

基于改进 YOLO11 的手腕骨折快速检测^①

张朋涛, 胡乃平

(青岛科技大学 信息科学与技术学院, 青岛 266061)
通信作者: 胡乃平, E-mail: hunaiping@qust.edu.cn



摘要: 针对急救场景下医学影像分析中手腕骨折检测精度不足、模型推理效率低等问题, 本文提出一种基于改进 YOLO11 的轻量化手腕骨折检测算法 MDM-YOLO. 首先设计多尺度特征提取 (multi-scale feature extraction module, MSFE) 模块, 通过多个并行分支提取不同尺度的信息, 解决复杂骨折形态的多尺度表征问题; 其次, 提出混合空间局部注意力 (mixed spatial and local attention, MSLA) 机制, 结合局部和全局特征, 显著提升了对细微骨折的关注程度; 最后, 设计动态深度可分离卷积 (dynamic depthwise separable convolution, DDSConv), 在保持检测精度的同时显著降低计算复杂度并加快推理速度, 使模型更加轻量化. 实验表明, MDM-YOLO 在 GRAZPEDWRI-DX 数据集上的精确率达到 92.6%, 召回率达到 88.1%, mAP50 达到 95.1%, 较原始模型提升 1.7%、2.5% 和 1.5%. 在相同的硬件环境下, 检测速度提升 37%, 参数量仅为原模型的 73.3%, 验证了轻量化设计的有效性. 为应急场景下的快速手腕骨折诊断提供了高效解决方案.

关键词: 手腕骨折检测; YOLO11; 多尺度特征; 注意力机制; 轻量化卷积

引用格式: 张朋涛, 胡乃平. 基于改进 YOLO11 的手腕骨折快速检测. 计算机系统应用. <http://www.c-s-a.org.cn/1003-3254/10085.html>

Rapid Wrist Fracture Detection Based on Improved YOLO11

ZHANG Peng-Tao, HU Nai-Ping

(School of Information Science and Technology, Qingdao University of Science and Technology, Qingdao 266061, China)

Abstract: To address the issues of insufficient detection accuracy and low model inference efficiency in wrist fracture detection for medical image analysis in emergency scenarios, this study proposes a lightweight wrist fracture detection algorithm based on the improved YOLO11, named MDM-YOLO. First, a multi-scale feature extraction module is designed to address the multi-scale representation challenges of complex fracture shapes by extracting information at different scales through multiple parallel branches. Second, a mixed spatial local attention mechanism is proposed, which combines local and global features to significantly enhance the attention to subtle fractures. Finally, a dynamic depthwise separable convolution (DDSConv) is designed to reduce computational complexity and accelerate inference speed while maintaining detection accuracy, thus making the model more lightweight. Experiments show that MDM-YOLO achieves an precision of 92.6%, a recall rate of 88.1%, and an mAP50 of 95.1% on the GRAZPEDWRI-DX dataset, representing improvements of 1.7%, 2.5%, and 1.5%, respectively, compared with the original model. Under the same hardware conditions, the detection speed increases by 37%, and the number of parameters is only 73.3% of the original model, verifying the effectiveness of the lightweight design. This provides an efficient solution for rapid wrist fracture diagnosis in emergency scenarios.

Key words: wrist fracture detection; YOLO11; multi-scale feature; attention mechanism; lightweight convolution

^① 收稿时间: 2025-08-12; 修改时间: 2025-09-04; 采用时间: 2025-09-26; csa 在线出版时间: 2026-01-08

手腕骨折是临床实践中最常见的创伤类型之一, 约占所有骨折的 20%^[1]. 在急救场景下, 实现快速、准确的骨折诊断对后续治疗决策至关重要. 目前, 诊断主要依赖 X 光和 CT 扫描等影像学检查. 本研究选用 X 光影像开展手腕骨折检测, 因其二维成像足以达到快速识别关键骨折信息的要求, 同时兼具数据易获取、成本低和处理复杂度低的优势, 与提升检测效率的目标高度契合. 传统诊断依赖医生人工审阅影像, 逐个定位并标注骨折部位, 以此制定治疗方案^[2]. 研究显示, 骨折误诊在急诊科所有误诊案例中占比达到 24%^[3], 这凸显了提升诊断效率与准确性的迫切临床需求与现有方法能力之间的显著差距.

深度学习驱动的计算机辅助诊断 (CAD) 系统可显著提高诊断准确率, 具有重要应用价值^[4]. 针对手腕骨折检测的时效性需求, 单阶段目标检测模型凭借高效推理特性成为快速检测的核心技术, 并在相关任务中得到实证支持. Ahmed 等^[5]研究表明, YOLO 系列单阶段模型在手腕骨折检测的关键性能指标上表现出色, 为高性能诊断工具的构建提供了坚实基础. Ju 等^[6]验证了 YOLOv8 在手腕骨折检测中的有效性, 并开发了原型系统, 为后续研究奠定了实践基础. Chien 等^[7]将注意力机制集成到 YOLOv8 框架中, 提出了 YOLOv8-AM 模型. 该模型引入注意力权重增强对细微骨折特征的特征, 显著提升了对潜在手腕骨折病灶区域的聚焦与特征提取能力. 为满足快速诊断需求, Ju 等^[8]提出了 FCE-YOLOv8 模型, 通过引入特征上下文增强模块, 显著降低推理时间, 加速手腕骨折诊断流程. Ferdi^[9]提出了 G-YOLOv11 模型, 采用幽灵卷积技术, 减少计算复杂度和参数量, 提升推理速度, 为快速手腕骨折 CAD 系统提供了技术路径.

在近期研究中, 胡远江等^[10]使用多核并行卷积模块 PKIModule 及上下文锚点注意力 CAA 对模型骨干中的 C3K2 进行重新设计, 以增强模型对遮挡目标及不同尺度目标的特征提取效果. 刘巧等^[11]设计了一种协同架构, 将双向特征金字塔网络 BiFPN 与全局-局部自注意力机制 GLSA 相结合, 有效捕获全局热分布模式与局部细节特征. 彭哲轩等^[12]将 YOLO11 的 C3K2 模块与 HetConv 模块融合构建 C3K2-H 模块, 实现了在保持检测精度的同时降低计算量. 这些研究为急救场景下手腕骨折快速检测提供了有益思路.

手腕骨折的类型复杂多样, 这种多样性不仅体现

在骨折的形态上, 还涉及骨折的部位和骨折程度上. 其次, 细微骨折征象往往在影像学检查中表现较为隐匿, 模型检测时可能会遗漏一些关键的细微骨折线索, 无法及时、准确地识别出所有潜在的细微骨折. 在急救场景下, 对诊断模型的检测速度和轻量化提出了更高的要求, 需要在较短的时间内完成对腕部骨折的准确判断, 以便及时为患者提供治疗建议. 针对上述挑战, 本研究提出一种基于改进 YOLO11 的轻量化手腕骨折检测模型 MDM-YOLO. 该模型设计旨在有效克服现有方法的不足, 优化快速检测性能.

