

光电工程

Opto-Electronic Engineering

中文核心期刊 中国科技核心期刊
Scopus CSCD

融合边缘特征与细节感知网络的YOLOv8s髌关节关键点检测算法

吕佳, 段训禄, 陈欣

引用本文:

吕佳, 段训禄, 陈欣. 融合边缘特征与细节感知网络的YOLOv8s髌关节关键点检测算法[J]. 光电工程, 2025, 52(3): 240281.

Lv J, Duan X L, Chen X. Edge feature and detail-aware network integrated YOLOv8s algorithm for hip joint keypoint detection[J]. *Opto-Electron Eng*, 2025, 52(3): 240281.

<https://doi.org/10.12086/oe.2025.240281>

收稿日期: 2024-12-02; 修改日期: 2025-02-24; 录用日期: 2025-02-25

相关论文

惯性稳定平台中的双观测器宽频扰动抑制技术

边启慧, 苗青青, 唐涛, 马浩统

光电工程 2025, 52(3): 240305 doi: 10.12086/oe.2025.240305

THI-YOLO: 改进YOLOv8的非机动车驾驶员头盔检测

孙光灵, 王薪博

光电工程 2024, 51(12): 240244 doi: 10.12086/oe.2024.240244

改进YOLOv8的路基边坡裂缝检测算法

牛晓富, 黄河, 张红民, 胥铁峰

光电工程 2024, 51(11): 240171 doi: 10.12086/oe.2024.240171

基于YOLOv8优化改进的太阳能电池片缺陷检测模型

彭自然, 王思远, 肖仲平

光电工程 2024, 51(11): 240220 doi: 10.12086/oe.2024.240220

更多相关论文见光电期刊集群网站 

 | 光电工程
Opto-Electronic Engineering

<http://cn.ojournal.org/oe>

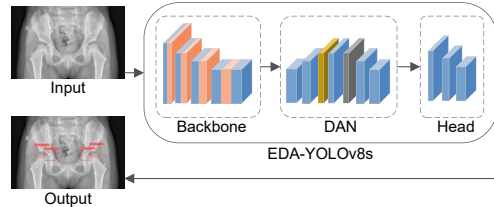


OE_Journal



Website

融合边缘特征与细节感知网络的 YOLOv8s 髋关节关键点检测算法



吕佳^{1,2}, 段训禄¹, 陈欣^{3*}

¹重庆师范大学计算机与信息科学学院, 重庆 401331;

²重庆师范大学重庆国家应用数学中心, 重庆 401331;

³重庆医科大学附属儿童医院国家儿童健康与疾病临床医学研究中心儿童发育与疾病教育部重点实验室, 重庆 400014

摘要: 髋关节关键点的准确识别对于提高发育性髋关节发育不良诊断精度具有重要意义。然而, 在儿童髋关节 X 射线图像中, 关键点所在的骨骼区域通常对比度低和边缘模糊, 导致边缘特征不明显。同时, 在特征提取过程中, 下采样操作会进一步弱化边缘信息。此外, 关键点邻域内的关键结构易受背景干扰, 这些因素均限制了关键点的精确定位。为此, 本文提出了一种融合边缘特征与细节感知网络的 YOLOv8s 髋关节关键点检测算法。该算法在网络中设计了边缘特征强化模块, 以捕获关键点周围空间信息并增强其所在的边缘特征; 同时, 提出细节感知网络, 对多层次特征进行融合与优化, 增强对图像中细微结构的感知能力。本文使用重庆医科大学附属儿童医院影像科提供的髋关节 X 射线图像数据集进行实验, 结果显示, 关键点的平均定位误差和平均角度误差降低至 4.2090 pixel 和 1.4872°, 相较于 YOLOv8s 降低了 6.8% 和 9.9%, 显著优于现有方法。实验证明, 本文算法有效提升了关键点的检测精度, 为临床诊断提供了重要参考。

关键词: 发育性髋关节发育不良; 关键点检测; YOLOv8s; 边缘特征强化; 细节感知网络

中图分类号: TP183

文献标志码: A

吕佳, 段训禄, 陈欣. 融合边缘特征与细节感知网络的 YOLOv8s 髋关节关键点检测算法 [J]. 光电工程, 2025, 52(3): 240281
Lv J, Duan X L, Chen X. Edge feature and detail-aware network integrated YOLOv8s algorithm for hip joint keypoint detection[J]. *Opto-Electron Eng*, 2025, 52(3): 240281

Edge feature and detail-aware network integrated YOLOv8s algorithm for hip joint keypoint detection

Lv Jia^{1,2}, Duan Xunlu¹, Chen Xin^{3*}

¹ College of Computer and Information Sciences, Chongqing Normal University, Chongqing 401331, China;

² National Center for Applied Mathematics in Chongqing, Chongqing Normal University, Chongqing 401331, China;

³ Department of Radiology, Children's Hospital of Chongqing Medical University, National Clinical Research Center for

收稿日期: 2024-12-02; 修回日期: 2025-02-24; 录用日期: 2025-02-25

基金项目: 国家自然科学基金重大项目 (11991024); 重庆市教委重点项目 (KJZD-K202200511); 重庆市自然科学基金创新发展联合基金重点项目 (CSTB2024NSCQ-LZX0090)

*通信作者: 陈欣, b2309@126.com。

版权所有©2025 中国科学院光电技术研究所

Child Health and Disorders, Ministry of Education Key Laboratory of Child Development and Disorders, Chongqing Key Laboratory of Pediatrics, Chongqing 400014, China

Abstract: The accurate identification of the hip joint keypoint is vital for diagnosing developmental dysplasia of the hip. However, in pediatric hip X-ray images, bone regions around key points often exhibit low contrast and blurred edges, resulting in unclear edge features. Furthermore, down-sampling operations during feature extraction further weaken edge information. Key structures surrounding the keypoint are highly susceptible to background interference. Such factors hinder the precise localization of key points. An edge feature and detail-aware integrated YOLOv8s algorithm was proposed for hip joint key point detection. The algorithm designs an edge feature enhancement module to capture spatial information around key points and strengthen edge features. A detail-aware network was designed to integrate and refine multi-level features, enhancing image perception of fine structures. Experiments used a hip X-ray dataset from the Department of Radiology, Children's Hospital of Chongqing Medical University. Results showed reductions in average keypoint localization and angular errors to 4.2090 pixel and 1.4872°, respectively. These reductions, which are 6.8% and 9.9% compared to those of YOLOv8s, highlight significant improvements in detection accuracy. The algorithm enhances keypoint detection precision and provides valuable support for clinical diagnosis.

Keywords: developmental dysplasia of the hip; keypoint detection; YOLOv8s; edge feature enhancement; detail-aware network

1 引言

发育性髋关节发育不良 (Developmental dysplasia of the hip, DDH) 是一种在新生儿中常见的骨科疾病, 发病率约 1.15%^[1]。若未及时诊断, 可能需要手术介入治疗并易引起手术并发症, 导致患者生活质量下降和医疗费用增加^[2]。因此, 尽早准确诊断 DDH 对患者至关重要。DDH 的临床诊断方法多样^[3-6], 常通过对髋关节 X 射线图像进行髋臼指数 (acetabular index, AI) 测量, 并结合患者年龄提供分析结果作为诊断参考依据。具体操作为手工标记图像盆骨结构中表现特殊的点, 并基于这些点的位置使用量角仪器进行相应角度测量^[7], 测量结果即 AI 值, 发育性髋关节发育不良诊断参考见图 1^[8], 其中 a 点为右髋臼上外侧缘, b 点为右侧三放射状软骨中心, c 点为左侧三放射状软骨中心, d 点为左髋臼上外侧缘。AI 值为评估髋关节的发育状况提供了关键依据, 故准确识别并正确标记关键点是诊断 DDH 的重要前提。

在临床诊断中, 手工标记髋关节关键点不仅要求医生具备扎实的专业知识和丰富的临床经验, 而且这一过程极其耗时和易受主观偏差影响。为应对上述挑战, 文献 [9] 使用边缘检测器并基于模板匹配技术识别髋关节关键点, 但这种方法缺乏对上下文信息的考虑, 在面对光照变化、遮挡和图像旋转等挑战时, 其鲁棒性和泛化能力不足。

近年来, 深度学习模型通过从大量医学图像数据中自动学习复杂的视觉特征和理解图像的上下文信息, 显著优化了图像识别的准确性和泛化能力^[10]。基于深度学习的 DDH 关键点检测早期算法^[8]是将识别目标从髋关节关键点转化为关键点的局部邻域, 通过学习这些邻域内的局部特征进行定位, 然后取局部邻域的中心点作为髋关节关键点。尽管这种方法能够提升关键点的定位精度和算法的鲁棒性, 但其存在噪声干扰和邻域信息利用不足等问题。为此, 一些研究者^[11-13]利用金字塔非局部模块、局部-非局部模块和双级路由注意力等方法, 增强对关键点邻域信息的关注, 并通过逐步缩小注意力聚焦区域以减少背景噪声干扰, 但这些方法因依赖深层特征进行注意力计算忽视了局部邻域内的细节边缘特征。另一些研究者^[14-15]则使用 (high-resolution network, HRNet) 网络和 Hourglass 网络, 从精细化特征层面来挖掘关键点邻域信息。然而, 这类方法在细化特征过程中, 因频繁上下采样加重了边缘信息的损失, 而骨骼边缘是辅助识别髋关节关键点的重要特征。此外, 在医学图像中, 许多关键结构在图像中占据很小的像素面积^[16], 注意力机制虽然在深层网络中表现出色, 但在具有丰富细节信息的浅层网络上大多因计算量大而难以应用, 这限制了其捕捉细微结构的能力。综上, 现有的工作未充分考虑关键点所在的骨骼边缘区域信息, 同时在识别局部邻

