

闫建红, 张雷. 基于多实例冗余抑制的骶髂关节疾病分类方法[J]. 智能计算机与应用, 2026, 16(2): 171-176. DOI: 10.20169/j. issn. 2095-2163. 25101403

# 基于多实例冗余抑制的骶髂关节疾病分类方法

闫建红, 张雷

(太原师范学院 计算机科学与技术学院, 山西 晋中 030619)

**摘要:** 骶髂关节疾病的精准分类对临床诊疗具有重要意义,但现有基于CT影像的计算机辅助诊断方法存在数据细粒度标注困难、CT连续切片间信息冗余,以及关键病理特征在空间上分布不均匀等问题。为此,文章提出一种基于多实例学习的骶髂关节疾病分类模型,在仅有患者级标注的CT数据下实现疾病分类。设计余弦相似度判别模块,通过滑动窗口动态剔除高度相似的冗余切片。构建解剖引导的高斯加权模块,依据骶髂关节病变集中于中段切片的先验知识,对关键区域特征赋予更高权重。利用通道与空间并行注意力机制,增强病灶相关特征的表达能力。实验结果表明,与传统MIL方法相比,该模型能有效降低冗余计算,提升可靠性,为简单标注、冗余度高、特征分布不均的医学影像分类任务提供了一种有效的解决方案。

**关键词:** 图像分类; 骶髂关节; 多实例学习; 余弦相似度; 高斯加权

中图分类号: TP391.4

文献标志码: A

文章编号: 2095-2163(2026)02-0171-06

## Classification method for sacroiliac joint diseases based on multi-instance learning and redundancy suppression

YAN Jianhong, ZHANG Lei

(College of Computer Science and Technology, Taiyuan Normal University, Jinzhong 030619, Shanxi, China)

**Abstract:** The precise classification of sacroiliac joint diseases is of great significance for clinical diagnosis and treatment. However, existing computer-aided diagnosis methods based on CT images have problems such as the difficulty of fine-grained data annotation, information redundancy between consecutive CT slices, and uneven spatial distribution of key pathological features. To address these issues, this paper proposes a sacroiliac joint disease classification model based on multi-instance learning, which can achieve disease classification with only patient-level annotations of CT data. A cosine similarity discrimination module is designed to dynamically eliminate highly similar redundant slices through a sliding window. An anatomy-guided Gaussian weighting module is constructed, which assigns higher weights to features in key regions based on the prior knowledge that sacroiliac joint lesions are concentrated in the middle slices. A parallel attention mechanism of channels and space is utilized to enhance the expression ability of lesion-related features. Experimental results show that compared with traditional MIL methods, this model can effectively reduce redundant computations, improve reliability, and provide an effective solution for medical image classification tasks with simple annotations, high redundancy, and uneven feature distribution.

**Key words:** image classification; sacroiliac joint; multi-instance learning; cosine similarity; Gaussian weighting

## 0 引言

骶髂关节疾病的早期精准诊断对临床干预及预后评估至关重要。计算机断层扫描(CT)因其高空间分辨率成为评估骶髂关节结构性病变的金标准<sup>[1]</sup>。然而,传统人工阅片存在主观性强、效率低

下等局限,亟需发展自动化智能诊断系统。近年来,随着深度学习、尤其是卷积神经网络在图像处理领域的广泛应用,学者们在骶髂关节疾病自动分类领域进行了很多研究,取得了很大进展。2019年,Shenkman等学者<sup>[2]</sup>提出了首个完全根据CT影像自动化检测和分级骶髂关节炎的方法,使用U-Net分

**基金项目:** 山西省科技战略研究专项重点项目(202304031401011); 山西省重点研发计划项目(202102010101008); 山西省研究生精品教学案例项目(2024AL27)。

**作者简介:** 闫建红(1972—),女,博士,教授,主要研究方向:机器学习,计算机视觉。Email: zhanglei\_raynell@163.com; 张雷(1995—),男,硕士研究生,主要研究方向:深度学习,图像分类。

收稿日期: 2025-10-14



$F$ 。最后,结合空间注意力(SA)与通道注意力(CA)的并行注意力机制,增强关键特征,实现精准的病灶表征。

## 2.1 余弦相似度判别模块

在CT影像中,相邻层的切片间隔往往小于1 mm,解剖结构高度相似,特征表示十分相近(见图2)。红框标注了以灰度差异计算、差异度超过10%以上的部分。

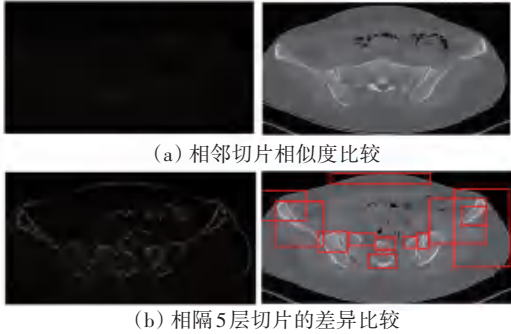


图2 不同层间切片差异对比

Fig. 2 Comparison of slice differences between different layers

图2(a)展示了相邻切片的差异度,由于相邻切片过于相似,导致无法计算出显著差异,因此没有红框标注的部分;图2(b)为间隔5层切片之间的差异度,切片间具有明显差异。

将全部切片特征纳入计算的成本很高,但对模型实际的提升效果却并不明显。为减少相似切片可能造成的冗余计算,可以使用余弦相似度判别方法<sup>[13-14]</sup>,通过计算2个向量夹角的余弦值衡量二者之间的相似度,其值越接近1则相似度越高。将相似度高于阈值的切片从计算中去除,能够大幅减少模型的计算量。