本文主要贡献如下.

(1) 提出多尺度特征提取模块, 增强模型对于多尺度特征的捕捉能力, 有效突破了传统单一尺度特征提取存在的局限性.

(2) 提出混合空间局部注意力机制, 巧妙地融合局部与全局特征以及通道和空间信息, 进而提升了模型对于细微骨折的关注程度.

(3) 设计动态深度可分离卷积, 在大幅削减模型参数量与计算复杂度的基础上, 提高了模型的推理速度, 使其能够更加高效地契合对实时性要求较高的应用场景.

1 MDM-YOLO 模型

1.1 YOLO11 模型

YOLO11 是由 Ultralytics 团队推出的单阶段目标检测模型, 继承了 YOLOv8 的快速检测优势, 并通过模块重构、多尺度特征增强与计算效率优化, 显著提升了目标检测的精度与速度平衡性. 在架构创新方面, YOLO11 的骨干网络引入了 C3K2 模块替代 C2f, 通过可定制卷积核尺寸与通道分离策略, 提升复杂特征提取能力. 同时, C2PSA 模块集成位置敏感注意力, 结合多头注意力机制与前馈网络, 强化空间-通道特征交互. 颈部网络采用动态特征融合机制, 通过跨阶段部分融合策略拆分特征流, 降低冗余计算. 头部网络则采用深度可分离卷积替代标准卷积, 减少参数量同时维持分类精度, 并引入 PIoU 损失优化定位精度.

1.2 模型选择

YOLO11 系列模型提供多种尺寸的检测算法, 以适应不同场景对精度和计算资源的需求. 如表 1 所示, 较大的模型精度更高, 但参数量和计算复杂度随之增加, 推理速度慢, 不适合急救场景. 因此, 本文选择 YOLO11s 作为基线模型, 以平衡检测性能、内存需求和检测速度.

表 1 不同 YOLO11 模型实验参数对照

Model	Depth	Width	mAP50 (%)	Params (M)	FPS (f/s)
YOLO11n	0.33	0.25	85.4	2.4	141.8
YOLO11s	0.33	0.50	93.6	3.0	101.4
YOLO11m	0.67	0.75	93.8	20.2	83.4
YOLO11l	1.00	1.00	94.4	25.2	60.7
YOLO11x	1.00	1.25	94.0	56.8	47.6

1.3 MDM-YOLO 模型概述

如图 1 所示, 为了有效捕捉手腕骨骼复杂的微观结构特征, 本文在主干网络中设计了多尺度特征提取模块 (MSFE) 以提升模型对骨折区域细微解剖结构的表征能力. 为实现轻量化目标, 本文在主干网络中广泛引入动态深度可分离卷积 (DDSCov) 替代标准卷积

模块, 显著提高了模型的推理速度. 此外, 为了增强模型对骨折局部区域的聚焦能力, 本文在颈部网络中引入混合空间局部注意力机制 (MSLA) 提升模型对细微骨折征象的感知精度.

1.4 多尺度特征提取模块

在手腕骨折的临床诊断中, 骨折特征常表现出多尺度与形态复杂的特点. 为应对这一挑战, 本研究提出了一种多尺度特征提取模块. 通过并行分支对输入特征图进行分解, 提取多尺度信息, 并对融合后的输出特征图进行整合, 构建融合特征图. MSFE 模块的结构如图 2 所示.