域内的关键结构特征方面存在不足, 导致难以进一步提升关键点的定位精度。

为解决上述问题, 本文提出一种融合边缘特征与细节感知网络的 YOLOv8s 髋关节关键点检测算法 (edge feature and detail-aware network integrated YOLOv8s algorithm for hip joint keypoint detection, EDA-YOLOv8s), 通过强化关键点所在骨骼边缘信息和增强网络对细微结构的捕捉能力, 以提升关键点的检测精度。本文主要贡献有以下三点:

1) 提出边缘特征强化模块 (edge feature enhancement module, EFEM), 用于捕获关键点周围空间特征, 并同步强化所在骨骼边缘区域的特征。

2) 提出细节感知网络 (detail-aware network, DAN), 通过对原有 Neck 网络进行改进, 整合不同尺度的特征信息, 提升网络对关键点局部邻域内重要细微结构的捕捉能力。

3) 本文算法在髋关节 X 射线图像数据集上进行了关键点检测的有效性验证。实验结果表明, 关键点的平均定位误差和平均角度误差分别降低至 4.2090 pixel 和 1.4872°, 显著提升了髋关节关键点检测精度, 为临床应用提供了更为可信的科学依据。

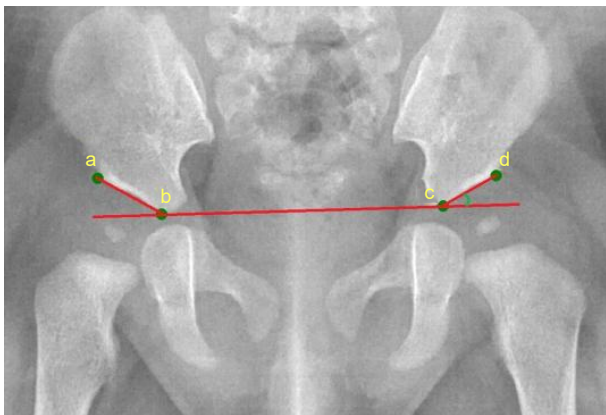


图 1 发育性髋关节发育不良诊断参考^[8]

Fig. 1 Diagnostic reference for developmental dysplasia of the hip^[8]

2 相关工作

2.1 DDH 关键点检测

髋关节发育在时间和空间上存在显著变化, 导致其病理状态的表现形式多样化^[8]。这种多样性增加了在髋关节图像中进行手工标记关键点的复杂性, 使得这一任务不仅对专业知识依赖性强, 而且耗时和易受主观性影响。为此, Al-Bashi 等人^[9]使用边缘检测器

捕捉骨骼轮廓并采用模板匹配方法定位髋关节关键点, 但传统的模板匹配方法缺乏对图像全局特征的语义理解。因此, Liu 等人^[8]基于深度学习技术将关键点的检测转换为关键点局部邻域检测, 利用邻域内的空间相关性捕捉关键点的空间分布特征, 进而识别关键点局部邻域。但该方法未能充分利用关键点的邻域信息。为此, Liu 等人^[11]使用金字塔非局部注意力模块, 增强网络对关键点邻域信息的关注, 不足的是网络在注意力交互过程中易受众多背景干扰。为降低复杂背景对局部邻域定位的制约, Wu 等人^[12]引入局部-非局部模块, 将注意力聚焦在盆骨区域矩形框内以减少背景带来的干扰, 但框内仍存在大量不利于局部邻域定位的背景噪声。为进一步减少背景噪声的不利影响, Lv 等人^[13]采用双级路由注意力, 将注意力计算限定在与关键点局部邻域内最相关的几个区域块, 并设计角度损失约束网络对局部邻域的选取。然而, 实施角度损失约束应以关键点的高精度定位为前提, 且上述方法过度依赖深层网络特征做注意力计算, 可能会忽略浅层的细节特征, 未关注到关键点周围细微结构特征对局部邻域定位的重要性。因此, Xu 等人^[14]使用 Mask R-CNN 分割髋关节 X 射线图像中的骨盆区域, 再通过 HRNet 网络专注于提取与关键点相关的细节特征, 实现对关键点局部邻域的定位, 但第二阶段局部邻域的准确定位受第一阶段分割质量的直接影响。为此, Xu 等人^[15]提出 Hourglass 网络, 通过在编码器中逐层下采样和在解码器中与低层网络融合特征后再上采样, 然后在解码器中恢复细节信息和细化特征, 从而提高关键点局部邻域的定位精度。然而这种方法因频繁采样使得边缘信息被丢失, 未考虑到关键点局部邻域内边缘线对于辅助识别骨骼边界中关键点的重要性。

2.2 边缘检测

边缘检测是计算机视觉领域的重要技术, 主要用于识别图像中亮度显著变化的区域, 从而定位物体边界。传统边缘检测算法通常先对图像进行预处理, 再应用 Sobel、Prewitt 和 Laplacian 等算子计算图像的梯度, 最后通过后处理进行标记和细化边缘^[17]。这种方法虽然在简单场景下效果良好, 但在复杂场景中难以准确识别边缘, 这源于其深度依赖局部像素变化检测边缘而忽视了图像的上下文信息^[18]。近年来, 深度学习模型通过学习图像上下文信息, 显著提高了处理复

杂图像的能力。因此, 基于传统的边缘检测算法逐渐向深度学习方法转变, Xie 等人^[19]提出了全嵌套边缘检测 (holistically-nested edge detection, HED) 网络, 通过摒弃传统边缘检测算子, 直接使用卷积神经网络 (convolutional neural network, CNN) 和深度监督策略, 对不同层级特征进行边缘提取, 并在每个侧输出层执行深度监督以指导特征学习, 逐步提炼出更准确的边缘信息。然而, 该方法通常对计算资源要求较高, 且主要关注于边缘优化而未考虑细节特征处理, 限制了在其他任务中的适用性。为提高边缘检测技术的可扩展性, 研究者开始探索将传统边缘检测技术与深度学习相结合的方法。Kong 等人^[20]在图像超分辨率重建任务中, 先使用 CNN 提取图像细节信息, 然后应用边缘检测技术添补图像的边缘信息, 提高了网络对图像细节和边缘信息的双重提取能力。由于髋关节 X 射线图像中关键点位于骨骼边界上, 故在现有目标检测网络中, 如何融合边缘检测技术来捕捉关键点相关的边缘区域特征, 可能成为提升关键点检测精度的重要因素。

2.3 YOLO 目标检测算法

YOLO^[21] 作为一阶段目标检测算法, 通过将目标检测任务转化为回归问题, 利用整张图像作为网络输入, 直接预测边界框的位置和目标的类别, 这使得该算法在处理速度上具有明显优势。随着技术的不断进步, 基于 YOLO 思想的检测算法已经发展至 YOLOv10。

YOLOv8 是现阶段常用的目标检测算法, 实现了目标检测在高效率和高精度之间的平衡。本文对 YOLOv8、YOLOv9 和 YOLOv10 检测算法及其不同规模的网络进行了对比实验。最终, 选择在本文数据集上综合性能表现最优的 YOLOv8s 作为本算法基础框架。该算法的网络由 Backbone 网络、Neck 网络和 Head 检测头三部分组成。其中, Backbone 负责图像基本特征信息的提取; Neck 采用特征金字塔网络 (feature pyramid network, FPN)^[22] 和路径聚合网络 (path aggregation network, PAN)^[23] 有效融合多尺度特征用以提升网络对目标物体的识别敏感度; Head 则采用解耦设计, 将分类和回归任务分开优化处理以产生最终预测结果。此外, 模型还引入分类损失和回归损失函数, 并采用动态策略平衡正负样本, 从而提高检测网络输出的一致性。

3 本文方法

3.1 网络结构

现有的 DDH 关键点检测算法通常是先将关键点检测转换为关键点的局部邻域检测, 再通过各种方法提升网络对局部邻域信息的利用, 但这类方法难以捕捉到能够辅助识别关键点的边缘信息和其局部邻域的细微结构信息, 导致关键点定位精度提升效果不佳。因此, 本文在此基础上通过增强边缘信息和提升局部邻域内细微结构的提取能力, 进而提出 EDA-YOLOv8s 算法。EDA-YOLOv8s 算法的整体架构如图 2 所示, 其网络结构由 Backbone 网络、Neck 网络和 Head 检测头三部分组成。在 Backbone 的四个 C2f 模块中均嵌入 EFEM 模块, 以增强关键点的骨骼边缘特征。而原有的 YOLOv8s 中的 Neck 部分用 DAN 网络替换, 旨在加强网络对关键点周围重要细微结构的提取能力。其中, DAN 网络先利用双尺度特征融合模块 (dual-scale fusion module, DSFM) 对不同尺寸大小的浅层特征图进行特征融合和优化; 然后, 将优化后的特征图与深层网络上采样后的特征图进行拼接处理; 最后, 将拼接后的特征图输入到跨阶段部分全核模块 (cross-stage partial omni-kernel module, CSP-OKM)^[24] 中, 以实现从语义信息到细节信息的全面整合。

3.2 EFEM 模块

由于髋关节 X 射线图像中关键点所在的骨骼边缘特征不明显, 且在特征提取过程中, 下采样操作会导致边缘信息逐渐被削弱^[25], 从而限制了网络对关键点的识别能力。文献 [20] 利用传统边缘检测技术在特征提取阶段结束后添补丢失的边缘信息, 但这种后处理方式可能会导致边缘信息与语义信息之间存在较大差距。受其启发, 本文设计了 EFEM 模块, 使网络在捕获关键点周围空间信息时, 同步强化骨骼边缘特征。此外, 在特征提取过程中多次应用该模块以逐步增强边缘特征, 引导网络聚焦于关键点的边缘区域。

EFEM 模块结构图如图 3 所示, 其采用的是三支结构: 边缘增强分支、普通卷积分支和残差分支, 尺寸大小为 $C \times H \times W$ 的特征图 f_m 送入三个分支后分别进行对应处理。具体处理过程为: 先在边缘增强分支中, 应用 Sobel 算子分别在水平方向和垂直方向进行边缘特征提取, 并将两种方向提取的结果相加, 得到增强边缘后的特征图 f_1 ; 同时在普通卷积分支中, 使用 3×3 的普通卷积捕获图像局部信息得到特征图 f_2 。

