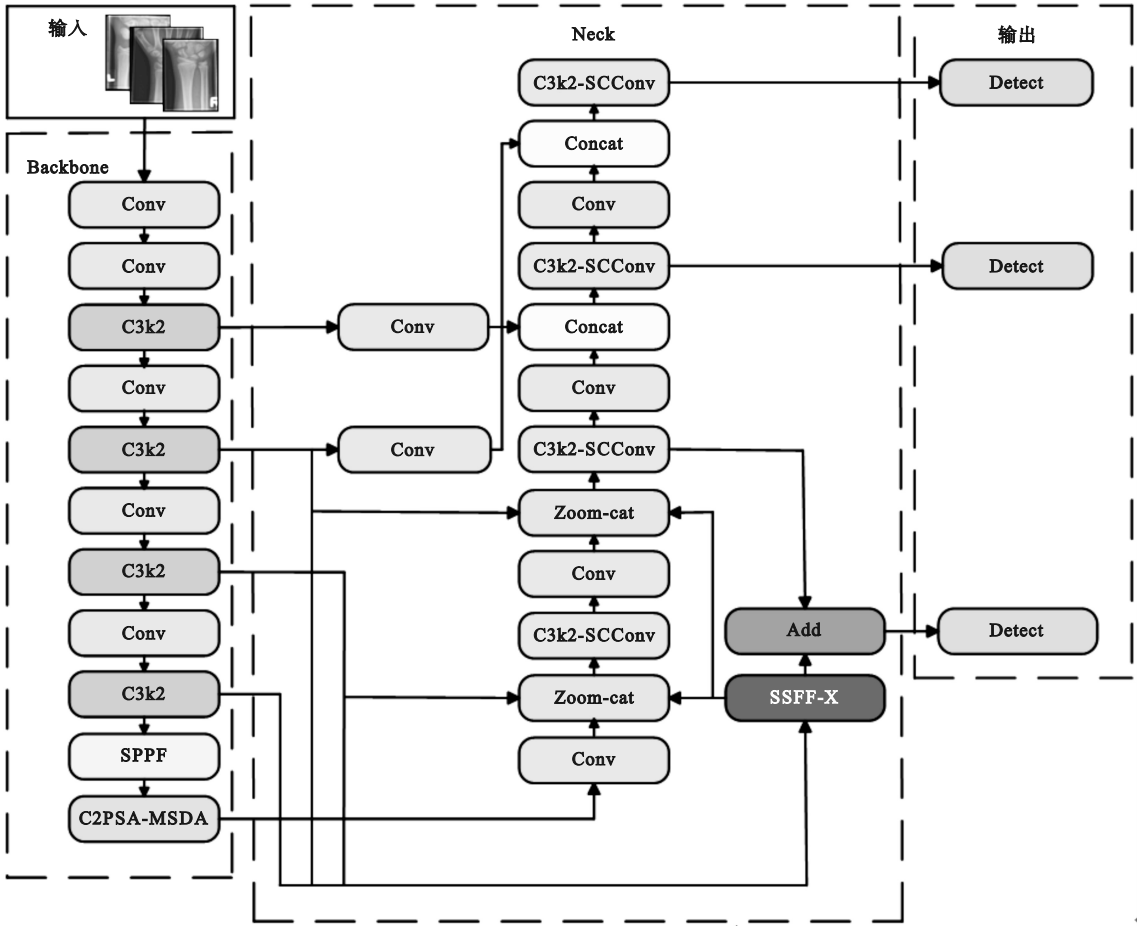


图 1 MSEF-YOLO 算法网络结构

Fig. 1 MSEF-YOLO algorithm network structure

计算归一化相关性权重  $W, W = \{w_i\}$ , 用于评估不同特征图的重要性,  $w_i$  计算式为

$$w_i = \frac{\gamma_i}{\sum_{j=1}^C \gamma_j} \quad (2)$$

式中:  $C$  为特征图的通道数;  $\gamma_i$  和  $\gamma_j$  分别为第  $i$  个和第  $j$  个通道的学习参数 ( $i, j = 1, 2, \dots, C$ )。