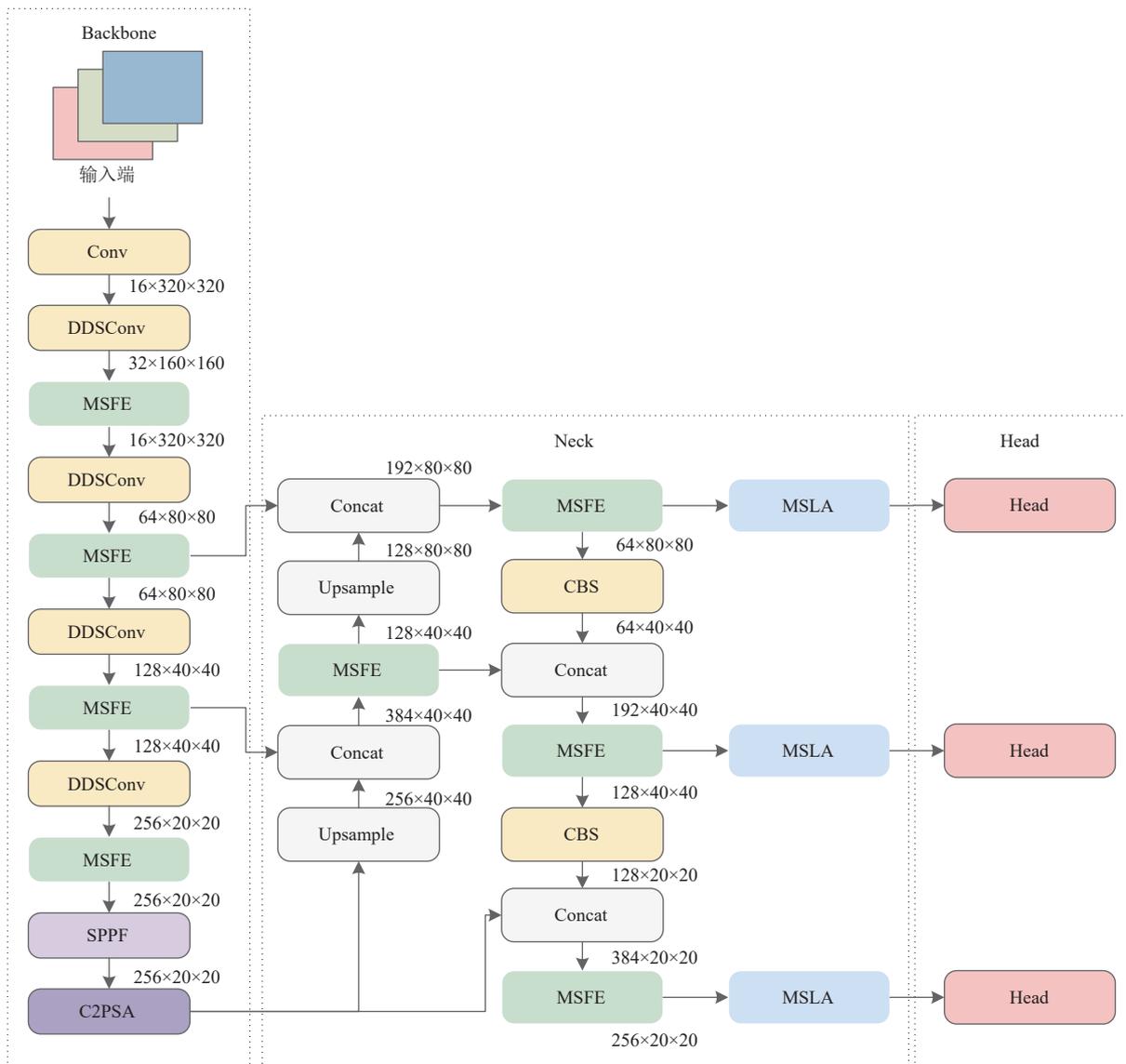


图 1 MDM-YOLO 结构图

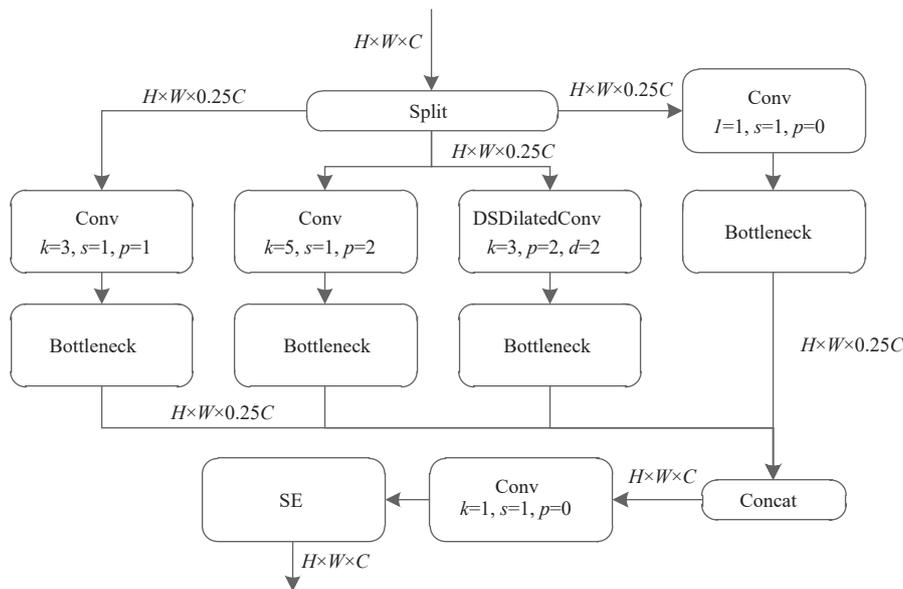


图2 多尺度特征提取模块 MSFE

MSFE 首先会按照通道维度将输入特征图划分为 4 个部分. 这一划分操作作为后续的特征提取奠定了基础, 使得每个部分都能专注于特定的特征信息. 接下来, MSFE 模块通过 4 条并行的路径分别对这 4 个部分进行处理. 每条路径都针对特定的尺度进行特征捕捉, 从而能够提取出对应尺度的特征信息. 这种设计使得模块能够同时处理多个尺度的特征, 极大地提升了特征提取的效率和准确性.

在路径 1 中采用标准的 3×3 卷积核以实现局部特征的提取, 不仅能够有效捕捉图像中的边缘、纹理等细节信息, 还有助于敏锐地捕捉骨折区域局部结构的微小变化.

路径 2 则使用标准的 5×5 卷积核对中尺度特征进行提取. 与 3×3 卷积核相比, 5×5 卷积核具有更大的感受野, 能够捕获更为丰富和广泛的上下文信息, 从而在检测中等尺寸目标时增强特征表达的有效性与鲁棒性.

在路径 3 中, 我们引入了扩张率为 2 的 3×3 深度可分离空洞卷积. 与传统卷积局限于局部邻域运算不同, 深度可分离空洞卷积通过在卷积核元素之间插入固定间隔, 实现了更大范围的信息采集, 能够更加全面地捕获全局及长距离依赖信息, 同时逐点卷积负责进行通道融合显著降低了模型的参数量与计算复杂度.

路径 4 采用 1×1 卷积实现全局特征的提取及计算量的降低. 该卷积核能够在保持特征图空间分辨率不变的前提下调整通道数, 从而实现高效的通道间特征

融合和降维处理.

在本研究中, 为了进一步提升特征提取的效率和质量, 我们对多尺度特征提取模块进行了优化设计. 具体而言, 在每条并行路径的末端, 引入 Bottleneck 模块, 通过巧妙的结构设计, 在不显著增加计算复杂度的前提下, 能够有效地增强特征的表达能力和传递效率. 具体结构如图 3 所示, 该模块由 3 个卷积层组成: 首先, 1×1 卷积用于降低通道维度, 减少计算复杂度; 其次, 3×3 卷积在降维后的特征空间中进行特征提取; 最后, 通过 1×1 卷积将通道维度恢复至原始规模. 这种设计策略显著降低了模型参数量和计算成本, 同时保持了特征表达的丰富性和准确性.

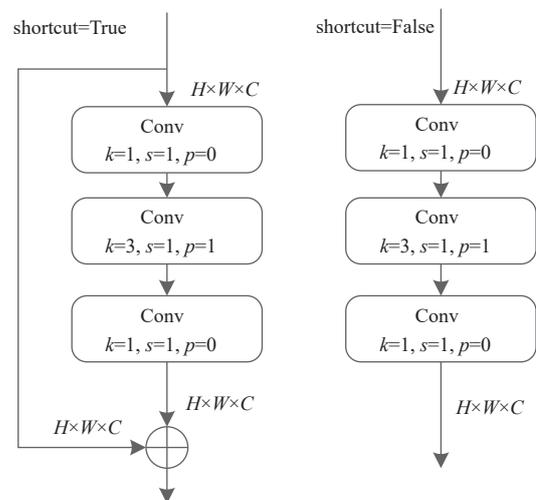


图3 Bottleneck 结构图

本文将4条路径的输出特征图沿通道维度进行拼接,在保持空间分辨率不变的前提下,融合不同尺度的特征.然而,拼接操作会导致通道数显著增加,进而加重后续计算负担.为此,本文在拼接后引入了 1×1 卷积层,用于调整通道数至预设的输出通道数 C .同时 1×1 卷积不仅能够高效地实现通道间的信息融合和降维处理,还能在不引入额外空间复杂度的情况下,增强特征

的非线性表达能力.

为进一步增强多分支模块中各子分支对不同尺度目标特征的选择性响应能力,本文在通道拼接后引入了SE(squeeze-and-excitation)注意力机制.通过对通道间的依赖关系进行建模,自适应地调整各通道特征的权重,从而提升模型对关键特征的关注度,增强其在复杂场景下的判别能力.SE结构如图4所示.