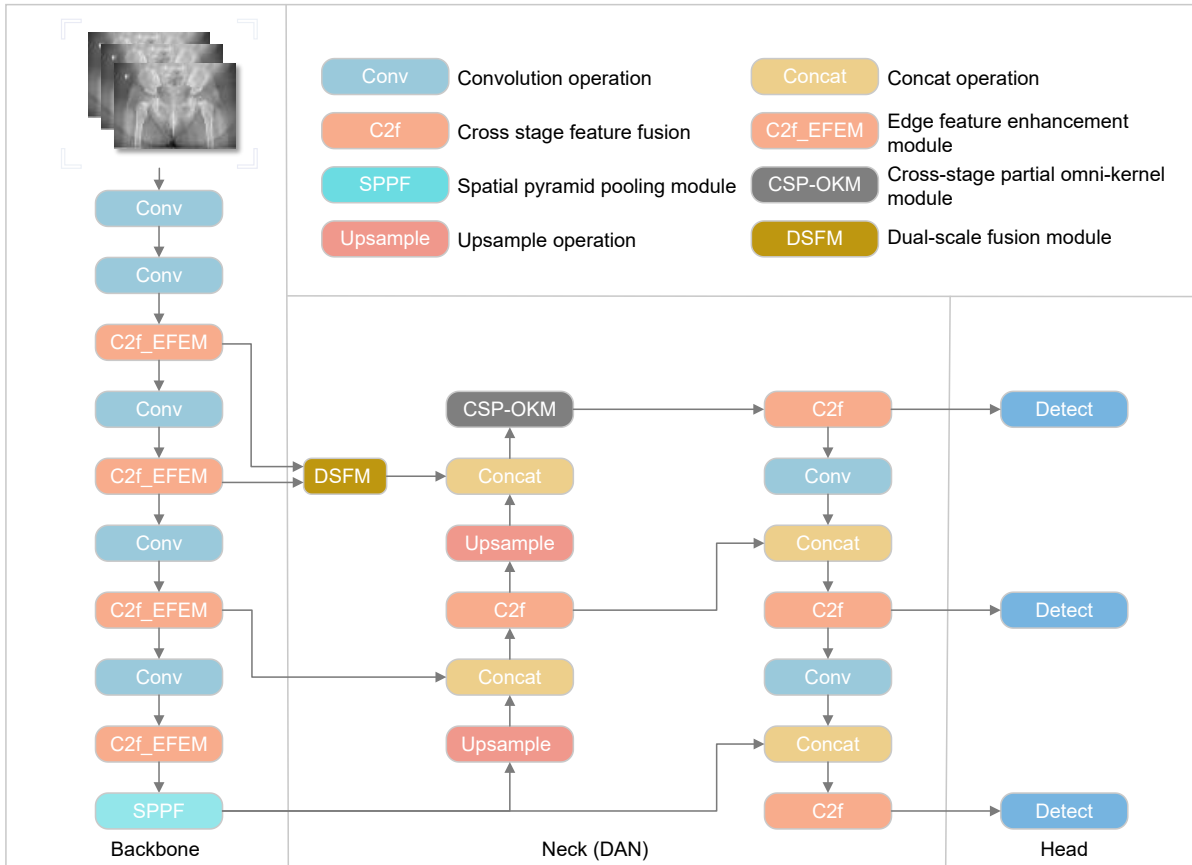


图 2 EDA-YOLOv8s 整体网络结构图

Fig. 2 Overall network structure diagram of EDA-YOLOv8s

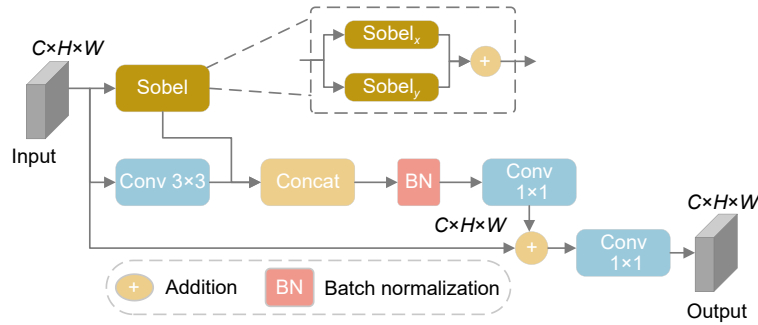


图 3 EFEM 模块结构图

Fig. 3 Structural diagram of EFEM module

此外, 考虑到边缘增强分支和普通卷积分支的输出特征图 f_1 和 f_2 在数值尺度上可能存在差异, 故将两者在通道维度上拼接后进行批量归一化 (batch normalization, BN)^[26] 处理, 以确保不同类型特征图的数值尺度一致性; 随后, 使用 1×1 普通卷积进一步调整将特征图通道数缩减一半, 得到特征图 f_3 。为补充原始特征信息, 模块在残差分支中将 f_{in} 和 f_3 进行加法操作, 并利用 1×1 普通卷积再次调整以生成更精细的特征表示, 得到目标特征图 f_{out} 。具体计算过程如下:

$$f_1 = \text{Sobel}_x(f_{in}) + \text{Sobel}_y(f_{in}), \quad (1)$$

$$f_2 = \text{Conv}_{3 \times 3}(f_{in}), \quad (2)$$

$$f_3 = \text{Conv}_{1 \times 1}(\text{BN}(\text{Concat}(f_1, f_2))), \quad (3)$$

$$f_{out} = \text{Conv}_{1 \times 1}(f_{in} + f_3). \quad (4)$$

通过融合普通卷积分支和边缘增强分支各自对图像的提取特征, EFEM 模块使得网络在提取图像局部特征时, 能够同时增强对边缘特征的敏感度, 从而为网络提供更为丰富和精确的特征信息。

3.3 DAN 网络

在膝关节 X 射线图像中, 关键点局部邻域内的细微关键结构能够辅助网络识别关键点。为了提升网络对细微结构的感知能力, 可以在 Neck 中, 采取优化特征融合策略或融入高分辨率特征图方法以实现目的, 图 4 展示了基于 YOLOv8s 的三种不同 Neck 网络结构。

在 YOLOv8s 的 Neck 中, 最高分辨率输入特征图 P3 的尺寸仅为 80 pixel × 80 pixel, 如图 4(a) 所示, 该层特征图每个像素点负责压缩并表达初始输入图像的 64 个像素区域信息。随着网络层级深入, 每个像素点需要表示的区域面积显著增加, 这种高负荷的信息压缩虽然有助于特征的整合, 但也制约了网络对图像中细微结构的提取和表达。为提升网络对细微结构的表达以及小目标的识别能力, 文献 [27-29] 在 YOLOv8s 中利用拥有丰富细节信息的 P2 特征图构建小目标检测层, 辅助网络提取图像中的细节特征以检测存在漏检的小目标, 如图 4(b) 所示。然而, P2 特征图也存在大量无关细节信息, 这类方法加大了计算量且未对其进行优化处理, 致使网络易受无关信息干扰。基于此, 本文在 Neck 基础上设计了 DAN 网络, 如图 4(c) 所示。DAN 先将 P2 和 P3 特征图共同输入到 DSFM 模块中执行特征融合和优化处理, 以更好的保留细微结构信息。随后, 再将其与深层网络上采样的特征图进行拼接后, 通过使用 CSP-OKM 模块将精细化的浅层细节信息与深层语义信息进行全面整合, 实现对关键细微结构的捕捉。

3.3.1 DSFM 模块

引入高分辨率的 P2 特征有助于提升网络对细节

特征识别能力。然而, 直接使用通过 P2 特征构建小目标检测层的方法^[27-29], 不仅会增加计算负担, 还可能引入无关细节。为降低计算量和优化细节特征, 本文设计了 DSFM 模块, 如图 5 所示。

为实现不同尺度特征图的有效融合, 首先 DSFM 模块会对 P2 特征图执行下采样操作。在下采样过程中, 采用文献 [30] 中的 Adown 模块对尺寸大小为 $C \times H \times W$ 的 P2 特征图 f_{in1} 进行下采样处理, 以便最大限度地保留图像的细节信息, 使得 P2 特征图尺寸缩小 1/2 而通道数保持不变; 同时, 对尺寸大小为 $2C \times H/2 \times W/2$ 的 P3 特征图 f_{in2} 使用 1×1 普通卷积进行调整, 实现通道数减半。然后, 在通道维度上将 f_{in1} 和 f_{in2} 对应的输出特征图进行拼接得到特征图 f_1 , 以丰富细节特征的多样性。为能够更好地捕获图像中的复杂结构, 采用文献 [31] 方法将 f_1 分别输入到三个深度可分离卷积分支中处理, 并在残差连接后进行加法操作得到特征图 f_2 , 以增强网络对细微结构的感知能力。最后, 使用 1×1 普通卷积对 f_2 进行特征平滑处理, 再与残差分支进行加法操作得到目标特征图 f_{out} 。具体操作过程为

$$f_1 = \text{Concat}(\text{Adown}(f_{in1}), \text{Conv}_{1 \times 1}(f_{in2})), \quad (5)$$

$$f_2 = f_1 + \sum_{i=1}^n \text{Dwconv}(k, f_1), \quad (6)$$

$$f_{out} = \text{Conv}_{1 \times 1}(f_2) + f_1, \quad (7)$$

式中: n 为深度可分离卷积分支数量, 本文设置为 3; k 为卷积核的大小, 公式中 $k=2i+1$; f_{in1} 为 P2 特征; f_{in2} 为 P3 特征; f_1 和 f_2 均为模块内中间层特征; f_{out} 为目标输出特征。

