

本文引用格式: 肖军峰, 丁鹏. 基于改进 YOLOv8s 的膝骨关节炎自动诊断算法[J]. 自动化与信息工程, 2024, 45(6): 29-35.

XIAO Junfeng, DING Peng. Automatic diagnosis algorithm for knee osteoarthritis based on improved YOLOv8s[J]. Automation & Information Engineering, 2024, 45(6): 29-35.

基于改进 YOLOv8s 的膝骨关节炎自动诊断算法*

肖军峰¹ 丁鹏²

(1. 中国石油广东销售广州分公司, 广东 广州 510000

2. 东华理工大学信息工程学院, 江西 南昌 330000)

摘要: 膝骨关节炎是导致老年人活动能力受限和身体残疾的主要原因之一, 早期发现和干预对于延缓病情发展、改善患者的生活质量具有重要意义。针对现有膝骨关节炎诊断算法检测精度低的问题, 提出一种基于改进 YOLOv8s 的膝骨关节炎自动诊断算法。该算法提出一种改进的卷积块注意力机制模块 (CBAM), 使网络能够更加关注膝关节图像的关键信息, 提高膝骨关节炎的检测精度; 设计一种基于多尺度线性注意力的 Focal Modulation 模块, 以提高网络的多尺度特征表达能力。实验结果表明, 该算法在测试集上的平均精度均值为 0.791, 有效实现了膝骨关节炎的自动诊断。

关键词: 膝骨关节炎; YOLOv8s; Transformer; 注意力机制

中图分类号: TP242.2

文献标志码: A

文章编号: 1674-2605(2024)06-0004-07

DOI: 10.3969/j.issn.1674-2605.2024.06.004

开放获取

Automatic Diagnosis Algorithm for Knee Osteoarthritis Based on Improved YOLOv8s

XIAO Junfeng¹ DING Peng²

(1. China National Petroleum Corporation Guangdong Sales Guangzhou Branch, Guangzhou 510000, China

2. School of Information Engineering, East China University of Technology, Nanchang 330000, China)

Abstract: Knee osteoarthritis is one of the main causes of limited mobility and physical disability in the elderly. Early detection and intervention are of great significance for delaying the progression of the disease and improving the quality of life of patients. Aiming at the problem of low detection accuracy of existing knee osteoarthritis diagnosis algorithms, a knee osteoarthritis automatic diagnosis algorithm based on improved YOLOv8s is proposed. This algorithm proposes an improved attention mechanism module (CBAM) of the convolution module, which enable the network to pay more attention to the key information of knee joint images and improve the detection accuracy of knee osteoarthritis; Design a Focal Modulation module based on multi-scale linear attention to improve the multi-scale feature representation ability of the network. The experimental results show that the average accuracy of the algorithm on the test set is 0.791, effectively achieving automatic diagnosis of knee osteoarthritis.

Keywords: knee osteoarthritis; YOLOv8s; Transformer; attention mechanism

0 引言

膝骨关节炎是一种常见的关节疾病, 其典型症状主要表现为膝关节疼痛和活动受限^[1]。目前, 我国 60 岁以上的老年人群中膝骨关节炎的发病率约为 58%, 而 70 岁以上的老年人群中这一比例高达 75%^[2]。传

统的膝骨关节炎临床诊断主要依据膝关节 X 光影像确定其 Kellgren-Lawrence (KL) 分级, 诊断的准确性依赖于医生的专业知识和主观判断。据统计, 约有 20% 的膝骨关节炎患者在首次就诊时遭遇误诊或漏诊^[3], 这不仅延误了治疗时机^[4], 还可能导致关节畸形

或残疾^[5]。早期发现并及时治疗可以延缓膝骨关节炎的发展。因此，开发一种能够提高诊断准确性的膝骨关节炎自动诊断算法，对改善患者的治疗效果具有重要意义。

随着深度学习技术的快速发展，一系列成熟的卷积神经网络（convolutional neural networks, CNN）分类模型，如 ResNet、InceptionV3 和 DenseNet 等，已被广泛应用于医学图像分类任务中^[6-8]。文献[9]提出一种基于深度学习的图像分析方法，用于对混合型肝细胞癌的重新分类，能够提高肝细胞癌的诊断准确性，降低误诊率。文献[10]提出一种结合多尺度特征融合和注意力机制的方法，增强模型对肿瘤结构和形态特征的捕捉能力，实现对肿瘤及其周围组织的更精确分割。文献[11]提出一种融合 VGG-16 和 ResNet50 模型的改进型 Faster R-CNN 网络，该网络在处理特征不明