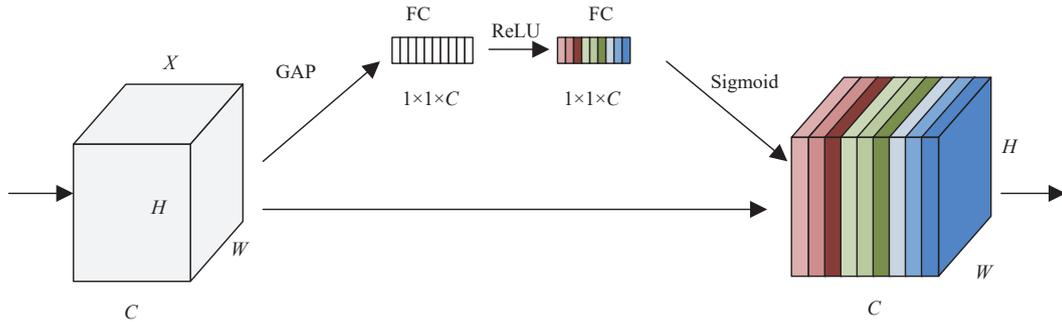


图4 SE注意力机制

首先,对输入的特征图($C \times W \times H$)进行压缩操作,将每个通道的信息压缩成一个单值,形成通道描述向量.公式表达:

$$S_c = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W X_{c(i,j)} \quad (1)$$

其中, $X_{c(i,j)}$ 代表输入特征图第 c 个通道上,位于 (i,j) 位置的像素值.

然后,对压缩操作输出的 $1 \times 1 \times C$ 特征图进行激励操作,学习通道间依赖关系,生成各通道的权重,并提升网络的非线性能力,通过 1×1 卷积降低通道数,使其减少计算量,同时压缩信息.

最后,通过缩放操作,采用加权平均的方式连接输出的特征图.经过SE模块,模型对目标的特征提取能力提升,同时摒弃不重要的特征.

通过上述融合策略,模型能够从多个尺度对同一图像进行综合分析,在捕捉复杂骨折特征的精准度上有所提升.同时,在高层次语义信息的提取与解析能力方面也获得了增强.

1.5 混合空间局部注意力机制

鉴于手腕骨折中普遍存在细微骨折特征,这一现象对精准诊断提出了更高要求.为增强模型对细微骨折特征的检测效能,本文在颈部网络中引入了混合空间局部注意力机制.该机制通过融合局部与全局特征

以及通道与空间信息,显著增强了模型对局部细节的关注能力,其结构如图5所示.

对于输入特征图首先被划分为多个子区域,每个子区域包含特征图的一部分通道.这种划分方式旨在对特征图进行分块处理,以便更高效地提取局部空间信息.对于每组特征图会通过局部平均池化(local average pooling, LAP),能够捕捉局部空间信息.

局部特征池化的计算公式为:

$$X_{loc}^{c,i,j} = \frac{1}{k_s^2} \sum_{m=0}^{k_s-1} X^{c,i+m,j+n} \quad (2)$$

其中, c 表示通道维度, i 和 j 表示空间位置维度.

为了充分提取特征图的多尺度信息并增强特征的表达能力,局部池化特征图通过两个独立的分支进行处理.分别采用不同的策略来捕捉特征图的全局特征和局部池化特征,从而实现了对特征图的多角度分析和特征提取.

首先通过全局平均池化(global average pooling, GAP)和全局最大池化(global max pooling, GMP)全面地提取特征图的全局信息.

GAP通过对特征图的每个通道进行全局平均操作,能够捕捉到特征图的整体平均信息,而GMP则通过提取每个通道的最大值,捕捉到特征图中最显著的特征信息.对于GAP和GMP处理后得到的特征图,通

过卷积层和 Sigmoid 激活函数进行处理, 生成空间注意力特征图得到全局特征 Y_{glb} .

然后通过特征转换生成局部池化特征, 先将其转换成维度为 $C \times k_s^2$ 的张量, 然后通过 1D 卷积进行特征

转换, 得到局部特征转换后的张量 Y_{loc} , 其计算方式为:

$$Y_{\text{loc}}^{c,m} = \sum_{n=0}^{k-1} K_{\text{loc}}^{c,n} \cdot Y_{\text{loc}}^{c,m+n} \quad (3)$$

其中, K_{loc} 表示局部分支的 1D 卷积.

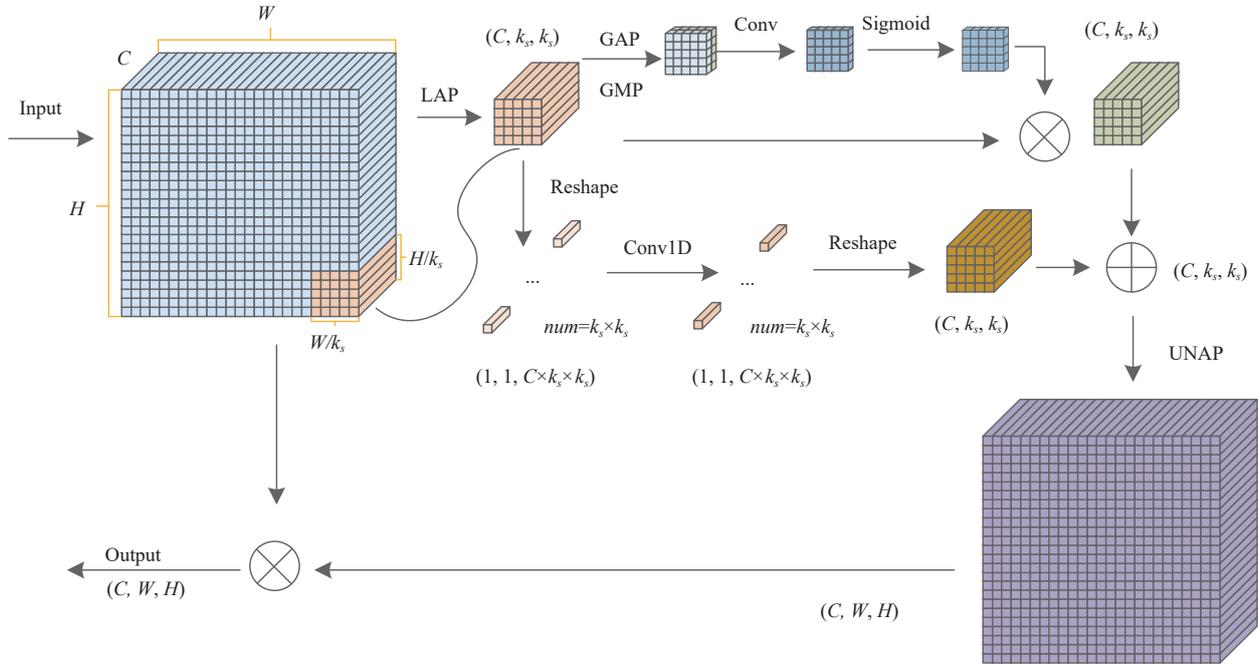


图 5 混合空间局部注意力 MSLA

通过加法操作将全局特征与局部池化特征相结合, 这个步骤在特征图中融合了全局上下文信息. 随后再通过反池化 (UNAP) 操作, 恢复到原始的空间维度得到注意力特征图 A .