通过 Sigmoid 函数将通道权重映射到  $(0, 1)$  区间, 并设置门控阈值。若权重高于阈值, 则将权重设为 1, 由此获得信息权重  $W_1$ ; 若权重低于阈值, 则将权重设为 0, 由此获得非信息权重  $W_2$ 。

在重构阶段, 将输入特征  $X$  分别乘  $W_1$  和  $W_2$ , 得到两个加权特征。为减少空间冗余, 引入交叉重构操作, 拼接交叉重构的特征  $X^{w1}$  和  $X^{w2}$ , 得到空间细化特征图  $X^w$ 。整个重构过程表示如下。

$$\begin{cases} X_1^w = W_1 \otimes X \\ X_2^w = W_2 \otimes X \\ X_{11}^w \oplus X_{22}^w = X^{w1} \\ X_{21}^w \oplus X_{12}^w = X^{w2} \\ X^{w1} \cup X^{w2} = X^w \end{cases} \quad (3)$$

式中:  $\otimes$  表示元素相乘;  $\oplus$  表示元素求和;  $\cup$  表示拼接操作;  $X_1^w$  为信息量较大的加权特征图;  $X_2^w$  为信息量较小的加权特征图;  $X_{11}^w$  与  $X_{12}^w$  是由  $X_1^w$  拆分的子特征;  $X_{21}^w$  与  $X_{22}^w$  是由  $X_2^w$  拆分的子特征。

### 2) CRU 模块

CRU 用于减少通道维度上的冗余, 从而提高计算效率, 其网络结构由分离、变换、融合三部分组成。

在分离阶段, 先将从 SRU 结构得到的空间细化特征  $X^w$  分离成两个通道, 一个通道数为  $\alpha C$ , 另一个通道数为  $(1 - \alpha) C$ , 其中  $\alpha$  为分割比。然后分别使用  $1 \times 1$  卷积压缩通道, 将空间细化特征  $X^w$  划分为上部特征  $X_{up}$  与下部特征  $X_{low}$ 。

在变换阶段, 对同一  $X_{up}$  进行分组卷积 (GWC) 和逐点卷积 (PWC) 后, 将输出进行整合, 得到特征图  $Y_1$ 。对  $X_{low}$  仅进行 PWC 操作, 生成浅层细节特征, 并将其与原  $X_{low}$  整合得到特征图  $Y_2$ 。

在融合阶段, 采用简化的选择性卷积网络 (SKNet)<sup>[12]</sup> 自适应合并特征  $Y_1$  和  $Y_2$ 。首先通过池化聚合两个通道上的全局空间信息  $S_1$  与  $S_2$ , 然