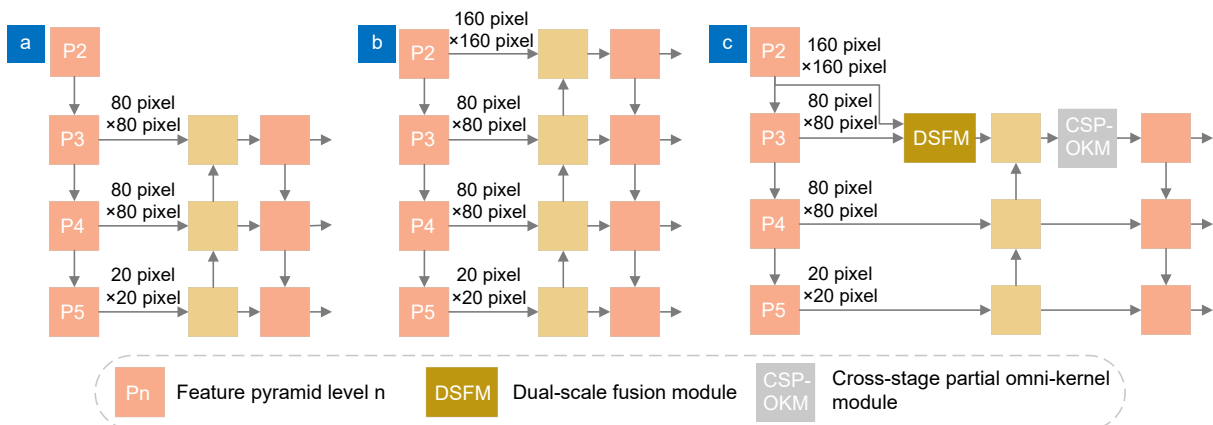


图 4 YOLOv8s 中三种 Neck 网络结构对比图。(a) Neck; (b) Small-neck; (c) DAN

Fig. 4 Comparison diagram of three neck network structures in YOLOv8s. (a) Neck; (b) Small-neck; (c) DAN

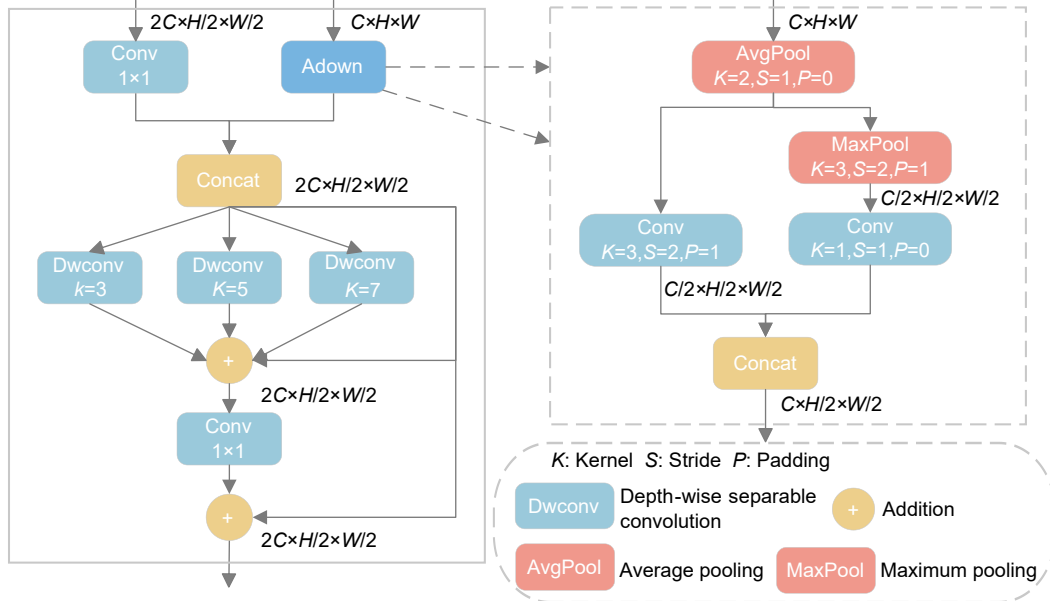


图 5 DSFM 模块结构图

Fig. 5 Structural diagram of DSFM module

在模块中引入高分辨率层级特征并进行特征融合和优化处理, 不仅能增强网络对图像局部特征的理解, 而且能提升对图像中细微结构的敏感度。

3.3.2 CSP-OKM 模块

在 Neck 中, DSFM 模块通过对 P2、P3 特征图执行双尺度特征融合, 提供了精细化的细节信息, 深层网络上采样后的特征图则提供了较强的语义信息^[32]。为对拼接后的两者特征图进行有效融合和减少计算量, 本文引入 OKM 模块, 并将其嵌入到跨阶段部分 (cross stage partial, CSP) 模块内部实现轻量化, 如图 6(a) 所示。OKM 由三个重要分支构成: 局部分支、大尺度分支和全局分支, 如图 6(d) 所示。

在 OKM 模块特征融合过程中, 首先, 输入特征图通过 1x1 普通卷积进行优化调整; 然后, 调整后的特征图分别被送入三个分支进行处理, 具体而言, 局部分支通过 1x1 深度卷积补充精细的局部信息。大尺度分支利用不同形状的大卷积核捕获更广泛的空间信息。全局分支则使用注意力机制获取全局信息, 如图 6(b) 和图 6(c) 所示, 通过基于频率的空间注意力模块 (frequency-based spatial attention module, FSAM) 和双域通道注意力模块 (dual-domain channel attention module, DCAM) 在不同维度上增强对关键点邻域内细微结构的关注, 以帮助网络在全局范围内理解关键点的位置和特征分布; 最后, 将三个分支的输出特征图进行加法操作, 并使用 1x1 普通卷积进行调整得到输

出特征图。OKM 模块通过多分支结构, 能够感知并提取关键点附近的重要结构信息, 这些信息有助于辅助网络进一步准确识别关键点。

4 实验结果与分析

4.1 实验环境与参数设置

本实验使用的 CPU 为 Intel® Core™ i5-13400F, GPU 为 NVIDIA GeForce RTX 3060 12 G, 操作系统为 Windows 11。使用的软件环境为: Python 3.8.0, Cuda 11.7, PyTorch 1.13.1 深度框架。网络训练轮次总数设置为 400 轮, 批量大小设置为 16, 采用优化算法 Adam 训练网络, 初始学习率为 0.01, 余弦退火参数为 0.2。

4.2 数据集采集

为确保算法在临床诊断中的有效性和实用性, 数据集采自于重庆医科大学附属儿童医院影像科, 且由该医院的资深儿科影像学医生对髋关节 X 射线图像进行关键点标注。本文收集了 1000 张婴儿髋关节 X 射线图像, 图像的分辨率为 2602 pixel x 1166 pixel。数据集按照训练集、验证集和测试集进行随机划分, 划分比例为 7 : 2 : 1, 训练时关键点局部邻域大小为 110 pixel x 110 pixel。

4.3 评价指标

关键点定位误差 (Point-to-point error for landmark,

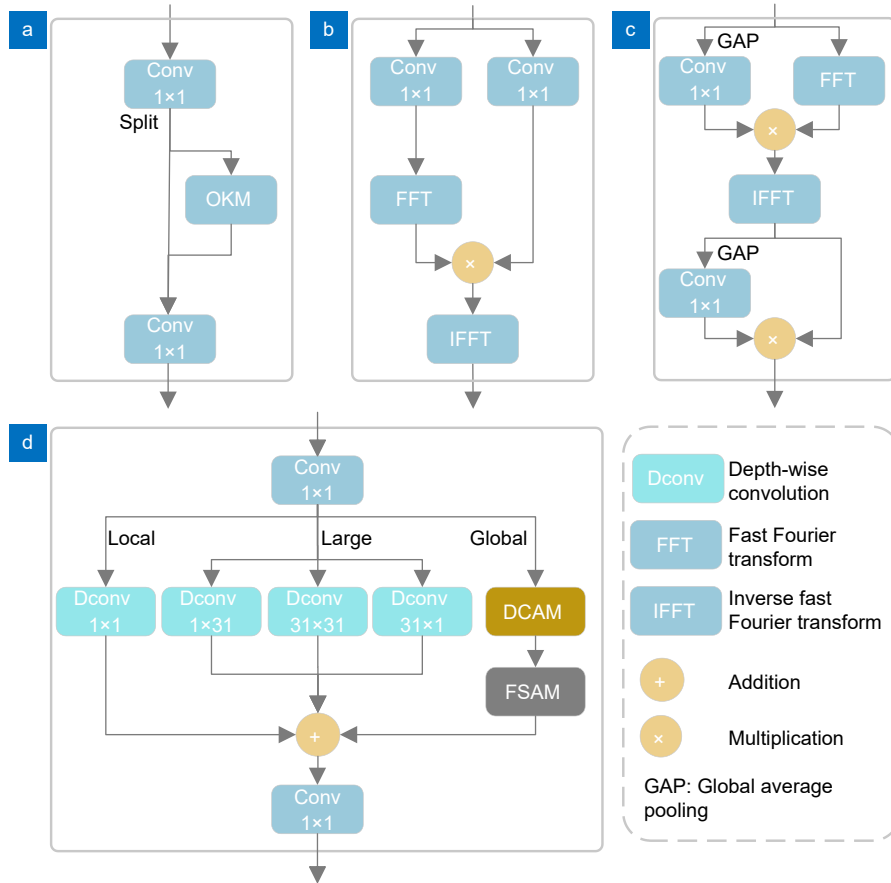


图 6 CSP-OKM 模块及其子模块结构图。(a) CSP-OKM; (b) FSAM; (c) DCAM; (d) OKM

Fig. 6 Structure diagram of CSP-OKM module and associated submodules. (a) CSP-OKM; (b) FSAM; (c) DCAM; (d) OKM

PEL_l) 表示在测试集所有图像中, 类型 l 关键点的预测坐标与对应真实坐标的 L_2 距离的平均值。误差的单位为像素。数值越小, 代表类型 l 关键点的定位越精确。