UNAP 通过插值和上采样将特征图的空间分辨率恢复到与输入特征图一致, 同时保留注意力权重的精度. 最后将原始输入特征图 X 与注意力特征图 A 进行逐元素相乘, 得到增强后的特征图.

通过引入混合空间局部注意力机制, 提升模型对局部关键信息的捕捉能力, 有效降低细微骨折漏检的风险.

1.6 动态深度可分离卷积

为契合急救场景下手腕骨折快速检测的临床需求, 本研究提出了一种动态深度可分离卷积 (DDSCConv), 在保持检测精度的同时, 进一步降低计算复杂度并加快推理速度. DDSCConv 结构如图 6 所示.

DDSCConv 模块利用深度可分离卷积^[13]的思想, 将标准卷积分解为两个连续的操作. 本文在深度可分离卷积的基础上引入了动态路由机制. 通过全局上下文感知生成权重, 动态调制逐点卷积输出. 其过程包括:

利用 GAP 提取通道描述符, 通过轻量级全连接层和非线性激活函数生成权重, 最后将权重与逐点卷积输出相乘, 实现特征增强与抑制.

具体来说, 对于输入特征图 (C, H, W) , 深度卷积为每个通道配备一个独立的卷积核进行卷积运算. 对于某一通道 c 的输出特征图 y_c 在空间位置 (i, j) 的计算公式可以写为:

$$y_{c,i,j} = \sum_{u=0}^{K-1} \sum_{v=0}^{K-1} w_{c,u,v} \cdot x_{c,i+u-P,j+v-P} \quad (4)$$

其中, K 为卷积核的尺寸, P 为填充大小.

从参数量上看, 标准卷积的参数量为 $C_{\text{in}} \times C_{\text{out}} \times K \times K$, 而深度卷积的参数量仅为 $C \times K \times K$, 这大幅降低了计算复杂度和内存占用.

然后对深度卷积后的特征图进行逐点卷积, 使用 1×1 的卷积核以融合通道之间的信息. 对于输入特征图 X 和权重矩阵 W_{pw} , 输出特征图 Y 的计算公式为:

$$Y = W_{\text{pw}} \times X \quad (5)$$

其中, W_{pw} 的尺寸为 $C \times C_1$, C_1 为输出通道数.

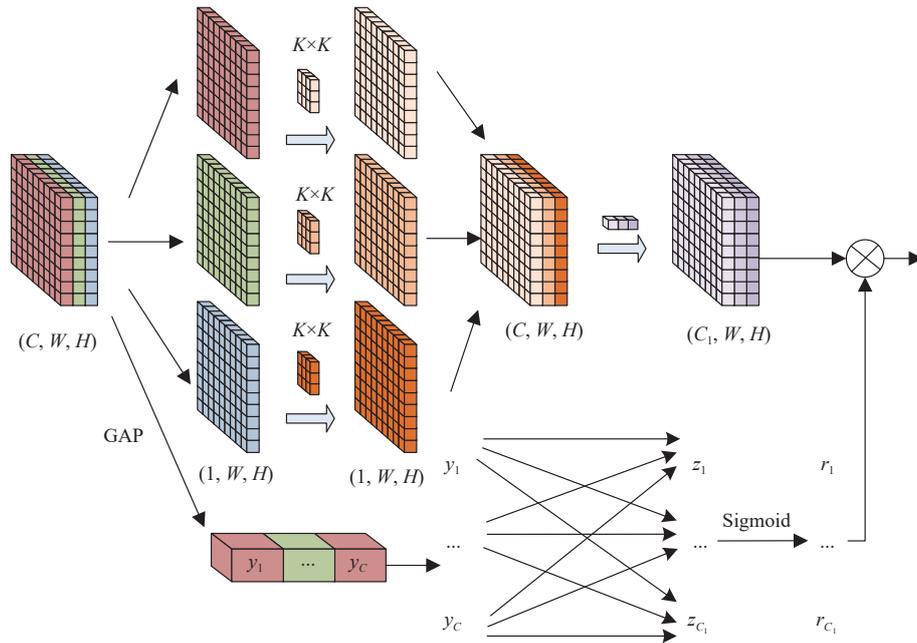


图6 动态深度可分离卷积 DDSCONV

引入动态路由机制对输入特征图进行权重计算,并对深度可分离卷积进行动态加权处理。

具体而言,对输入特征图的每个通道在空间维度上计算平均值,从而将一个二维的特征图压缩为一个标量.这一过程可以用式(6)来描述:

$$y_c = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W x_{c,i,j} \quad (6)$$

其中, $x_{c,i,j}$ 表示通道 c 在位置 (i, j) 的特征值. H 和 W 分别为该特征图的高度和宽度, y_c 则是经过全局平均池化后,通道 c 得到的输出。

将全局平均池化得到的特征向量输入到一个全连接层以生成动态路由权重,其中全连接层输入维度为 C ,输出维度为 C_1 ,该层将特征向量映射到多个不同的卷积核组.全连接层的权重矩阵为 W ,路由权重 r 的计算公式为:

$$r = \sigma(W \cdot GAP) \quad (7)$$

其中, σ 表示 Sigmoid 激活函数, W 的尺寸为不同的卷积核组的数量.经过路由层后,得到的 r 表示每个卷积核组的权重。

将逐点卷积的输出与动态路由权重相乘,一方面模型能够有效抑制冗余计算,避免对不重要的特征进行不必要的处理,从而节省计算资源并提高网络的效

率.另一方面,它能够突出关键特征,增强对重要信息的提取能力,使得网络在处理复杂的输入数据时能够更加精准地捕捉到对任务最有价值的特征,进而提升模型的整体性能。

2 实验结果

2.1 数据集

GRAZPEDWRI-DX^[14]由奥地利格拉茨医科大学发布,是目前公开可用的规模较大、标注详尽的手腕创伤 X 光影像数据集.该数据集汇总了 2008–2018 年间 6091 名患者的 10643 次检查,共 20327 张影像.尽管官方标注包含骨折、骨病变、金属植入物和软组织肿胀等标签,但手腕骨折样本的数量和质量显著优于其他创伤亚型.其余类别(如软组织损伤、骨病变)样本量少、分布稀疏,部分标签存在歧义或缺失.鉴于研究目标聚焦于手腕骨折检测,为确保训练数据的均衡性和可靠性,本文仅选用标注为手腕骨折的影像及其对应标签作为模型训练与评估的数据源.按照 7:2:1 的比例将数据集划分为训练集、验证集和测试集。

为了全面评估所提模型的泛化能力,本文引入了 FracAtlas^[15]数据集进行验证. FracAtlas 数据集是一个专注于肌肉骨骼放射影像的骨折分类、定位和分割的数据集,由 Abedeen 等人^[15]于 2023 年创建.该数据集的

构建基于从孟加拉国中3家主要医院收集的14068张X光图像,经过严格筛选和处理后,最终保留了4083张与手、肩、髋、腿等部位相关的高质量图像.其中,717张图像包含922个骨折实例,其余图像则为非骨折样本.所有图像的标注工作由两位放射科专家和一位骨科医生完成,确保了标注的准确性和可靠性.