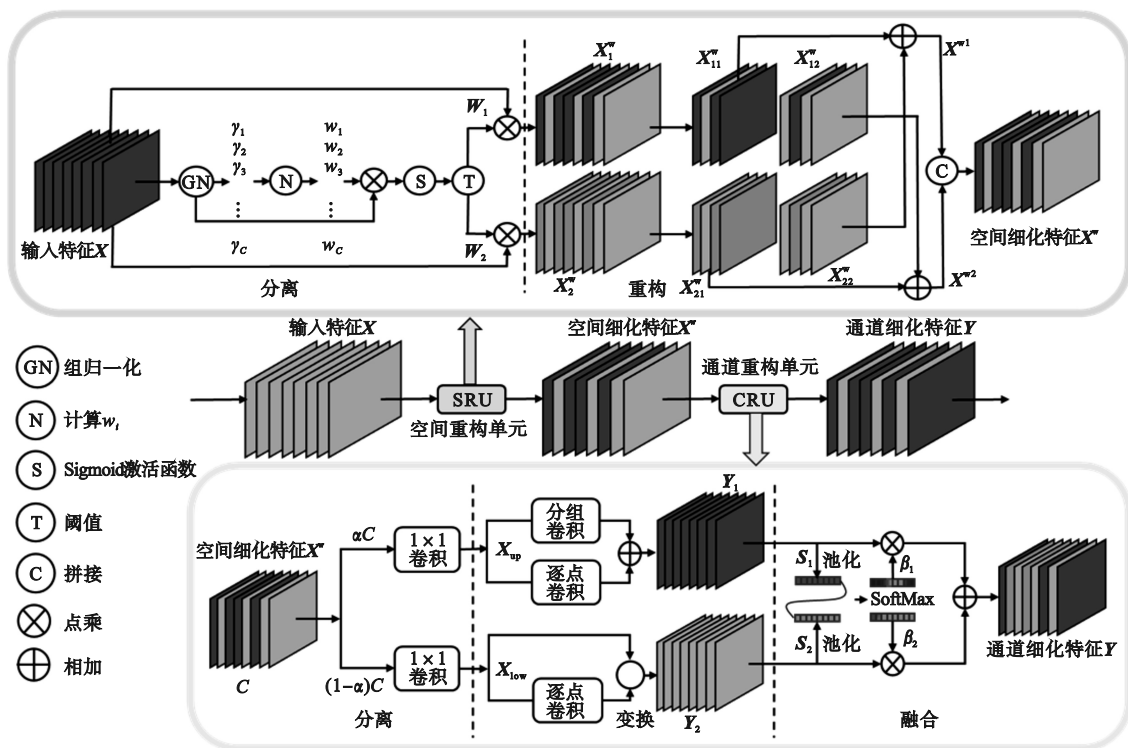


图 2 SCConv 模块网络结构

Fig.2 SCConv module network structure

后通过 SoftMax 激活函数生成特征重要性向量  $\beta_1$  和  $\beta_2$ , 最后将  $Y_1$  与  $Y_2$  合并得到通道细化特征  $Y$ , 表达式为

$$Y = \beta_1 Y_1 + \beta_2 Y_2 \quad (4)$$

在骨折检测中, 小目标更加依赖底层局部细节, 在 X 光图像中易被忽略。SRU 模块能够增强小目标边缘信息, 且通过门控保留有效特征。CRU 模块通过 GWC 提取局部细节、通过 PWC 整合全局信息, 提高小目标的检测精度。SCConv 采用  $1 \times 1$  卷积, 其参数量较传统的  $3 \times 3$  卷积减少, 模型更加高效。C3k2-SCConv 模块通过平衡特征提取能力和计算效率, 增强对小目标的感知能力, 从而减少漏检和误检。

### 1.2 MSDA 机制

MSDA 机制的核心思想是通过引入不同的扩张率实现多尺度的特征提取, 增强特征表示。

在儿童腕部骨折 X 光图像中, 骨折呈现出不同的大小与形态, 扩张率的大小直接影响感受野的范围。采用较小的扩张率有助于捕捉微小裂缝和细微骨折, 提高对隐匿骨折的检测效果; 采用较大的扩张率则能够扩展感受野, 便于识别大范围的骨折情况; 采用中等扩张率可在细节特征与全局特征之间保持平衡。

MSDA 模块的结构如图 3 所示, 图中  $r$  表示

扩张率。首先, 输入特征图  $X$  经过线性投影, 生成查询 (Q)、键 (K) 和值 (V) 矩阵。然后, 将输入特征图的通道划分为多个头, 每个头在独立的空间中执行滑动窗口扩张注意力操作, 在不同的头中以不同的扩张率在 Q 和 V 之间执行自注意力操作, 以提取不同尺度的特征信息。将每个头执行完自注意力后的结果拼接在一起, 形成更加丰富的特征表示。最后, 这些拼接后的特征在一个线性层中进行聚合, 生成最终的特征表示。