显的眼底图像时，检测精度可达 97.42%，提升了糖尿病视网膜病变筛查的效率和准确性。文献[12]开发了一种基于眼底图像的深度学习系统，实现了糖尿病视网膜病变进展的风险预警及时间预测。尽管深度学习技术在医学影像分类领域取得了显著成果，但在利用 X 光影像诊断膝骨关节炎的应用方面仍处于探索阶段。

本文针对膝骨关节炎检测精度低的问题，提出一种基于改进 YOLOv8s 的膝骨关节炎自动诊断算法。通过改进的 YOLOv8s 网络结构，提高膝骨关节炎的检测精度。

1 YOLOv8s 网络

YOLOv8s 网络主要由输入端（Input）、主干网络（Backbone）、颈部网络（Neck）以及检测头（Head）4 个部分组成，结构如图 1 所示。

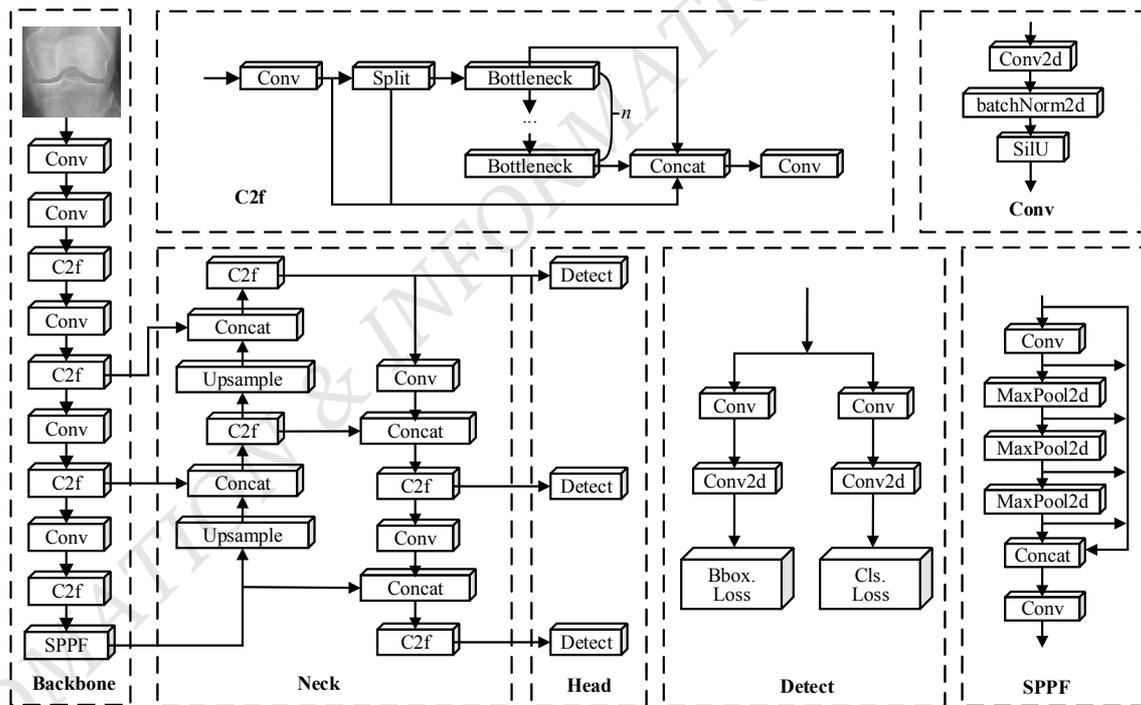


图 1 YOLOv8s 网络结构

- 1) 输入端: 采用 Mosaic 数据增强策略，通过拼接不同的图像，生成更多的训练样本，以丰富数据集。
- 2) 主干网络: 用 C2f 模块替换传统的 C3 模块。C2f 模块的多分支结构增加了梯度回传支路，使 YOLOv8s 网络在保证模型轻量化的同时，获得更加