2.2 类别不平衡处理策略

GRAZPEDWRI-DX数据集中骨折与非骨折样本数量存在显著不平衡.为了缓解此类不平衡可能导致

的模型偏见问题,本文采用结合针对性数据增强与定向过采样的组合策略.具体而言,针对数据集中稀缺的非骨折样本,对原始的非骨折训练集样本应用了一系列鲁棒性强且符合医学图像特性的空间与像素级增强变换,如图7所示.包括空间变换:随机旋转($\pm 15^\circ$)、随机平移($\pm 10\%$)、随机缩放(0.9–1.1倍).像素变换:随机调整图像对比度(因子范围[0.8, 1.2]).这些变换在每次训练迭代中随机组合,显著增加了非骨折样本的表现多样性.

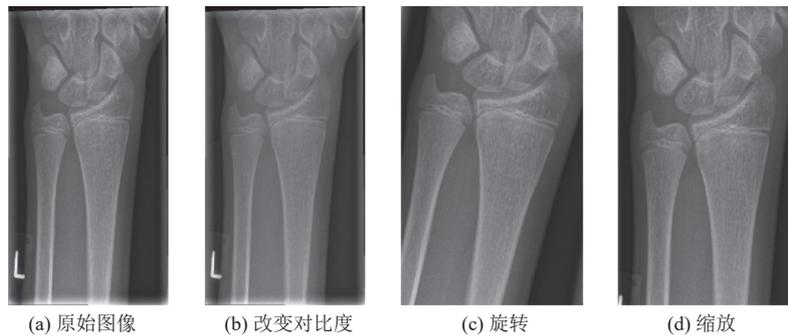


图7 数据增强示意图

在数据增强的基础上,本文进一步对增强后的非骨折样本池进行定向过采样,确保每个训练周期内,模型能够接触到足够多的非骨折样本,从而均衡地学习两类特征.

2.3 实验环境及参数配置

本文采用PyTorch深度学习框架搭建模型,实验环境如表2所示,训练参数如表3所示.为了提升模型的收敛速度与特征提取能力,本文采用MS COCO数据集上预训练的权重来初始化模型.该预训练策略使模型能够充分的利用在大规模数据集上学习到的通用特征表示,有效提升训练效率并在手腕骨折检测任务上取得更优的泛化性能.

表2 实验环境

硬件	参数
CPU	i5-12600KF
GPU	GeForce RTX 3070 12 GB
操作系统	Windows 11
Python版本	3.7.11
CUDA版本	12.0
PyTorch版本	1.9.1

2.4 实验评价指标

本文采用精确率 (precision, P)、召回率 (recall, R)

和平均精度均值 (mean average precision, mAP) 评价模型的检测性能.

表3 训练参数

参数名称	参数
批量 (batchsize)	8
优化器	SGD
迭代次数 (epoch)	100
输入图片尺寸	640×640
初始学习率	0.01
动量	0.937

精确率体现了模型对骨折的判断的可信度.召回率反映了模型对所有手腕骨折病例的覆盖能力.在手腕骨折检测中,高召回率至关重要,有助于减少漏检情况.平均精度综合考虑模型的精确率和召回率,提供对模型预测质量的全面评估.本文采用mAP50作为具体指标.

此外,本文引入参数量 (Params) 和检测速度 (FPS) 评估模型的参数大小和计算效率.其中,Params表示网络各层的参数之和;FPS则表示网络每秒检测的图像数量.

2.5 消融实验

为系统评估改进策略的有效性,本文设计并完成

了9组消融实验,实验结果如表4所示。

模型1作为基线,采用MS COCO数据集预训练权重初始化YOLO11s的全部骨干参数,以验证迁移学习对后续任务的影响。模型2则完全摒弃预训练策略,模型权重以随机初始化方式从零开始训练,其余参数、数据增强策略及训练周期均与模型1保持一致,

以排除干扰因素,从而精确量化预训练策略对最终检测性能的贡献。表4中“√”标记代表在对应改进点基础上进行的实验,旨在量化单独改进策略对整体性能的贡献。通过对比不同实验组与对照组之间的性能差异,本文能够准确地评估每种改进策略对模型性能的提升效果。

表4 改进模块在GRAZPEDWRI-DX数据集上的消融实验

Model	YOLO11s	预训练策略	MSFE	MLSA	DDSCConv	P (%)	R (%)	mAP50 (%)	Params (M)	FPS (f/s)
1	√	√	—	—	—	90.9	85.6	93.6	3.0	101.9
2	√	—	—	—	—	90.8	85.3	93.4	3.0	101.3
3	√	√	√	—	—	91.6	86.4	94.2	2.9	105.8
4	√	√	—	√	—	92.1	86.6	94.3	2.8	107.6
5	√	√	—	—	√	91.9	85.3	93.5	2.4	133.6
6	√	√	√	√	—	92.5	86.3	94.8	2.7	113.5
7	√	√	√	—	√	92.4	87.3	94.1	2.6	134.9
8	√	√	—	√	√	92.2	86.4	94.3	2.3	135.4
9	√	√	√	√	√	92.6	88.1	95.1	2.2	138.7

由表4可知,模型1借助预训练策略,迁移大规模数据通用表征,显著提升检测精度。模型3通过引入MSFE,使模型克服了单一尺度特征提取的局限性,模型的精确率和召回率分别提高了0.7%、0.8%,mAP50提高了0.6%,在提升识别率的同时,实现了检测效率的适度优化。在模型4中,通过在颈部网络引入MLSA,显著增强了模型对细节特征的关注能力,从而使模型精确率提高了1.2%,召回率提高了1.0%,在显著降低模型漏检的基础上,进一步优化了检测速度。模型5中,DDSCConv的引入显著降低了模型的参数量和计算复杂度,使推理速度提升至133.6 f/s,同时保持了较高的识别率和mAP50。在模型6中,MSFE和MLSA的协同作用使模型对手腕骨折特征的敏感度进一步提高,模型mAP50相较于YOLO11s提升了1.2%,推理速度达到113.5 f/s,实现了识别性能的稳步提升。模型8中,随着MLSA和DDSCConv的加入,模型在保持较高识别率的同时,推理速度进一步提升至135.4 f/s,体现了良好的速度-精度平衡。在模型9中,所有改进模块的综合应用充分彰显了MDM-YOLO模型在手腕骨折检测任务中的卓越性能。该模型的精确率和召回率分别达到92.6%和88.1%,mAP50达到95.1%,同时检测速度达到了138.7 f/s,成功实现了在保证识别率和漏检率的前提下显著提升检测速度的目标,达到了两者之间的有效平衡,为急救场景下手腕骨折的快速准确检测提供了有力支持。