关键点平均定位误差 (Average point-to-point error, APE) 表示所有不同类型关键点定位误差的平均值, 能够衡量算法的关键点检测性能。

角度误差 (Error for angle, EA_k) 表示在测试集所有图像中, 不同关键点坐标连线组成的类型角度 k , 计算预测关键点与真实关键点形成的类型角度 k 之间的平均误差。单位为角度, 其数值越小, 代表类型 k 的角度预测越接近真实值, 准确性越高。

平均角度误差 (Average angle error, AAE) 表示所有不同类型角度误差的平均值。其数值越小, 反映检测算法计算的 AI 值与真实 AI 值一致性越高, 诊断参考结果在医学应用中越具可信性。

成功检测率 (Successful detection rate, SDR) 表示特定误差范围内, 关键点被成功检测的数量与关键点总数的比例。

漏检率 (Missed detection rate, MDR) 表示测试集中未检测出关键点的样本数量占实际样本数量的百分比。

浮点运算数 (Floating point operations, FLOPs) 表示模型在进行推理或训练过程中实际进行的浮点运算总量, 反映了模型计算复杂度。

在 DDH 关键点检测中, L_2 距离考虑了髌关节图像平面上所有方向上的距离变化, 有助于更准确地评价预测关键点与真实标签关键点之间的差异。因此, 本文采用 L_2 距离计算实验中关键点的定位误差, 并使用 PEL_{RASM} 、 PEL_{RTCC} 、 PEL_{LTCC} 、 PEL_{LASM} 、APE、 EA_R 、 EA_L 、AAE、SDR、MDR 和 FLOPs 等指标对检测算法进行定性分析, 其部分评价指标计算公式见表 1。在实验中, 粗体和下划线分别表示在指标中算法的最优和次优效果。

在表 1 中, n 代表被检测的图像数量, i 代表图像的编号, T 代表医生标记的真实标签, P 代表网络模型的预测标签, l 代表图像中不同关键点类型的代号, 在盆骨 X 射线图像中有 RASM、RTCC、LTCC

和 LSAM 四种关键点类型, T_{li} 、 P_{li} 分别表示第 i 张图像中关键点类型 l 的真实坐标、预测坐标, k 代表角度类型, 在 X 射线图像中有 R 和 L 两种角度类型, T_{ki} 、 P_{ki} 分别表示第 i 张图像中角度类型 k 的真实角度值、预测角度值, z 代表关键点的定位误差范围, x 、 y 分别代表关键点在 x 、 y 轴上的位置, m 代表模型预测结果中存在漏检的图像数量。

4.4 消融实验

4.4.1 EFEM 模块实验分析

为验证 EFEM 模块强化边缘特征的设计有效性, 首先, 本文通过固定边缘检测算子验证模块的有效性, 实验结果如表 2 所示, 其 APE、AAE 指标相较于 YOLOV8s 均得到了有效降低。其次, 为验证 BN 在模块内部不同特征拼接时对缓解数值尺度差异的有效性, 从表 2 可知, 使用 BN 后网络在所有指标上效果均优于直接拼接操作, 证明其减少了因不同特征数值尺度差异对网络性能的影响。最后, 模块使用 Sobel、Prewitt 和 Laplacian 三种边缘检测算子进行替换对比实验, 从表 2 可看出, EFEM 中的所有边缘检测算子在 APE 指标上均优于 YOLOV8s。由于 Sobel 算子在 APE 指标上表现最优, 且在 AAE 指标上也取得次优

效果, 故本文选择 Sobel 算子作为该模块的边缘检测算子。

4.4.2 DAN 网络实验分析

为验证 DAN 的优越性, 本文将 YOLOv8 原始 Neck、Small-neck 和 DAN 在 YOLOv8s 上进行对比实验, 将其与 DAN 以及基准模型进行对比。实验结果如表 3 所示, DAN 网络在 APE、AAE 上均取得最优结果, 且对比于 Small-neck, FLOPs 降低了 5.7%, 而 Small-neck 结果次于 YOLOv8 原始 Neck。实验结果表明其并不适合迁移到关键点局部邻域检测, 可能因该网络未针对关键点局部邻域内的细节特征进行优化, 且专注于提取一般特征, 不足以描述局部邻域内复杂的局部特征, 导致无法有效提升关键点检测精度。

4.4.3 不同改进模块对算法的影响实验分析

为验证不同改进模块对关键点识别能力提升的有效性, 本文算法以 YOLOv8s 的网络结构为基础网络, 在 Backbone 中使用 EFEM 模块嵌入到 C2f 中, 在 Neck 中采用设计的 DAN 网络。实验结果如表 4 所示, 所有改进算法均有效提升了关键点的检测精度, 符合本文设计 EDA-YOLOv8s 的算法思路。

表 1 评价指标与计算公式

Table 1 Evaluation index and computational formula

Index	Calculation formula
PEL_l	$\left(\sum_{i=1}^n \sqrt{T_{li} - P_{li}}\right) / n$
APE	$(PEL_{RASM} + PEL_{RTCC} + PEL_{LTCC} + PEL_{LASM}) / 4$
EA_k	$\left(\sum_{i=1}^n T_{ki} - P_{ki} \right) / n$
AAE	$(EA_R + EA_L) / 2$
SDR	$\left\{i : \sqrt{(x_{T_{li}} - x_{P_{li}})^2 + (y_{T_{li}} - y_{P_{li}})^2} \leq z\right\} \times 100\% / n$
MDR	$\frac{m}{n} \times 100\%$

表 2 EFEM 模块内消融和边缘检测算子对比实验结果

Table 2 The ablation and edge detection operator comparison experiment results in the EFEM module

EFEM (edge detection operator)	BN	APE/pixel	AAE/(°)
×	×	4.5150	1.6504
Sobel	×	4.3652	1.5842
Sobel	√	4.3459	1.5565
Prewitt	√	<u>4.3470</u>	1.5555
Laplacian	√	4.4130	1.6934

表 3 三种 Neck 网络算法对比实验结果

Table 3 Comparison experiment results of three neck network algorithms

Neck network	APE/pixel	AAE/(°)	FLOPs/G
YOLOv8s (neck)	4.5150	1.6504	28.8
YOLOv8s+ (small-neck)	4.7150	1.7322	36.9
YOLOv8s+ (DAN)	4.3281	1.5451	34.8

表 4 不同改进模块对算法的影响

Table 4 The influence of different improved modules on the algorithm

Detection algorithm	PEL_{RASM}/pixel	PEL_{RTCC}/pixel	PEL_{LTCC}/pixel	PEL_{LASM}/pixel	APE/pixel	$EA_R/(\circ)$	$EA_L/(\circ)$	AAE/(\circ)
YOLOv8s	4.8476	3.7404	3.4674	6.0045	4.5150	1.6250	1.6757	1.6504
YOLOv8s+EFEM	4.6133	3.7692	3.2258	5.7752	4.3459	1.4886	1.6244	1.5565
YOLOv8s+DAN	4.1284	<u>3.6947</u>	3.3784	6.1110	<u>4.3281</u>	1.2587	1.8315	<u>1.5451</u>
EDA-YOLOv8s	4.2851	3.4870	<u>3.2530</u>	<u>5.8107</u>	4.2090	<u>1.3420</u>	<u>1.6324</u>	1.4872

在 YOLOv8s 中使用 EFEM 模块后, APE 为 4.3459 pixel, AAE 为 1.5565°, 相较于 YOLOv8s 分别降低了 3.7% 和 5.7%, 表明该模块通过增强骨骼区域边缘特征, 能够促进网络对边缘上的关键点识别。此外, 在 YOLOv8s 的 Neck 中使用 DAN 网络后, 其 APE 为 4.3281 pixel, AAE 为 1.5451°, 相较于 YOLOv8s 分别降低了 4.1% 和 6.4%, 表明该网络通过引入高分辨率特征, 并对特征进行优化后的细节信息能够增强网络对关键点的识别能力。在 YOLOv8s 中同时使用 EFEM 模块和 DAN 网络后, 本文算法在 PEL_{RASM} 、 PEL_{RTCC} 、 PEL_{LTCC} 和 PEL_{LASM} 四个不同类型关键点定位误差指标以及 EA_R 和 EA_L 两个不同类型关键点角度误差指标上均能取到最优或次优效果。相较于 YOLOv8s, 在 APE、AAE 两个主要指标上分别降低了 6.8%、9.9%。实验结果表明在 EDA-YOLOv8s 中, 经 EFEM 模块强化边缘后的不同层级特征图为 DAN 网络提供了更为丰富的边缘信息, 有助于网络进一步捕获关键点骨骼边缘区域存在的细微结构, 从

而有效地提升关键点的检测精度。

为直观表现各模块对 EDA-YOLOv8s 的影响, 本文对上述消融实验进行了可视化。可视化结果如图 7 所示, 图中红色圆点和矩形框分别表示真实标签下的关键点和关键点局部邻域。从图 7(a) 可以看出, YOLOv8s 在处理髋关节 X 射线图像时, 注意力分布较为分散, 且对关键点所在感兴趣区域的关注度相对较低。当分别在 YOLOv8s 中使用 EFEM 模块和 DAN 网络后, 如图 7(b) 和图 7(c) 所示, 网络对关键点的注意力均得到了显著改善。其中, EFEM 模块使网络注意力更多地集中于边缘区域, 对关键点区域的关注度有明显提升; 而 DAN 网络则能够有效强化对关键点局部邻域的关注, 并减少无关背景的干扰。从图 7(d) 可看出, 同时使用两者后的 EDA-YOLOv8s 进一步降低了背景干扰, 并且网络的注意力更加聚焦于关键点及其邻域。