详细的梯度流信息。

- 3) 颈部网络: 通过整合主干网络提取的多尺度特征，实现对目标的精确识别。其设计思想源于 YOLOv5 的特征金字塔网络（feature pyramid network, FPN）及自底向上的路径聚合网络（path aggregation

network, PAN)，增强了特征融合的效果。

4) 检测头：采用解耦头结构，通过两个独立的分支，消除目标分类与边界框预测回归任务之间的冲突。

2 改进的 YOLOv8s 网络

针对 YOLOv8s 网络对膝关节炎检测精度低的问题，本文对其结构进行如下改进：1) 提出一种改进的卷积块注意力机制模块 (convolutional block atten-

tion module, CBAM)，即采用基于 Transformer 架构的 BiFormer 注意力机制来改进 CBAM，并在主干网络和颈部网络之间引入改进的 CBAM，以提高检测精度；2) 设计一种改进的焦点调制 (Focal Modulation) 模块，用于替换 YOLOv8s 网络中的快速空间金字塔池化 (spatial pyramid pooling fast, SPPF) 模块，以提高多尺度特征的表达能力，进一步提升膝关节炎的检测精度。改进的 YOLOv8s 网络结构如图 2 所示。

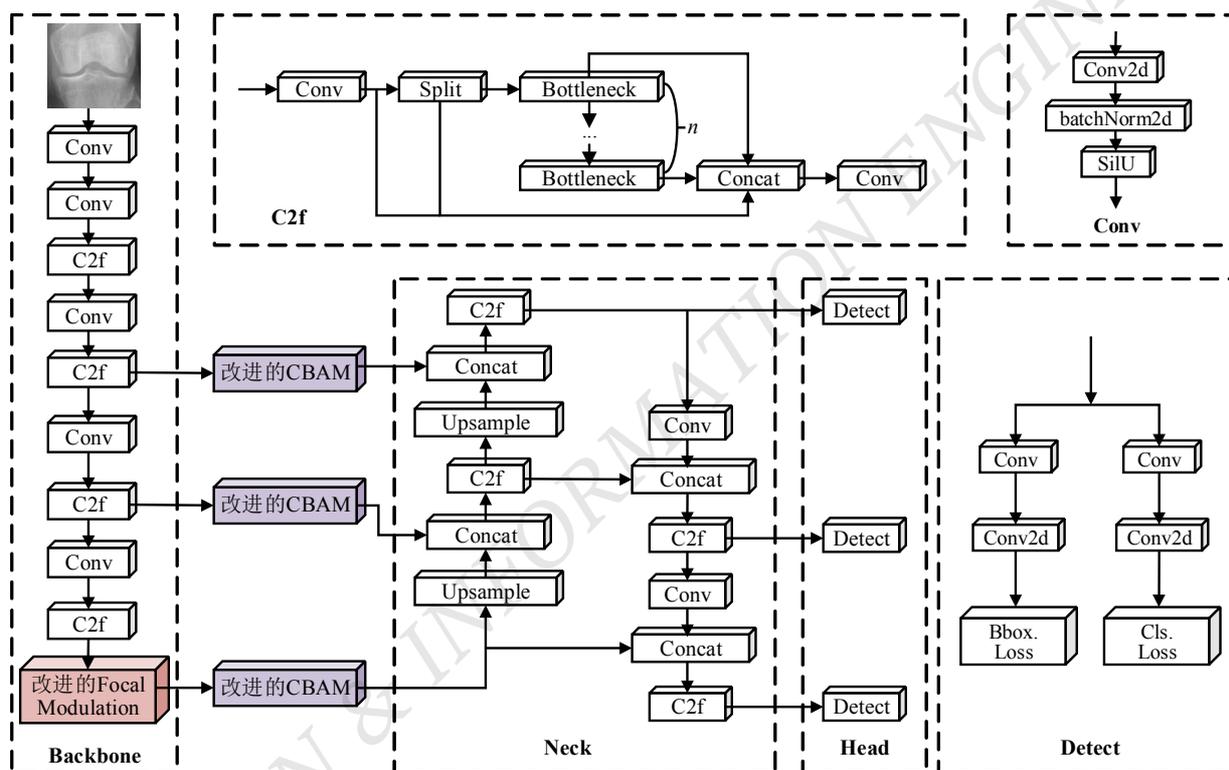


图 2 改进的 YOLOv8s 网络结构

2.1 改进的 CBAM

注意力机制通过权重调整来提升网络对关键特征的识别精度，优化目标检测模型的性能。CBAM 通过通道注意力机制模块 (channel attention module, CAM) 和空间注意力机制模块 (spatial attention module, SAM)，动态地调整特征图中各通道和空间区域的权重，以识别目标的通道和空间位置特征，其结构如图 3 所示。