为了直观地呈现MDM-YOLO模型的性能表现,本文以训练轮次(epoch)作为横坐标,以平均精度均值在50%阈值下的指标(mAP50)作为纵坐标,绘制了模型性能对比曲线图。如图8所示,该曲线图清晰地展示了MDM-YOLO模型在不同训练阶段的性能变化趋势,以及与基准模型的性能对比情况。

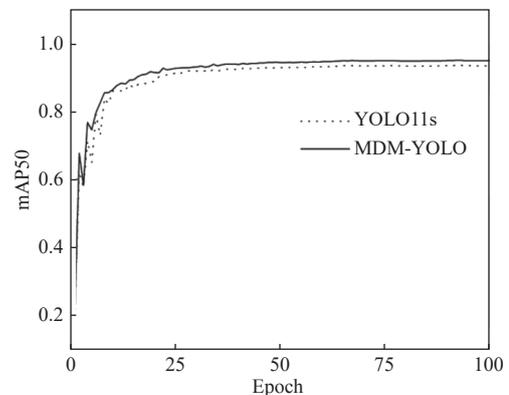


图8 模型对比曲线

此外,本研究亦呈现了模型的预测结果可视化图表,如图9所示。在该图中,模型能够精准地识别出目标对象,并准确地对其进行定位。在此基础上,模型进一步绘制了边界框并清晰地标注了目标对象的类别信息。这一过程充分展现了MDM-YOLO模型在手腕骨折目标检测任务中的高精度性以及复杂环境下的鲁棒性。通过将曲线图与预测结果图相结合,研究者可以从多个维度对MDM-YOLO模型的性能进行全面且深入的评估。

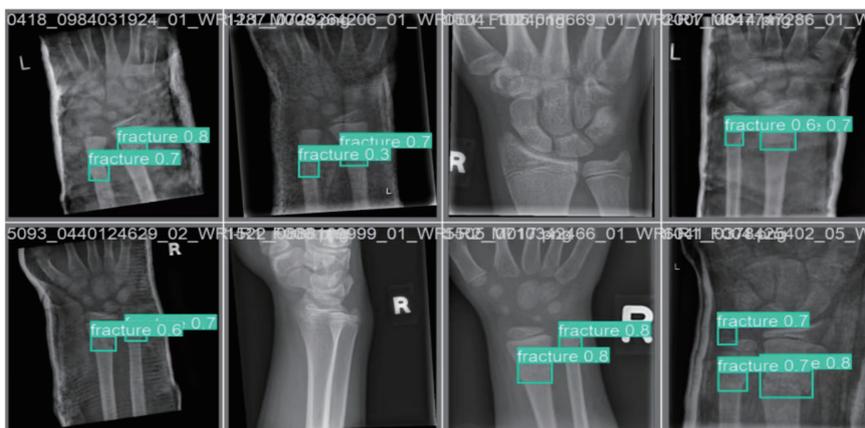


图9 MDM-YOLO 模型预测结果图

2.6 对比实验

2.6.1 不同检测算法对比实验

为了进一步验证所提算法的优越性,分别对比YOLOv5s^[16]、YOLOX^[17]、YOLOv8n^[18]、YOLOv10s^[19]、Hyper-YOLO^[20]、YOLO11s^[21]和RT-DETR^[22]这7种检测算法在GRAZPEDWRI-DX数据集检测效果,为了保证模型训练效果的一致性,所有对比实验均在硬件相同的环境下进行.实验结果如表5所示.

表5 不同检测算法的对比实验

Model	P (%)	R (%)	mAP50 (%)	Params (M)	FPS (f/s)
YOLOv5s	87.6	78.5	88.9	7.4	65.9
YOLOX	90.8	79.9	89.5	7.6	66.4
YOLOv8n	87.6	70.6	83.9	2.1	145.1
YOLOv10s	91.1	83.6	92.4	3.7	91.8
Hyper-YOLO	90.5	82.9	92.5	2.6	94.1
YOLO11s	90.9	85.6	93.6	3.0	101.4
RT-DETR	91.6	85.2	93.5	6.1	91.6
MDM-YOLO	92.6	88.1	95.1	2.2	138.7

由表5可知,在与其他多种检测算法的横向对比中,MDM-YOLO 彰显出优越的性能特质,其精确率召回率及mAP50指标与YOLOv5s、YOLOX、YOLOv10s 相比均有不同程度的提升,呈现出更精准、更高效的检测效能.与YOLOv8n 对比实验中,虽然MDM-YOLO 在检测速度指标上较YOLOv8n 稍有逊色,但其在精确率和召回率方面却大幅领先于YOLOv8n,显示出在检测精度上的突出表现.而从参数量的角度考量,尽管MDM-YOLO 与Hyper-YOLO 在参数量上较为接近,基本持平,但Hyper-YOLO 在检测精度和推理速度两个核心维度的表现却远远不及MDM-YOLO.综上所述,将MDM-YOLO 算法应用于应急场景下的手腕骨折检测任务时,能够凭借其在检测精度、推理速度

以及参数量等多方面的综合优势,取得更为理想、更为有效的检测成果,展现出其在该特定应用场景中的独特性能优势和实际应用价值.

2.6.2 样本平衡策略对比实验

为系统评估平衡样本对骨折检测性能的影响,本文引入3种样本平衡策略进行对比实验.采用常规数据增强、Focal Loss 损失函数加权以及组合策略.由表6可知,相同的实验条件下,采用组合策略取得了显著的性能提升.证实组合策略在缓解类别不平衡的同时可协同增强检测性能.

表6 不同数据平衡策略对比实验 (%)

Model	P	R	mAP50
YOLO11s	90.7	85.4	93.2
YOLO11s+数据增强	90.6	85.3	93.0
YOLO11s+Focal Loss	90.5	85.5	93.1
YOLO11s+组合策略	90.9	85.6	93.6

2.6.3 不同注意力机制对比实验

引入注意力机制的对比实验分析MSLA 注意力机制在手腕骨折数据集的检测效果,在颈部相同位置添加SE^[23]、CBAM^[24]、CA^[25]、ECA^[26]这4个注意力机制做对比.由表7可知,相同的实验条件下,添加MSLA 注意力机制的模型取得了显著的性能提升,精确率和召回率分别提升1.2%和1.0%,mAP50提升0.7%,优于当前其他主流的注意力机制.

2.7 模型泛化验证

为了验证模型的泛化能力,引入FracAtlas数据集作为泛化实验数据集.系统性地评估模型在不同数据环境下的鲁棒性和适用性.具体实验数据如表8所示.