4.4.4 APE 与 AAE 关联性实验分析

为探究不同改进模块下 APE 与 AAE 之间的关联

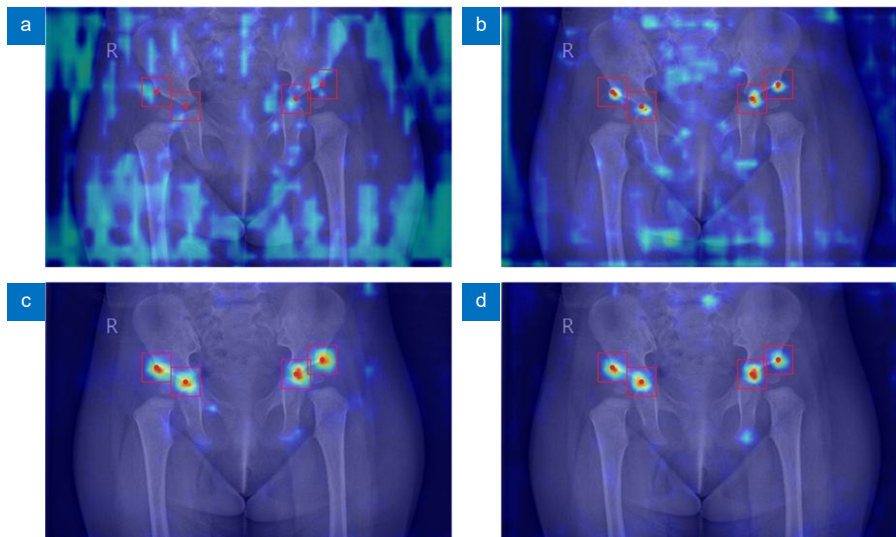


图 7 不同改进模块后的髋关节 X 射线图像热力图。(a) YOLOv8s; (b) YOLOv8s+EFEM; (c) YOLOv8s+DAN; (d) EDA-YOLOv8s

Fig. 7 Heatmaps of hip X-ray image after different improved modules. (a) YOLOv8s; (b) YOLOv8s+EFEM; (c) YOLOv8s+DAN; (d) EDA-YOLOv8s

性影响, 本文对其进行了相关性分析实验。实验结果如图 8 所示, 在 YOLOv8s 中加入不同改进模块后, APE 的降低也伴随着 AAE 的降低, 二者具有正向相关性。表明在关键点定位精度仍具有提升空间的情况下, 降低 APE 能够促进 AAE 下降。

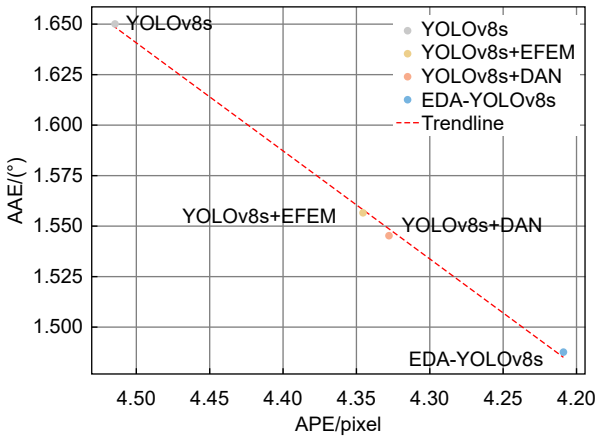


图 8 使用不同改进模块后的 APE 与 AAE 变化关系图
Fig. 8 The relationship diagram between APE and AAE after using different improvement modules

4.4.5 SDR 实验分析

为评估 EDA-YOLOv8s 在不同定位误差范围内关键点的成功检测率, 本文进行了髌关节关键点在误差范围 $z \in \{5 \text{ pixel}, 10 \text{ pixel}, 15 \text{ pixel}\}$ 的成功检测率实验。实验结果如图 9 所示, 在定位误差小于 10 pixel 时, RASM、RTCC 和 LTCC 关键点的成功检测率能达到 91.0% 以上; 在定位误差为 15 pixel 内, 四个类别关键点的成功检测率均在 95.0% 以上, 表明本文算法在髌关节关键点检测任务上实现了较高的检测精度。

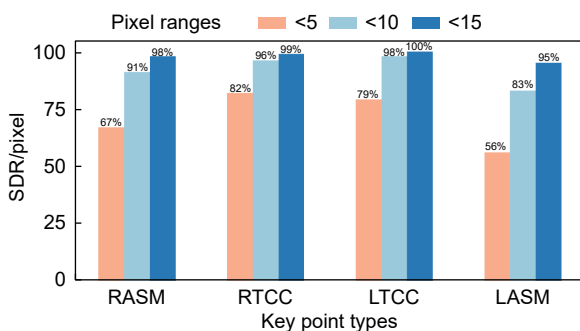


图 9 关键点成功检测率实验结果
Fig. 9 Experimental results of successful detection rate of keypoint

4.5 对比实验

4.5.1 不同关键点检测算法对比实验分析

为验证 EDA-YOLOv8s 算法的优越性, 本文选取

了四个通用型目标检测算法和三个髌关节关键点检测算法进行对比实验, 算法分别为 YOLOv8n、YOLOv8s、YOLOv9s^[30]、YOLOv10s、PN-UNet^[11]、CircleNet^[12] 和 CBA-YOLOv5s^[13]。实验结果如表 5 所示, 在四个通用性目标检测算法中, YOLOv8s 在 APE 指标上表现最佳。实验结果表明 YOLOv8s 采用 225 层的适中网络深度, 能够更好地平衡特征提取能力和细节信息保留, 满足关键点检测对精确定位的需求。相比之下, YOLOv9s 显著增加网络深度并使用可编程梯度信息技术解决深层网络中的信息瓶颈问题, 但过深网络在多层次特征变换中可能会丢失影响关键点定位的局部细节, 导致 APE 无法有效降低。YOLOv10s 通过采用双重分配策略代替非极大值抑制, 着重优化推理效率。但该算法在关键点检测任务中 MDR 为 6.0%, 这一高漏检率限制了其在实际应用中的可靠性^[13]。YOLOv8n 因侧重于轻量级和快速推理, 在特征提取能力上相对较弱, 导致在 APE 指标上表现较差。综上, YOLOv8s 在通用型目标检测算法中性能与效率平衡较好, 但其未针对关键点检测任务优化, 导致部分图像样本的关键点定位误差仍较大。而基于 YOLOv8s 的改进算法 EDA-YOLOv8s 在关键点检测精度上具有明显提升, 对比 YOLOv8n、YOLOv8s、YOLOv9s、YOLOv10s、PN-UNet、CircleNet 和 CBA-YOLOv5s 在 APE 指标上分别下降了 8.6%、6.8%、7.1%、4.9%、7.2%、6.5% 和 6.2%。此外, 本文算法在 PEL_{RASM} 、 PEL_{RTCC} 和 PEL_{LASM} 三个不同类型关键点定位误差指标以及 EA_L 类型关键点角度误差指标上取得最佳效果, 并在 APE 和平均角度误差 AAE 两个主要指标上同时取得最好表现, 表明该算法具有较优越的检测性能。

4.5.2 关键点定位误差分布对比分析

为确定定位误差较大的关键点的降低是否为 APE 下降的重要因素, 本文对 YOLOv8s 和 EDA-YOLOv8s 预测结果中定位误差大于 8 pixel 的关键点, 进行了定位误差分布对比实验。实验结果如图 10(a, b) 所示, YOLOv8s 预测的关键点定位误差上边缘、上四分位数和下四分位数分别为 25 pixel、17 pixel 和 10.5 pixel, 且误差分布的离散程度较大。而 EDA-YOLOv8s 预测的关键点定位误差上边缘、上四分位数以及下四分位数分别为 17.5 pixel、13.5 pixel 和 9 pixel, 离散程度较小。实验表明 APE 的降低, 能够反映出大部分定位误差较大的关键点得

到实质性改善, 使之更接近真实标签。此外, 提升定位误差较大关键点的检测精度, 还能保证依据其计算出的 AI 值具有可信性。

4.5.3 可视化对比

为直观展示本文算法的检测效果, 本文选取了对比度低和边缘模糊等低质量的髋关节图像进行关键点

检测结果和关键点角度可视化对比, 分别如图 11 和图 12 所示。在两个可视化对比结果图中均采用红色、绿色和蓝色圆点分别代表真实标签关键点、EDA-YOLOv8s 和 YOLOv8s 预测的关键点。此外, 在图 11 中, 不同颜色矩形框和直线分别表示各自关键点对应下的局部邻域大小和关键点连线; 在图 12 中, 红色虚

表 5 不同关键点检测算法对比实验结果

Table 5 Comparison experiment results of different keypoint detection algorithms

Detection algorithm	PEL_{RASM}/pixel	PEL_{RTCC}/pixel	PEL_{LTCC}/pixel	PEL_{LASM}/pixel	APE/pixel	$EA_R/(\circ)$	$EA_L/(\circ)$	AAE/(\circ)	MDR/%
YOLOv8n	5.5012	3.8100	3.2056	5.9099	4.6067	1.4698	1.7188	1.5943	0
YOLOv8s	4.8476	3.7404	3.4674	6.0045	4.5150	1.6250	1.6757	1.6504	0
YOLOv9s	4.6463	3.7893	3.4716	6.2252	4.5331	1.4878	1.8033	1.6456	0
YOLOv10s	4.6806	3.7820	3.3007	5.9486	<u>4.4280</u>	1.4594	1.8313	1.6454	<u>6.0</u>
PN-UNet	4.5411	3.7115	3.6881	6.2045	4.5363	1.4384	<u>1.6533</u>	<u>1.5458</u>	0
CircleNet	4.8348	3.7234	3.6113	<u>5.8375</u>	4.5017	1.4988	1.7996	1.6492	0
CBA-YOLOv5s	<u>4.3828</u>	<u>3.6464</u>	3.5506	6.3773	4.4893	1.3235	1.8488	1.5861	0
EDA-YOLOv8s	4.2851	3.4870	<u>3.2530</u>	5.8107	4.2090	<u>1.3420</u>	1.6324	1.4872	0