本文基于 Transformer 架构的 BiFormer 注意力机制来改进 CBAM，改进的 CBAM 结构如图 4 所示。

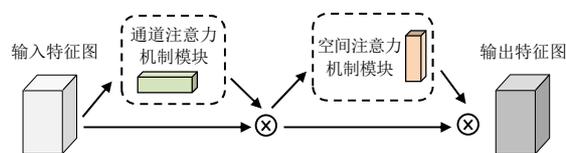


图 3 CBAM 结构

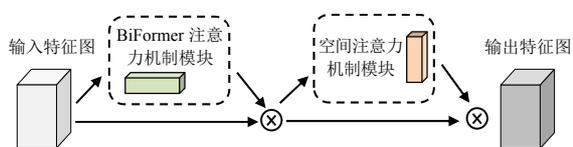


图 4 改进的 CBAM 结构

BiFormer 注意力机制^[13]采用视觉 Transformer 架构，通过双层路由注意力（Bi-level routing attention, BRA）来提升模型的计算效率和性能。首先，在粗粒度区域筛选与查询无关的键值对，减少计算负担；然后，在剩余的候选区域，应用细粒度的令牌对令牌注意力，实现更精确的信息处理。这种方法有效平衡了计算效率与模型性能。BiFormer 的计算分为以下两步：

1) 将图像划分为多个粗粒度区域，先通过自注意力机制模块提取这些区域的特征，再计算每两个粗粒度区域之间的相关性，构建一个关系矩阵；过滤掉最不相关的键值对，仅保留键值最大的 K 个元素；

2) 通过卷积操作进一步细化图像，BiFormer 注意力机制模块主要由 BRA 模块、深度卷积(depthwise convolutio, DWConv)、多层感知机 (multilayer perceptron, MLP) 和归一化层 (layer norm, LN) 构成，其结构如图 5 所示。

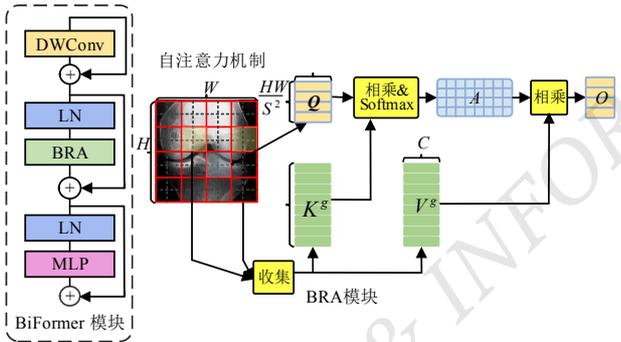


图 5 BiFormer 注意力机制模块结构

BRA 模块是 BiFormer 注意力机制模块的核心组成部分，其通过将输入图像 $[H, W, C]$ 划分为 $S \times S$ 个区域，并对每个区域进行线性映射，生成关键特征表示 \mathbf{Q} 、 \mathbf{K} 、 \mathbf{V} ，计算过程如下：

$$\mathbf{K}^s = \text{gather}(\mathbf{K}, I^r) \quad (1)$$

$$\mathbf{V}^s = \text{gather}(\mathbf{V}, I^r) \quad (2)$$

$$\mathbf{O} = \text{Attention}(\mathbf{Q}, \mathbf{K}^s, \mathbf{V}^s) + \text{LCE}(\mathbf{V}) \quad (3)$$

式中： \mathbf{Q} 为查询矩阵， \mathbf{K} 为键向量， \mathbf{V} 为值向量， I^r 为索引值， \mathbf{O} 为输出特征图， $\text{LCE}(\mathbf{V})$ 为 \mathbf{V} 的局部上下文增强 (local context enhancement, LCE) 值。

2.2 改进的 Focal Modulation 模块

YOLOv8s 网络的 SPPF 模块用于解决不同尺寸物体的检测问题，其通过在主干网络引入多尺度的 SPP，有效地融合了不同尺度的特征信息。虽然 SPPF 模块在多尺度特征融合方面表现出色，但对膝关节关节炎的检测精度不佳。为此，本文利用多尺度线性注意力机制来改进 Focal Modulation 模块，将全局感受野与多尺度学习相结合，提出一种改进的 Focal Modulation 模块，用于替代 SPPF 模块，以提升膝关节关节炎的检测精度。