具体而言,在FracAtlas数据集上,MDM-YOLO 的精确率和召回率分别达到了69.8%和54.9%,比原模

型提升了 5.1% 和 2.1%。mAP50 达到了 65.4%，同时检测速度达到 135.9 f/s，优于原模型的 89.1 f/s。实验结果表明，MDM-YOLO 在不同数据集上均取得显著的性能提升，体现了其较强的泛用性和适应性。

表 7 不同注意力机制对比实验 (%)

Model	P	R	mAP50
YOLO11s	90.9	85.6	93.6
YOLO11s+SE	91.3	85.6	93.7
YOLO11s+CBAM	90.8	85.7	93.9
YOLO11s+CA	89.9	85.4	92.9
YOLO11s+ECA	91.9	85.8	93.9
YOLO11s+MSLA	92.1	86.6	94.3

表 8 模型泛化实验

Model	P (%)	R (%)	mAP50 (%)	Params (M)	FPS (f/s)
YOLO11s	64.7	52.8	59.9	3.0	89.1
MDM-YOLO	69.8	54.9	65.4	2.2	135.9

3 结束语

本文提出的基于 MDM-YOLO 的手腕骨折检测算法在多方面实现改进，显著提升检测性能。多尺度特征提取模块有效解决复杂骨折特征信息丢失问题，增强多尺度目标特征捕捉能力。混合空间局部注意力机制缓解细微骨折漏检难题，提升检测精度。动态深度可分离卷积模块在保障性能的同时加快推理速度，满足应急场景需求。实验结果验证了算法优越性，其高精度与快速推理能力为应急环境下的快速诊断提供技术支撑。未来，本研究有望在更广泛的医学影像分析领域发挥重要作用，推动手腕骨折检测技术向更高效、更精准的方向发展。

参考文献

- Triwidodo A, Priambodo HN, Gani YI. Closed-reduction intramedullary screw in fractures of the proximal phalanges of the digits of the hand: A series of three cases. *Revista Brasileira de Ortopedia*, 2024, 59(S1): e34–e38.
- Lindsey R, Daluiski A, Chopra S, *et al.* Deep neural network improves fracture detection by clinicians. *Proceedings of the National Academy of Sciences of the United States of America*, 2018, 115(45): 11591–11596.
- 魏宗月, 仇大伟, 刘静, 等. 深度学习在上肢骨折诊断中的研究进展. *计算机科学与探索*, 2025, 19(9): 2341–2362.
- 吴仪, 沈思远, 梁锋, 等. 数字技术在早期膝关节炎诊疗中的应用. *中国临床研究*, 2025, 38(5): 681–684.
- Ahmed A, Imran AS, Manaf A, *et al.* Enhancing wrist

- fracture detection with YOLO. *arXiv:2407.12597*, 2024.
- Ju RY, Cai WM. Fracture detection in pediatric wrist trauma X-ray images using YOLOv8 algorithm. *Scientific Reports*, 2023, 13(1): 20077. [doi: [10.1038/s41598-023-47460-7](https://doi.org/10.1038/s41598-023-47460-7)]
- Chien CT, Ju RY, Chou KY, *et al.* YOLOv8-AM: YOLOv8 based on effective attention mechanisms for pediatric wrist fracture detection. *IEEE Access*, 2025, 13: 52461–52477. [doi: [10.1109/ACCESS.2025.3549839](https://doi.org/10.1109/ACCESS.2025.3549839)]
- Ju RY, Chien CT, Xieerke E, *et al.* Pediatric wrist fracture detection using feature context excitation modules in X-ray images. *arXiv:2410.01031*, 2024.
- Ferdi A. Lightweight G-YOLOv11: Advancing efficient fracture detection in pediatric wrist X-rays. *arXiv:2501.00647*, 2024.
- 胡远江, 杨翼, 吴湘宁, 等. 基于细节增强与多尺度特征融合的水下目标检测. *计算机系统应用*, 2025, 34(9): 69–78. [doi: [10.15888/j.cnki.csa.009925](https://doi.org/10.15888/j.cnki.csa.009925)]
- 刘巧, 章翔峰, 姜宏. 基于 HBGF-YOLO 的无人机航拍光伏板红外小目标缺陷检测. *激光与光电子学进展*, 2025: 1–22. <http://kns.cnki.net/kcms/detail/31.1690.TN.20250714.1547.074.html>. (2025-07-15).
- 彭哲轩, 应展烽, 葛昊, 等. ADE-YOLO: 基于改进 YOLO11 的退化环境下小目标检测算法. *计算机工程与应用*, 2025: 1–13. <http://kns.cnki.net/kcms/detail/11.2127.TP.20250704.1602.015.html>. (2025-07-04).
- Chollet F. Xception: Deep learning with depthwise separable convolutions. *Proceedings of the 2017 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Honolulu: IEEE, 2017. 1251–1258.
- Nagy E, Janisch M, Hrzić F, *et al.* A pediatric wrist trauma X-ray dataset (GRAZPEDWRI-DX) for machine learning. *Scientific Data*, 2022, 9(1): 222. [doi: [10.1038/s41597-022-01328-z](https://doi.org/10.1038/s41597-022-01328-z)]
- Abdeen I, Rahman A, Protyasha FZ, *et al.* FracAtlas: A dataset for fracture classification, localization and segmentation of musculoskeletal radiographs. *Scientific Data*, 2023, 10(1): 521. [doi: [10.1038/s41597-023-02432-4](https://doi.org/10.1038/s41597-023-02432-4)]
- Wang J, Yang P, Liu YS, *et al.* Research on improved YOLOv5 for low-light environment object detection. *Electronics*, 2023, 12(14): 3089. [doi: [10.3390/electronics12143089](https://doi.org/10.3390/electronics12143089)]
- Ge Z, Liu ST, Wang F, *et al.* YOLOX: Exceeding YOLO series in 2021. *arXiv:2107.08430*, 2024.
- Yaseen M. What is YOLOv8: An in-depth exploration of the internal features of the next-generation object detector. *arXiv:2408.15857*, 2024.

- 19 Wang A, Chen H, Liu LH, *et al.* YOLOv10: Real-time end-to-end object detection. arXiv:2405.14458, 2024.
- 20 Wu GN, Wu QH. Enhancing steel surface defect detection: A hyper-YOLO approach with ghost modules and hyper FPN. IAENG International Journal of Computer Science, 2024, 51(9): 1321–1330.
- 21 Khanam R, Hussain M. YOLOv11: An overview of the key architectural enhancements. arXiv:2410.17725, 2024.
- 22 Zhao Y, Lv WY, Xu SL, *et al.* DETRs beat YOLOs on real-time object detection. Proceedings of the 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle: IEEE, 2024. 16965–16974.
- 23 Hu J, Shen L, Sun G. Squeeze-and-excitation networks. Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018. 7132–7141.
- 24 Woo S, Park J, Lee JY, *et al.* CBAM: Convolutional block attention module. arXiv:1807.06521, 2018
- 25 Hou QB, Zhou DQ, Feng JS. Coordinate attention for efficient mobile network design. Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville: IEEE, 2021. 13713–13722.
- 26 Wang QL, Wu BG, Zhu PF, *et al.* ECA-Net: Efficient channel attention for deep convolutional neural networks. Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle: IEEE, 2020. 11534–11542.

(校对责编: 张重毅)