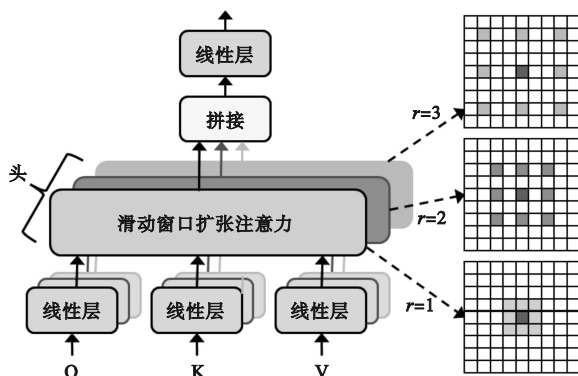


图 3 MSDA 模块结构

Fig.3 MSDA module structure

尽管 MSDA 能够增强特征提取的多样性与准确性, 但仍存在过拟合问题。为此, 在儿童腕部骨折 X 光图像检测中使用 Dropout 正则化。

Dropout 能够避免模型过度依赖特定特征,增强对小骨折、隐性骨折等不同骨折类型的适应能力。

在儿童腕部骨折 X 光图像中,MSDA 能够捕捉局部细节和全局特征,从而提高对不同类型和不同尺寸骨折的识别能力,提升目标检测的效率和鲁棒性。

### 1.3 SSFF-X 模块

本文对 SSFF 模块进行优化设计,提出 SSFF-X 模块,通过 3D 卷积使不同尺度的特征图能够在新的维度上进行深度融合,更好地捕捉骨折图像中的细微变化。SSFF-X 模块的输入特征图  $F_{\sigma}(x,y)$  按以下方式进行构造。

$$F_{\sigma}(x,y) = \sum_u \sum_v f(x-u,y-v) \cdot G_{\sigma}(u,v) \quad (5)$$

$$G_{\sigma}(m,n) = \frac{1}{2\pi\sigma^2} e^{-\frac{m^2+n^2}{2\sigma^2}} \quad (6)$$

式中:  $(x,y)$  表示输入特征图的位置坐标;  $f(x-u,y-v)$  表示二维特征图的像素位置;  $G_{\sigma}$  表示具有标准差  $\sigma$  的二维高斯核函数值;  $m$  和  $n$  表示像素点在水平方向和垂直方向的坐标。

SSFF-X 模块的结构如图 4 所示。首先使用  $1 \times 1$  的卷积将通道数统一为 512,避免后续拼接

过程中出现维度不匹配的情况。然后通过双线性插值(F.interpolate)将  $P4$  和  $P5$  的尺寸调整至与  $P3$  一致( $P3$ 、 $P4$ 、 $P5$  是对特征金字塔主干网络输出特征图通过不同采样方式得到的特征图)。3D 卷积的引入使模型能够在空间和深度两个维度上同时提取特征,克服采用传统 2D 卷积仅能处理图像高度和宽度信息的局限性。在儿童腕部骨折 X 光图像中,骨折特征不仅体现在平面上,还存在于不同灰度层次的深度信息中,3D 卷积能够更好地捕捉这些细节层次的信息,提升对细微骨折的检测能力。为使用 3D 卷积进行特征融合,采用 Unsqueeze 函数将特征图的形状由三维张量转换为四维张量。得到特征图后,先使用 Concat 将其拼接为包含不同尺度特征的 3D 特征块,再通过 3D 卷积进行多尺度特征融合。使用三维批量归一化(Batch Norm 3D)和 ReLU 激活函数<sup>[13]</sup> 提高训练的稳定性。为进一步融合不同尺度的信息,在 Z 轴方向进行 3D 最大池化(MaxPool 3D),压缩跨尺度的特征信息。最后,使用 Squeeze 操作去除多余的 Z 轴维度,使输出特征图恢复到三维张量(高度、宽度、通道数)。

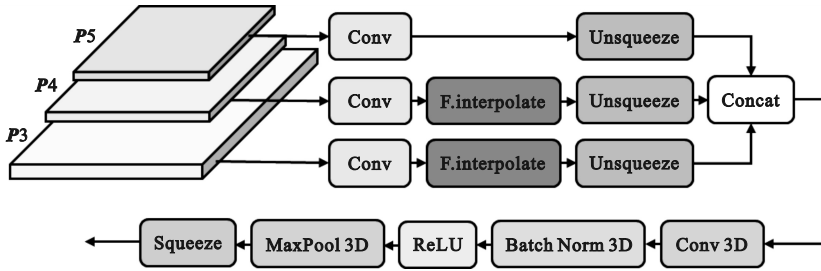


图 4 SSFF-X 模块结构

Fig.4 SSFF-X module structure

SSFF-X 模块通过尺度调整、维度转换、特征融合、归一化、激活以及池化操作,在不同尺度之间进行深度特征融合,捕获更加丰富的图像信息。对于儿童腕部骨折 X 光图像,SSFF-X 模块能够有效融合不同尺度的特征信息,提升检测精度,增强对不同骨折类型的适应性。