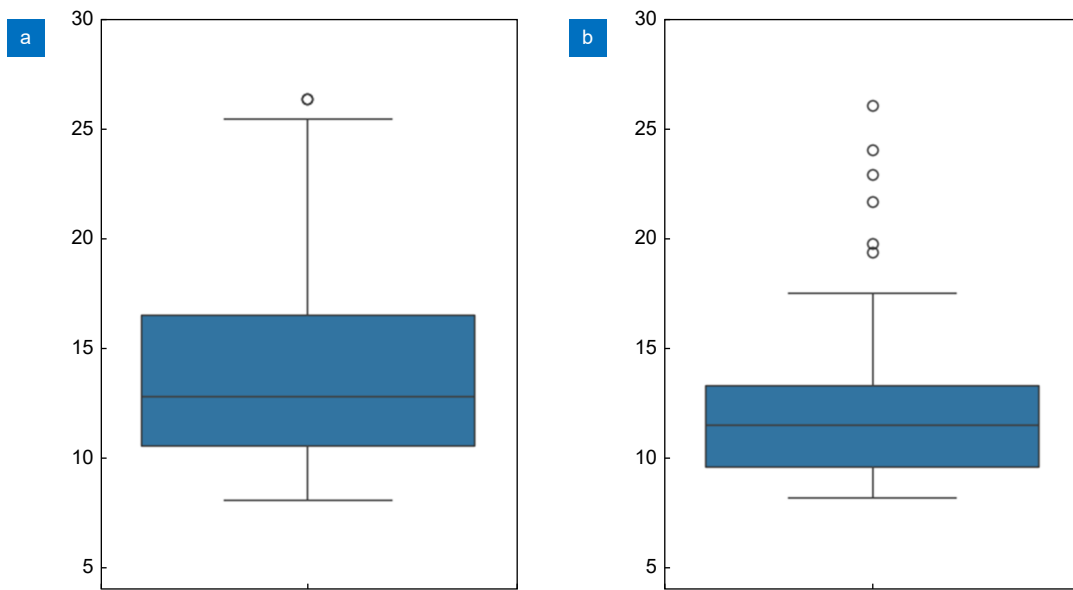


图 10 关键点定位误差分布对比。(a) YOLOv8s; (b) EDA-YOLOv8s

Fig. 10 Comparison of keypoint localization error distribution. (a) YOLOv8s; (b) EDA-YOLOv8s

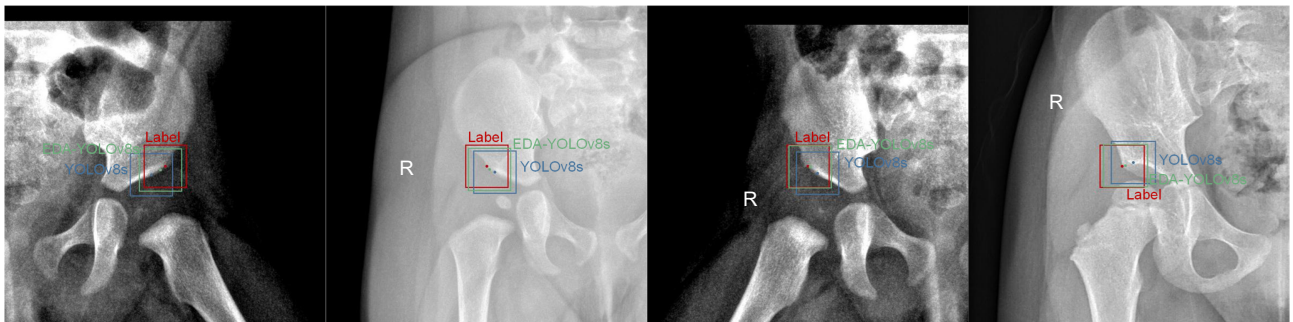


图 11 EDA-YOLOv8s 与 YOLOv8s 检测结果可视化对比

Fig. 11 Visualization comparison of detection results between EDA-YOLOv8s and YOLOv8s

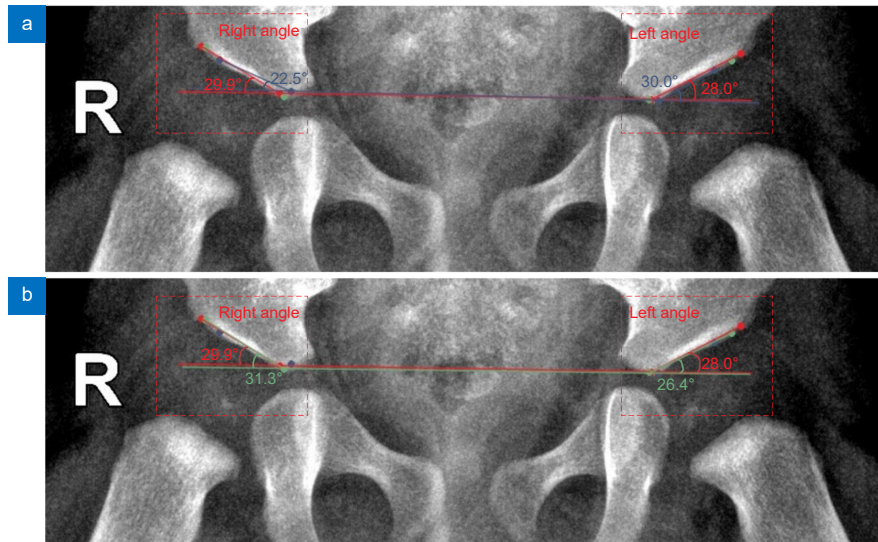


图 12 EDA-YOLOv8s 与 YOLOv8s 关键点角度可视化对比。(a) YOLOv8s; (b) EDA-YOLOv8s

Fig. 12 Visualization comparison of keypoint angle between EDA-YOLOv8s and YOLOv8s. (a) YOLOv8s; (b) EDA-YOLOv8s

线框则标记了不同关键点角度类型。从图 11 可看到, 相较于 YOLOv8s, EDA-YOLOv8s 能够更准确地识别关键点并最大程度地贴近真实标签中的关键点。这得益于 EDA-YOLOv8s 中的 EFEM 模块能够通过增强边缘特征辅助网络识别出关键点所在骨骼边缘。随后, DAN 网络利用强化的边缘特征和引入的高分辨率特征图提取关键点局部邻域细节信息, 从而提高网络对关键点的识别能力。而在图 12(a) 和图 12(b) 中的 Left Angle 和 Right Angle 虚线框内可看到, 与 YOLOv8s 相比, EDA-YOLOv8s 在关键点位置和关键点角度上更贴近真实标签, 特别在关键点角度类型 k 为右角度时, EDA-YOLOv8s 与真实标签的匹配度明显更高。这表明本文算法可以准确提取到识别关键点所需的重要特征, 从而在关键点的位置和角度上与真实标签的一致性更高, 展示了 EDA-YOLOv8s 在髋关节关键点检测任务中优越的检测性能。

5 结论

为解决髋关节 X 射线图像中边缘模糊和关键点邻域内细节信息提取不足的问题, 本文提出一种融合边缘特征与细节感知网络的 YOLOv8s 髋关节关键点检测算法。该算法利用 EFEM 模块, 使网络在捕获周围空间信息特征的同时, 增强了能够辅助识别关键点的边缘特征。此外, 使用 DAN 网络, 将经强化边缘后的多层级特征进行特征融合和优化, 提升了对关键点邻域内细微关键结构的捕捉能力, 从而提升网络对

关键点的检测精度。然而, 针对部分图像中存在的骨骼遮挡重叠问题, 算法缺乏对该区域信息的关注, 导致网络对这些图像中的关键点识别困难。因此, 提升网络对骨骼重叠区域中关键点的定位精度是我们下一步的工作方向。

利益冲突: 所有作者声明无利益冲突

参考文献

- [1] Bradley C S, Verma Y, Maddock C L, et al. A comprehensive nonoperative treatment protocol for developmental dysplasia of the hip in infants: a prospective longitudinal cohort study[J]. *Bone Joint J*, 2023, **105-B**(8): 935-942.
- [2] Alrashdi N, Alotaibi M, Alharthi M, et al. Incidence, prevalence, risk factors, and clinical treatment for children with developmental dysplasia of the hip in Saudi Arabia. a systematic review[J]. *J Epidemiol Glob Health*, 2024, **14**(3): 549-560.
- [3] Fujii M, Kawano S, Ueno M, et al. Clinical results of periacetabular osteotomy with structural bone allograft for the treatment of severe hip dysplasia[J]. *Bone Joint J*, 2023, **105-B**(7): 743-750.
- [4] Dornacher D, Lutz B, Fuchs M, et al. Acetabular deficiency in borderline hip dysplasia is underestimated by lateral center edge angle alone[J]. *Arch Orthop Trauma Surg*, 2023, **143**(7): 3937-3944.
- [5] Iyengar K P, Fitzpatrick J D, Michalos M, et al. Birmingham royal orthopaedic hospital (BROH) femoral offset—an ancillary measure of adult dysplasia of the hip[J]. *Indian J Radiol Imaging*, 2023, **33**(4): 471-477.
- [6] McQuivey K S, Secretov E, Domb B G, et al. A multicenter