Focal Modulation 模块^[14]主要由分层上下文、门控聚合、逐元素仿射变换 3 个核心组件组成，结构如图 6 所示。

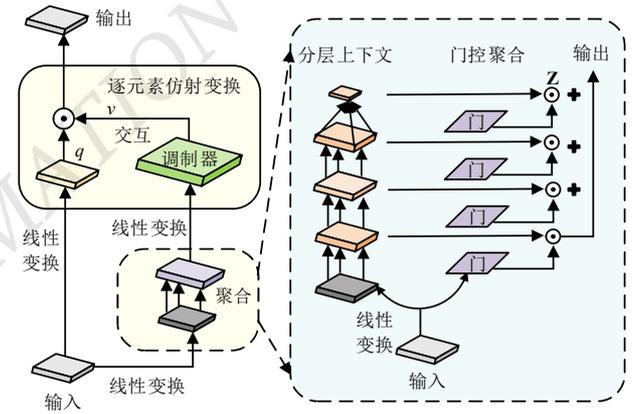


图 6 Focal Modulation 模块结构

分层上下文通过深度卷积有效地提取膝关节图像的上下文信息，捕获膝关节的局部细节和全局整体的结构特征，使网络能够在多个层次上解析图像内容。门控聚合能够筛选出与查询令牌相关的关键信息，以提高对膝关节关节炎病变特征的关注。逐元素仿射变换将聚合后的信息输入到每个查询令牌中，在增强明确指示病变特征的同时，抑制可能干扰判断的特征。

本文通过在 Focal Modulation 模块中添加多尺度线性注意力模块^[15-18]，提高 Focal Modulation 模块的局部信息提取和多尺度学习能力。多尺度线性注意力模块利用线性注意力机制实现全局感受野，并通过矩阵乘法的关联属性，将计算的复杂度从二次降低到线性。改进的 Focal Modulation 模块结构如图 7 所示。

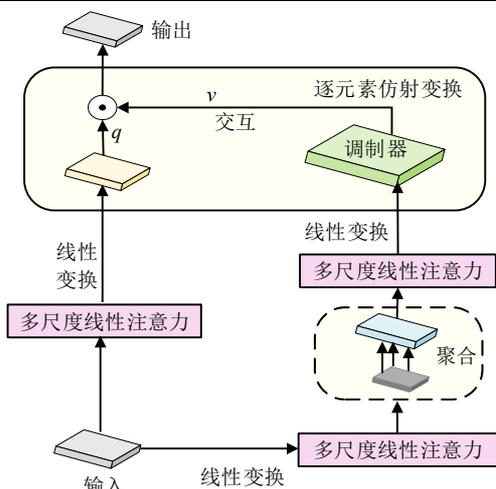


图7 改进的 Focal Modulation 模块结构

3 实验结果与分析

3.1 实验准备

依据 Kellgren-Lawrence (KL) 分级标准, 膝关节炎根据严重程度被划分为 5 个等级 (0~4 级)^[19]。本文基于某医院提供的膝关节 X 光影像, 构建了一个包含 9 786 幅图像的膝关节数据集。首先, 利用 Labelme 软件对这些图像进行手工标注, 即根据膝关节炎的 5 个等级分别标注为 0、1、2、3、4; 然后, 将标注好的数据集按照 9:1 的比例, 随机划分为训练集和测试集。

本实验在 Windows 10 操作系统上进行, 采用 PyCharm 集成开发环境和 Python 3.8 作为开发语言。网络训练采用的深度学习框架为 PyTorch 1.7.1, CUDA 版本为 11.1。硬件配置方面: 处理器为 R7 3700X, 图形处理单元为 Nvidia 3060Ti。

3.2 评价指标

为了评估本文算法在膝关节数据集上的有效性, 采用准确率 (precision)、召回率 (recall) 以及平均精度均值 (mean average precision, mAP) 作为评价指标。