## 2 实验与结果分析

### 2.1 实验平台及参数设置

本实验的操作系统为 Ubuntu 22.04,其详细实验环境配置如表 1 所示。

具体参数设置:批次大小为 32,输入图像尺寸为  $1024 \times 1024$  像素,采用随机梯度下降算法,初

始学习率为 0.01,学习率动量为 0.937,权重衰减系数为 0.0005,预热初始偏置学习率为 0.1。

表 1 实验环境配置

Table 1 Experimental environment configuration

实验环境	配置
中央处理器 (CPU)	AMD EPYC 7542 2.9 GHz
图形处理器 (GPU)	NVIDIA GeForce RTX 4090
编程语言及版本	Python 3.11
深度学习框架	PyTorch 2.1.2
加速计算框架	CUDA Toolkit 11.8.0

### 2.2 数据集与评价指标

本文采用奥地利格拉茨大学医学中心提供的 GRAZPEDWRI-DX 数据集<sup>[14]</sup>,该数据集涵盖了从 2008 年至 2018 年间 6091 名儿童腕部创伤的

X光检查图像,共20 327张。本文按照8:2的比例将数据集随机划分为训练集和验证集,训练集包含16 261张图像,验证集包含4 066张图像。此外,该数据集包含了9种类型的标签,具体标签情况如表2所示。

本文通过精确率( $P$ )、召回率( $R$ )、平均精度均值  $mAP@0.5$  (交并比为0.5时的平均精度均值)和  $mAP@0.5 \sim 0.95$  (交并比在0.5~0.95之间的平均精度均值)四个指标评价模型性能。

### 2.3 消融实验

为验证各改进模块在儿童腕部骨折检测中的有效性,将YOLOv11n作为基准模型,对C3k2-SCConv、MSDA和SSFF-X模块分别进行消融实验,结果如表3所示。

表2 数据集标签情况

Table 2 Label distribution of the dataset

异常情况	标签数量	图片数量
骨骼异常	276	192
骨损伤	45	42
异物	8	8
骨折	18 090	13 550
金属物	819	708
骨膜反应	3 453	2 235
旋前肌征	567	566
软组织	464	439
文字标注	23 722	20 274

表3 消融实验结果

Table 3 Results of ablation experiments

YOLOv11n	C3k2-SCConv	MSDA	SSFF-X	$P$	$R$	$mAP@0.5$	$mAP@0.5 \sim 0.95$
√				0.745	0.556	0.570	0.366
√	√			0.738	0.564	0.575	0.369
√		√		0.743	0.562	0.575	0.369
√			√	0.767	0.553	0.576	0.371
√	√	√		0.729	0.551	0.570	0.369
√	√		√	0.722	0.569	0.586	0.372
√		√	√	0.714	0.558	0.572	0.367
√	√	√	√	0.768	0.577	0.587	0.379

由表3可见,单独添加C3k2-SCConv模块后, $P$ 值略有下降, $R$ 值略有提升, $mAP@0.5$ 和 $mAP@0.5 \sim 0.95$ 均有所提升,表明该模块虽使精确率有所下降,但提升了模型的检测精度。C3k2-SCConv模块专注于小目标检测,单独使用时在复杂背景下会产生误判。单独添加MSDA模块后, $R$ 值提升, $mAP@0.5$ 和 $mAP@0.5 \sim 0.95$ 亦有提高,对模型检测能力有积极影响。单独添加SSFF-X模块后, $P$ 值升至0.767, $mAP@0.5$ 和 $mAP@0.5 \sim 0.95$ 同样上升,在精确度和平均精度均值方面表现良好。同时添加C3k2-SCConv和MSDA模块,整体检测精度有升有降。同时添加C3k2-SCConv和SSFF-X模块, $R$ 值提升显著, $mAP@0.5$ 和 $mAP@0.5 \sim 0.95$ 大幅上升,组合效果较好。MSDA模块的多尺度特征提取和SSFF-X模块的跨尺度特征融合,有助于模型更好地结合全局信息,增强特征的鲁棒性。但同时添加MSDA和SSFF-X模块,检测结果中平均精度均值提升不明显。本文采用三个模块都添加的方式, $P$ 、 $R$ 、 $mAP@0.5$ 和 $mAP@0.5 \sim 0.95$ 均达到较