- study of radiographic measures predicting failure of arthroscopy in borderline hip dysplasia: beware of the Tönnis angle[J]. *Am J Sports Med*, 2020, **48**(7): 1608–1615.
- [7] Li Q, Zhong L, Huang H N, et al. Auxiliary diagnosis of developmental dysplasia of the hip by automated detection of sharp's angle on standardized anteroposterior pelvic radiographs[J]. *Medicine*, 2019, **98**(52): e18500.
- [8] Liu C B, Xie H T, Zhang S C, et al. Misshapen pelvis landmark detection by spatial local correlation mining for diagnosing developmental dysplasia of the hip[C]//*Proceedings of the 22nd International Conference on Medical Image Computing and Computer Assisted Intervention*, 2019: 441–449. https://doi.org/10.1007/978-3-030-32226-7_49.
- [9] Al-Bashir A K, Al-Abed M, Sharkh F M A, et al. Algorithm for automatic angles measurement and screening for developmental dysplasia of the hip (DDH)[C]//*Proceedings of 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2015: 6386–6389. <https://doi.org/10.1109/EMBC.2015.7319854>.
- [10] Rana M, Bhushan M. Machine learning and deep learning approach for medical image analysis: diagnosis to detection[J]. *Multimed Tools Appl*, 2023, **82**(17): 26731–26769.
- [11] Liu C B, Xie H T, Zhang S C, et al. Misshapen pelvis landmark detection with local-global feature learning for diagnosing developmental dysplasia of the hip[J]. *IEEE Trans Med Imaging*, 2020, **39**(12): 3944–3954.
- [12] Wu H, Xie H T, Liu C B, et al. CircleNet for hip landmark detection[C]//*Proceedings of the 34th AAAI Conference on Artificial Intelligence*, 2020, **34**: 12370–12377. <https://doi.org/10.1609/aaai.v34i07.6922>.
- [13] Lv J, Che J L, Chen X. CBA-YOLOv5s: a hip dysplasia detection algorithm based on YOLOv5s using angle consistency and bi-level routing attention[J]. *Biomed Signal Process Control*, 2024, **95**: 106482.
- [14] Xu W Z, Shu L Q, Gong P, et al. A deep-learning aided diagnostic system in assessing developmental dysplasia of the hip on pediatric pelvic radiographs[J]. *Front Pediatr*, 2022, **9**: 785480.
- [15] Xu J Y, Xie H T, Tan Q F, et al. Multi-task hourglass network for online automatic diagnosis of developmental dysplasia of the hip[J]. *World Wide Web*, 2023, **26**(2): 539–559.
- [16] Kim N. Computational modelling for surgical planning of hip dysplasia[D]. Auckland: the University of Auckland, 2023.
- [17] Rahmawati S, Devita R, Zain R H, et al. Prewitt and canny methods on inversion image edge detection: an evaluation[J]. *J Phys: Conf Ser*, 2021, **1933**(1): 012039.
- [18] Tian B, Wei W. Research overview on edge detection algorithms based on deep learning and image fusion[J]. *Secur Commun Networks*, 2022, **2022**(1): 1155814.
- [19] Xie S N, Tu Z W. Holistically-nested edge detection[C]//*Proceedings of 2015 IEEE International Conference on Computer Vision*, 2015: 1395–1403. <https://doi.org/10.1109/ICCV.2015.164>.
- [20] Kong L Y, Wang F B, Yang F Y, et al. FISRCN: a single small-sized image super-resolution convolutional neural network by using edge detection[J]. *Multimed Tools Appl*, 2024, **83**(7): 19609–19627.
- [21] Redmon J, Divvala S, Girshick R, et al. You only look once: unified, real-time object detection[C]//*Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition*, 2016: 779–788. <https://doi.org/10.1109/CVPR.2016.91>.
- [22] Lin T Y, Dollár P, Girshick R, et al. Feature pyramid networks for object detection[C]//*Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition*, 2017: 936–944. <https://doi.org/10.1109/CVPR.2017.106>.
- [23] Liu S, Qi L, Qin H F, et al. Path aggregation network for instance segmentation[C]//*Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2018: 8759–8768. <https://doi.org/10.1109/CVPR.2018.00913>.
- [24] Cui Y N, Ren W Q, Knoll A. Omni-kernel network for image restoration[C]//*Proceedings of the 38th AAAI Conference on Artificial Intelligence*, 2024, **38**: 1426–1434. <https://doi.org/10.1609/aaai.v38i2.27907>.
- [25] Zheng J W, Shao A H, Yan Y D, et al. Remote sensing semantic segmentation via boundary supervision-aided multiscale channelwise cross attention network[J]. *IEEE Trans Geosci Remote Sens*, 2023, **61**: 4405814.
- [26] Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift[C]//*Proceedings of the 32nd International Conference on Machine Learning*, 2015: 448–456.
- [27] Pan J Y, Zhang Y J. Small object detection in aerial drone imagery based on YOLOv8[J]. *IAENG Int J Comput Sci*, 2024, **51**(9): 1346–1354.
- [28] Xu L Y, Zhao Y F, Zhai Y H, et al. Small object detection in UAV images based on YOLOv8n[J]. *Int J Comput Intell Syst*, 2024, **17**(1): 223.
- [29] Xie J, Deng Y M, Wang R M. Improved traffic sign detection algorithm based on YOLOv8s[J]. *Comput Eng*, 2024, **50**(11): 338–349. 谢竞, 邓月明, 王润民. 改进 YOLOv8s 的交通标志检测算法[J]. *计算机工程*, 2024, **50**(11): 338–349.
- [30] Wang C Y, Yeh I H, Mark Liao H Y. YOLOv9: learning what you want to learn using programmable gradient information[C]//*Proceedings of the 18th European Conference on Computer Vision*, 2024: 1–21. https://doi.org/10.1007/978-3-031-72751-1_1.
- [31] Muhammad W, Aramvith S, Onoye T. Multi-scale Xception based depthwise separable convolution for single image super-resolution[J]. *PLoS One*, 2021, **16**(8): e0249278.
- [32] Wang Y, Wang H H, Liu S D, et al. Small target detection in sonar images with multilevel feature screening and task dynamic alignment[J]. *Opto-Electron Eng*, 2024, **51**(10): 240196. 王燕, 王宏辉, 刘树东, 等. 多级特征筛选和任务动态对齐的声呐图像小目标检测[J]. *光电工程*, 2024, **51**(10): 240196.

作者简介



吕佳 (1978-), 女, 博士, 硕士生导师, 主要从事机器学习、数据挖掘及其在医学图像处理等方面的研究。

E-mail: lvjia@cqnu.edu.cn



【通信作者】陈欣 (1984-), 女, 博士, 副主任医师, 主要从事人工智能在儿童影像等方面的应用研究。

E-mail: b2309@126.com



段训禄 (1999-), 男, 硕士研究生, 主要从事深度学习及其在医学图像处理等方面的研究。

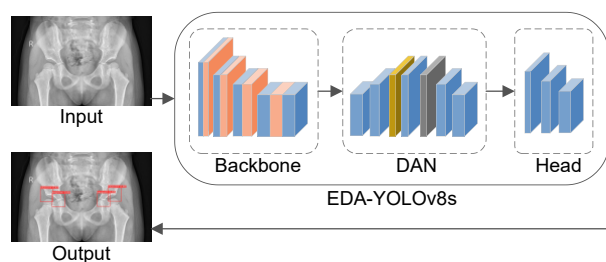
E-mail: 2023210516031@stu.cqnu.edu.cn



扫描二维码, 获取PDF全文

Edge feature and detail-aware network integrated YOLOv8s algorithm for hip joint keypoint detection

Lv Jia^{1,2}, Duan Xunlu¹, Chen Xin^{3*}



Schematic diagram of the key point detection process in EDA-YOLOv8s

Overview: Hip dysplasia is a common orthopedic disease in newborns, and timely and precise diagnosis is critical for optimal patient outcomes. In clinical diagnosis, specific key points on hip joint X-ray images are typically annotated, followed by the calculation of the acetabular index through angular measurements based on these points using goniometric tools. The diagnosis is then determined in combination with the age of the patient. However, manual annotation of key points in the hip joint not only demands that clinicians possess robust professional expertise and extensive clinical experience but also renders the process highly time-consuming and susceptible to subjective bias. Therefore, there is an urgent need for precise and automated key points detection technology to assist doctors in diagnosis. However, traditional template matching methods exhibit poor robustness and generalization when processing complex hip X-ray images, especially when faced challenges such as illumination changes, occlusions, and image rotations. To address these issues, researchers have enhanced the attention mechanism and extracted detailed information around key points using deep learning techniques, thereby improving the accuracy of key points localization. Nonetheless, these methods overlook the significance of bone edge information in assisting recognition and struggle with identifying local neighborhood key structural features, which limits further improvements in localization accuracy. To resolve these problems, an edge feature and detail-aware network integrated with the YOLOv8s algorithm for hip joint key points detection is proposed in this paper. This algorithm introduces an edge feature enhancement module to capture spatial features around key points and enhance the edge features of the bones where they are located. The module is applied multiple times during the feature extraction process of the network to progressively strengthen edge features and guide the network to focus on the key points edge areas. In addition, a detail-aware network is proposed to perform feature fusion and optimization on feature maps at different levels, enhancing the network's ability to capture important fine structures within the local neighborhood of key points. The algorithm was experimentally tested on the hip joint X-ray image dataset provided by the Department of Imaging of the Children's Hospital Affiliated to Chongqing Medical University. The results demonstrate that the average localization error and average angular error for key points have been reduced to 4.2090 pixel and 1.4872°, respectively, representing reductions of 6.8% and 9.9% compared with YOLOv8s, and significantly outperforms existing methods. The experimental findings confirm that the proposed algorithm effectively enhances the accuracy of key point detection, offering valuable insights for clinical diagnosis.

Lv J, Duan X L, Chen X. Edge feature and detail-aware network integrated YOLOv8s algorithm for hip joint keypoint detection[J]. *Opto-Electron Eng*, 2025, 52(3): 240281; DOI: 10.12086/oe.2025.240281

Foundation item: National Natural Science Foundation Projects (11991024), Chongqing Education Commission Key Project (KJZD-K202200511), and Chongqing Natural Science Foundation Innovation and Development Joint Fund Key Project(CSTB2024NSCQ-LZX0090)

¹College of Computer and Information Sciences, Chongqing Normal University, Chongqing 401331, China; ²National Center for Applied Mathematics in Chongqing, Chongqing Normal University, Chongqing 401331, China; ³Department of Radiology, Children's Hospital of Chongqing Medical University, National Clinical Research Center for Child Health and Disorders, Ministry of Education Key Laboratory of Child Development and Disorders, Chongqing Key Laboratory of Pediatrics, Chongqing 400014, China

* E-mail: b2309@126.com