准确率代表模型正确识别正类目标的能力, 计算公式为

$$P_{\text{precision}} = \frac{TP}{TP + FP} \quad (4)$$

式中: TP 为真正例数量, 即正确识别为正类目标的数量; FP 为假正例数量, 即错误识别为正类目标的数量。

召回率反映了模型成功检测到的正类目标数量占样本集中所有实际正类目标数量的比例, 计算公式为

$$R_{\text{recall}} = \frac{TP}{TP + FN} \quad (5)$$

式中: FN 为假负例数量, 即未被识别为正类目标的数量。

AP 可衡量单个类别的平均检测精度。mAP 是平均所有类别的 AP 得到的综合指标, 其值越大表示检测精度越高。

$$AP = \int_0^1 PdR \quad (6)$$

$$mAP = \sum_j^{\text{class number}} Class_AP_j \quad (7)$$

3.3 消融实验

本文通过消融实验验证在 YOLOv8s 模型中引入改进的 CBAM 和改进的 Focal Modulation 模块的有效性, 结果如表 1 所示。

表 1 改进的 YOLOv8s 模型的消融实验

YOLOv8s 模型	改进 CBAM	改进 Focal Modulation	0	1	2	3	4	mAP
✓			0.781	0.403	0.821	0.917	0.906	0.765
✓	✓		0.795	0.455	0.826	0.905	0.942	0.785
✓		✓	0.783	0.484	0.824	0.915	0.903	0.781
✓	✓	✓	0.823	0.514	0.834	0.875	0.911	0.791

由表 1 可知: YOLOv8s 模型在引入改进的 CBAM 后, mAP 提高了 0.020; 在仅引入改进的 Focal Modulation 模块后, mAP 提高了 0.016; 在引入改进的 CBAM +改进的 Focal Modulation 模块后, mAP 提高了 0.026, 证明了改进的 YOLOv8s 模型有效提升了检测精度。

将本文算法与 7 种目标检测算法 (EfficientDet、

RetinaNet、YOLOv4、YOLOv5、YOLOv7、YOLOXs、YOLOv8s) 进行对比实验, 并采用 mAP 作为评价指标, 实验结果如表 2 所示。

由表 2 可知, 本文算法的 mAP 比 YOLOv8s 提高了 0.026, 且优于 EfficientDet、RetinaNet、YOLOv4、YOLOv5、YOLOv7 和 YOLOXs 算法, 证明了本文算法的有效性。

为了更直观地展示本文算法的检测效果, 对膝关节关节炎的检测结果进行可视化处理, 如图 8 所示。

表 2 8 种目标检测算法对比实验结果

算法	0	1	2	3	4	mAP
Efficientdet	0.844	0.481	0.822	0.880	0.905	0.786
Retinanet	0.841	0.442	0.804	0.905	0.871	0.773
YOLOv4	0.807	0.267	0.69	0.833	0.903	0.707
YOLOv5	0.820	0.326	0.797	0.907	0.923	0.755
YOLOv7	0.823	0.475	0.799	0.875	0.905	0.775
YOLOXs	0.797	0.428	0.818	0.928	0.965	0.787
YOLOv8s	0.781	0.403	0.821	0.917	0.906	0.765
本文算法	0.823	0.514	0.834	0.875	0.911	0.791



图 8 改进 YOLOv8s 算法的膝关节关节炎检测结果

由图 8 可知, 本文算法能够有效定位膝关节 X 光影像的病变区域, 并准确地识别出膝关节关节炎的等级。

4 结论

本文提出了一种基于改进 YOLOv8s 的膝关节关节炎自动诊断算法。通过引入改进的 CBAM, 使网络能够更加关注膝关节图像中的关键信息, 从而提高膝关节关节炎检测精度; 引入改进的 Focal Modulation 模块, 提高了网络的多尺度特征表达能力。经实验验证, 该算法有效地提升了膝关节关节炎诊断的准确性。但膝关

节 X 光影像的数据量较少, 未来将通过迁移学习、自监督学习和无监督域自适应方法, 进一步提高网络在小样本情况下的泛化能力。

©The author(s) 2024. This is an open access article under the CC BY-NC-ND 4.0 License (<https://creativecommons.org/licenses/by-nc-nd/4.0/>)

参考文献

- [1] 郭天赐, 陈继鑫, 余伟杰, 等. 人工智能在骨关节炎诊疗中的应用进展[J]. 中国全科医学, 2023, 26(19): 2428-2433.