高水平,整体检测性能最优。

### 2.4 不同位置嵌入SCConv的对比实验

为确定SCConv模块的最佳嵌入位置,在不同位置的C3k2模块中嵌入SCConv,进行对比实验,结果如表4所示。表4中位置1表示主干网络的第4个C3k2模块,位置2表示主干网络的所有C3k2模块,位置3表示模型全局范围的C3k2模块,位置4表示颈部网络的所有C3k2模块,即本文采用的嵌入位置。

表4 不同位置的C3k2模块中嵌入SCConv的实验结果

Table 4 Experimental results of the SCConv embedded C3k2 modules at different positions

嵌入位置	$P$	$R$	$mAP@0.5$	$mAP@0.5 \sim 0.95$
不嵌入	0.745	0.556	0.570	0.366
位置1	0.698	0.558	0.572	0.365
位置2	0.733	0.558	0.570	0.361
位置3	0.706	0.559	0.576	0.364
位置4(本文)	0.738	0.564	0.575	0.369

由表 4 中各项指标来看:在颈部网络的所有 C3k2 模块中嵌入 SCCConv 时,在提升模型性能方面表现最为突出,尤其在召回率以及不同交并比下的平均检测精度上具有明显优势;在主干网络的第 4 个 C3k2 模块位置嵌入 SCCConv 时,仅在部分指标上取得了不错的表现;在主干网络的所有 C3k2 模块中嵌入 SCCConv 时,虽然保留了一定的精度,但各项指标数值相对较低;在模型全局范围的 C3k2 模块中嵌入 SCCConv 时,整体综合性能稍差。综上,颈部网络的 C3k2 模块是最优嵌入位置。

## 2.5 不同注意力机制的对比实验

为验证 MSDA 的效果,分别引入空间和通道协同注意力(SCSA)<sup>[15]</sup>与混合局部通道注意力(MLCA)<sup>[16]</sup>机制进行对比实验,结果如表 5 所示。

由表 5 可以看出,添加不同的注意力机制后, $P$  值均有所下降, $R$  值变化较小,mAP@0.5 及 mAP@0.5~0.95 均有一定提升。综合对比,添加 MSDA 时算法在 mAP@0.5 和 mAP@0.5~

0.95 指标上表现最优,一定程度上提升了模型检测性能。

表 5 不同注意力机制的对比实验结果

Table 5 Comparative experimental results of different attention mechanisms

算法	$P$	$R$	mAP@0.5	mAP@0.5~0.95
YOLOv11n	0.745	0.556	0.570	0.366
YOLOv11n + SCSA	0.735	0.556	0.571	0.367
YOLOv11n + MLCA	0.722	0.554	0.570	0.362
YOLOv11n + MSDA(本文)	0.743	0.562	0.575	0.369

## 2.6 对比实验分析

为验证本文算法在儿童腕部骨折检测任务中的性能,将 MSEF-YOLO 算法与 YOLOv3-tiny、YOLOv5n、YOLOv6n、YOLOv8n、YOLOv9-s、YOLOv10n 以及 YOLOv11n 算法进行对比实验,结果如表 6 所示,表中 fracture 表示骨折类别。

表 6 不同算法的对比实验结果

Table 6 Comparative experimental results of different algorithms

算法	$P$	$R$	mAP@0.5	mAP@0.5~0.95	$P$ (fracture)	$R$ (fracture)	mAP@0.5(fracture)	mAP@0.5~0.95(fracture)
YOLOv3-tiny	0.678	0.530	0.584	0.333	0.852	0.889	0.939	0.532
YOLOv5n	0.737	0.529	0.557	0.343	0.855	0.891	0.930	0.542
YOLOv6n	0.728	0.526	0.547	0.339	0.850	0.887	0.925	0.539
YOLOv8n	0.750	0.554	0.566	0.348	0.868	0.902	0.938	0.551
YOLOv9-s	0.724	0.558	0.576	0.361	0.882	0.897	0.936	0.553
YOLOv10n	0.742	0.516	0.551	0.339	0.874	0.886	0.932	0.545
YOLOv11n	0.745	0.556	0.570	0.366	0.882	0.901	0.940	0.554
MSEF-YOLO	0.768	0.577	0.587	0.379	0.880	0.906	0.946	0.556