- [2] 黄霖, 车圳, 李明, 等. 人工智能在骨科疾病诊治中的研究进展[J]. 山东大学学报(医学版), 2023, 61(3): 37-45.
- [3] KATZ J N, ARANT K R, LOESER R F. Diagnosis and treatment of hip and knee osteoarthritis: A review[J]. *Jama*, 2021, 325(6): 568-578.
- [4] TOSUN O, SINCI K A, BAYSAN C, et al. Phenotypic variations in knee osteoarthritis: Insights from MRI and radiographic comparisons[J]. *Skeletal Radiology*, 2024: 1-10.
- [5] KINGER S. Deep learning for automatic knee osteoarthritis severity grading and classification[J]. *Indian Journal of Orthopaedics*, 2024: 1-16.
- [6] SARWINDA D, PARADISA R H, BUSTAMAM A, et al. Deep learning in image classification using residual network (ResNet) variants for detection of colorectal cancer[J]. *Procedia Computer Science*, 2021, 179: 423-431.
- [7] KV S, SELVI S, NANDA P, et al. Deep learning approach to nailfold capillaroscopy based diabetes mellitus detection[J]. 2022 *International Journal of Online and Biomedical Engineering (iJOE)*, 2022, 18(6): 95-109.
- [8] JIANG H, YIN Y, ZHANG J, et al. Deep learning for liver cancer histopathology image analysis: A comprehensive survey [J]. *Engineering Applications of Artificial Intelligence*, 2024, 133: 1-22.
- [9] CALDERARO J, SERAPHIN T P, LUEDDE T, et al. Artificial intelligence for the prevention and clinical management of hepatocellular carcinoma[J]. *Journal of Hepatology*, 2022, 76(6): 1348-1361.
- [10] ZHU P, WANG C, SUN Z, et al. Segmentation of liver cancer pathology images based on multi-scale feature fusion[C]// *Proceedings of 2021 Chinese Intelligent Systems Conference: Volume III*. Springer Singapore, 2022: 596-605.
- [11] DAIL L, SHENG B, CHEN T, et al. A deep learning system for predicting time to progression of diabetic retinopathy[J]. *Nature Medicine*, 2024, 30(2): 584-594.
- [12] SANDLER M, HOWARD A, ZHU M, et al. Mobilenetv2: Inverted residuals and linear bottlenecks[C]// *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018: 4510-4520.
- [13] ZHU L, WANG X, KE Z, et al. Biformer: Vision transformer with bi-level routing attention[C]// *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023: 10323-10333.
- [14] YANG J, LI C, DAI X, et al. Focal modulation networks[J]. *Advances in Neural Information Processing Systems*, 2022, 35: 4203-4217.
- [15] CAI H, LI J, HU M, et al. Efficientvit: Multi-scale linear attention for high-resolution dense prediction[J]. *arXiv preprint arXiv:2205.14756*, 2022.
- [16] 来春庆, 黄勇, 朱喆, 等. 基于特征融合多尺度卷积网络的光伏组件红外图像故障诊断[J]. *机电工程技术*, 2023, 52(9): 145-151.
- [17] 周庆辉, 葛馨远, 孙峥, 等. 融合少样本学习与注意力端到端网络的小目标在线检测研究[J]. *机床与液压*, 2024, 52(17): 130-135.
- [18] 赵彦龙, 钟震宇. 基于注意力机制的异常行为识别方法[J]. *自动化与信息工程*, 2023, 44(3): 17-22.
- [19] QUICKE J G, CONAGHAN P G, Corp N, et al. Osteoarthritis year in review 2021: Epidemiology & therapy[J]. *Osteoarthritis and Cartilage*, 2022, 30(2): 196-206.

作者简介:

肖军峰, 男, 1982年生, 高级政工师, 主要研究方向: 目标检测。E-mail: gdxiaojf@petrochina.com.cn

丁鹏(通信作者), 男, 1992年生, 博士研究生, 副教授, 主要研究方向: 计算机视觉、深度学习和视觉导航。E-mail: 202460055@ecut.edu.cn

(上接第 19 页)

作者简介:

王楚晗, 男, 2001年生, 在读硕士研究生, 主要研究方向: 激光定向能量沉积。E-mail: the_yshdl@163.com

齐超琪, 女, 1994年生, 硕士研究生, 工程师, 主要研究方向: 激光选区融化。E-mail: qichaoqi_nuaa@126.com