由表 6 可以看出,对于所有类别的目标检测,本文 MSEF-YOLO 算法在 GRAZPEDWRI-DX 数据集上检测的  $P$ 、 $R$ 、mAP@0.5 和 mAP@0.5~0.95 分别为 0.768、0.577、0.587 和 0.379,均高于其他主流算法。由骨折类别的检测结果可知:所有算法的  $P$  值均高于 0.850,其中 MSEF-YOLO 算法的  $P$  值为 0.880,预测准确性较好;MSEF-YOLO 算法的  $R$  值为 0.906,高于其他算法,有效减少了漏检;MSEF-YOLO 算法的 mAP@0.5 和 mAP@0.5~0.95 分别为 0.946 和 0.556,证明其在骨折检测方面的优势显著。在儿童腕部骨折检测任务中,MSEF-YOLO 算法在各项指标上均表

现较好,相较于其他 YOLO 系列算法,具有更高的召回率和检测精度。

## 2.7 检测结果

为验证算法的性能,使用 YOLOv11n 算法与优化改进后的 MSEF-YOLO 算法对验证集进行检测,部分检测结果如图 5 所示。从图中可以清晰看出:原 YOLOv11n 算法在骨折类(fracture)检测中存在漏检和误检的情况,且在小目标骨膜反应类(periostealreaction)检测中也存在漏检的情况;改进后的 MSEF-YOLO 算法能够有效减少漏检和误检情况,增强了对细微骨骼异常情况的检测能力,提高了检测效果。

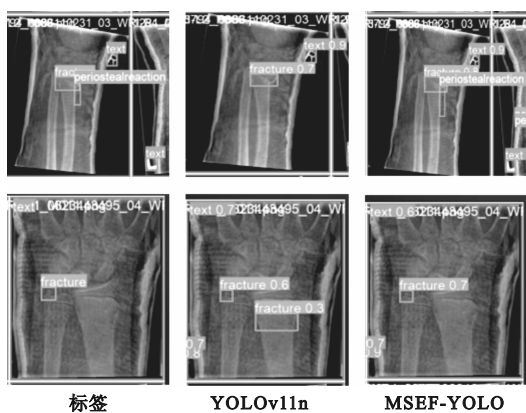


图5 部分检测结果对比

Fig. 5 Comparison of partial detection results

### 3 结论

针对儿童腕部骨折的检测问题,本文提出了基于 YOLOv11n 改进的 MSEF-YOLO 目标检测算法。C3k2-SCConv 模块利用空间和通道重构卷积,有效减少了模型参数量并提升了小目标检测能力;MSDA 机制通过采用不同的扩张率实现了多尺度特征提取;SSFF-X 模块借助 3D 卷积显著提升了多尺度特征融合能力。实验结果表明,本文提出 MSEF-YOLO 算法的精确率、召回率、平均精度均值( $mAP@0.5$  和  $mAP@0.5 \sim 0.95$ ) 相比于 YOLOv11n 分别提高了 3.1%、3.8%、3.0% 和 3.5%,能够有效辅助放射科医生对儿童腕部骨折进行检测,在医学影像分析领域具有良好的应用前景。

尽管 MSEF-YOLO 算法在本文实验数据集上的表现较为优异,但在临床应用中可能会面临一些挑战,例如不同成像设备的差异和多样化的拍摄角度等。针对这些问题,未来研究将通过数据增强和模型优化提升泛化性与鲁棒性,从而进一步提升模型的临床适用性,助力智能医疗发展。

### 参考文献 (References):

[1] THAIYALNAYAKI K, KAVYAA L, SUGUMAR J. Automated bone fracture detection using convolutional neural network [J]. Journal of Physics: Conference Series, 2023, 2471(1): 012003.

[2] THIAN Y L, LI Y T, JAGMOHAN P, et al. Convolutional neural networks for automated fracture detection and localization on wrist radiographs [J]. Radiology: Artificial Intelligence, 2019, 1(1): e180001.

[3] ZOU J T, ARSHAD M R. Detection of whole body bone fractures based on improved YOLOv7 [J]. Biomedical Signal Processing and Control, 2024, 91: 105995.

[4] WU H T, FU Q S. An innovative deep learning approach to spinal fracture detection in CT images [J]. Annali Italiani Di Chirurgia, 2024, 95(4): 657 - 668.

[5] 周涛, 杜玉虎, 茆晶晶, 等. 面向下颌骨骨折检测的主辅网络 YOLOv5 模型 [J]. 中国科技论文, 2023, 18(11): 1257 - 1266.

ZHOU T, DU Y H, MAO J J, et al. YOLOv5 model of primary and auxiliary networks for mandibular fracture detection [J]. China Sciencepaper, 2023, 18(11): 1257 - 1266. (in Chinese)

[6] ZHOU T, WANG H W, DU Y H, et al. M<sup>3</sup>YOLOv5: feature enhanced YOLOv5 model for mandibular fracture detection [J]. Computers in Biology and Medicine, 2024, 173: 108291.

[7] LIU X H, ZHOU Y H, DONG Z H, et al. LSDASCFormer: a transformer-like network with long-short-distance attention and SCConv for heart rate measurement [J]. IEEE Sensors Journal, 2024, 24(17): 28007 - 28017.

[8] 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.

[9] 王震洲, 杨榕, 宿景芳. 改进 YOLOv8 的烟盒缺陷检测算法 [J]. 电子测量技术, 2024, 47(13): 110 - 119.

WANG Z Z, YANG R, SU J F. Improved YOLOv8 cigarette box defect detection algorithm [J]. Electronic Measurement Technology, 2024, 47(13): 110 - 119. (in Chinese)

[10] 蒋仁杰, 袁贞明, 吴英飞. HR-SCNet 在儿童发育性髋关节发育不良诊断中的应用研究 [J]. 中国数字医学, 2024, 19(10): 1 - 7.

JIANG R J, YUAN Z M, WU Y F. Application of HR-SCNet in the diagnosis of developmental dysplasia of the hip in children [J]. China Digital Medicine, 2024, 19(10): 1 - 7. (in Chinese)

[11] 冯迎宾, 刘艾妮. 基于 YOLOv8n 的夜间车辆检测 [J]. 沈阳理工大学学报, 2025, 44(2): 1 - 6, 12.

FENG Y B, LIU A N. Night-time vehicle detection based on YOLOv8n [J]. Journal of Shenyang Ligong University, 2025, 44(2): 1 - 6, 12. (in Chinese)

[12] 周秦汉, 贾杰, 陈昊, 等. 基于 SKNet 改进 YOLOv5s 的无人机对道路小目标的检测 [J]. 南昌航空大学学报(自然科学版), 2023, 37(4): 39 - 45.

ZHOU Q H, JIA J, CHEN H, et al. Improving YOLOv5s based on SKNet for unmanned aerial vehicle detection of small road targets [J]. Journal of Nanchang Hangkong University (Natural Sciences), 2023, 37(4): 39 - 45. (in Chinese)

[13] 匡向阳, 景任杰, 董立红. 3D 密集全卷积的高光谱地物分类算法研究 [J]. 计算机工程与应用, 2023, 59(3): 112 - 117.

SHE X Y, JING R J, DONG L H. Research on hyperspectral ground object classification algorithm based on 3D dense full convolution [J]. Computer Engineering and Applications, 2023, 59(3): 112 - 117. (in Chinese)

[14] NAGY E, JANISCH M, HRŽIČF, et al. A pediatric wrist trauma X-ray dataset (GRAZPEDWRI-DX) for machine learning [J]. Scientific Data, 2022, 9(1): 222.

[15] WAN D H, LU R S, SHEN S Y, et al. Mixed local channel attention for object detection [J]. Engineering Applications of Artificial Intelligence, 2023, 123: 106442.

[16] WANG Y Q, ZHANG L Y, XIE J J, et al. DFEMF: fusing frequency correlation and mel features for robust edge bird audio detection [J]. Avian Research, 2025, 16(2): 100232.

(责任编辑: 宋颖韬)