为了在减少冗余计算的同时多保留对正确分类有利的切片,可以采用滑动窗口的余弦相似度判别方法。将相邻的 $k$ 层切片特征放到统一窗口内构成特征矩阵 $F$ ,然后计算滑动窗口内(窗口大小 $k=5$ ,步长 $s=5$ )特征的余弦相似度,计算公式如下:

$$\cos(O_i, O_m) = \frac{O_i \cdot O_m}{\|O_i\| \cdot \|O_m\|} \quad (1)$$

其中, $O_i$ 表示第 $i$ 个切片的特征向量, $O_m$ 表示中心基准切片。 $\cos(O_i, O_m)$ 的值越接近1,切片 $i$ 和窗口中心的基准切片在特征空间中方向越一致,相似度越高。冗余切片判别方法如下:

$$\text{Redundancy} = \begin{cases} 1, & \text{if } \cos(O_i, O_m) > \tau \\ 0, & \text{otherwise} \end{cases} \quad (2)$$

设置阈值 $\tau=0.98$ ,当 $\text{Redundancy}=1$ 时,当前

切片与基准切片高度相似,舍弃该切片;当 $\text{Redundancy}=0$ 时,当前切片与基准切片差异度超过阈值,保留该切片。

## 2.2 高斯加权模块

在骶髂关节CT中,关节缝隙的病灶表征是判断疾病类别的重要依据。不同段的骶、髋关节CT切片如图3所示。图3中,CT中的前、后段切片关节间隙不可见,而中段切片则更能清晰展现骶髂关节缝隙和关节表面的硬化和侵蚀。

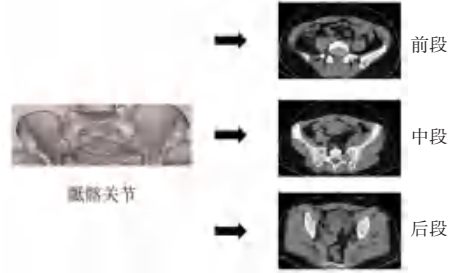


图3 不同段的骶髂关节CT切片

Fig. 3 CT slices of the sacroiliac joint at different segments

CT图像在空间上存在连续性,能够显示骶髂关节的区域随切片在空间中的位置变化,因此可以根据轴向位置对切片赋予权重。将有利于分类的中段切片赋予更高权重,而两侧切片则随位置变化权重逐渐递减。这种权重分布十分符合高斯分布的特点,因此本文采用空间位置引导的高斯加权机制对切片特征加权,数学公式如下:

$$w_i = \exp\left(-\frac{(i-\mu)^2}{2\sigma^2}\right) \quad (3)$$

其中, $i$ 表示第 $i$ 层切片; $\mu$ 表示CT中段中心坐标; $\sigma$ 控制权重衰减速度,为了保证解剖合理性,通过设置 $\mu$ 与 $\sigma$ 的最大值与最小值作为参数约束,使其始终在解剖的合理区间。将位置权重 $w_i$ 与特征拼接,输入全连接层生成融合视觉特征的联合表示。

完全基于注意力机制的模型过于优先考虑显著切片的特征,从而忽略实例的真实标签和特征之间的因果因素,导致非因果实例得到更高的注意力分数。高斯加权机制能够很好地平抑注意力机制潜在的风险,通过先验知识使模型更关注关键特征。

## 2.3 并行注意力模块

注意力机制是一种模仿人类“聚焦”能力的技术,能够让模型在处理信息时自动挑选出对当前任务最关键的部分,从而提升表达和推理能力。为了使网络更关注CT图像中骶髂关节疾病的病灶区域,强化对关键通道的响应,抑制冗余通道,本文引入了通道注意力(CA)和空间注意力(SA)并行的双

注意力结构<sup>[15-16]</sup>(见图4)。

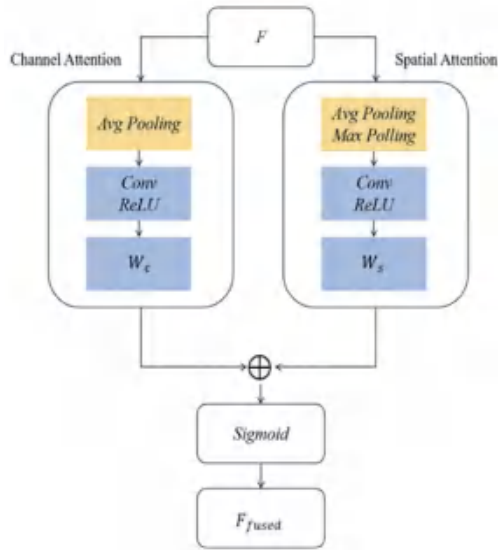


图4 并行注意力结构

Fig. 4 Parallel attention structure

特征矩阵  $F$  分别进入通道和空间注意力模块。其中,通道注意力通过全局平均池化压缩空间信息,生成通道重要性权重,增强关键通道、抑制低贡献通道;空间注意力融合通道维度的最大池化和平均池化,生成空间权重图,突出重要区域。

### 3 实验与结果

#### 3.1 数据集与预处理

实验数据集来源于山西某大型三甲医院,数据源自2023年至2024年期间采用CT机收集的77例骶髂关节疾病样本,包括骶髂关节炎20例、退行性变20例、致密性骨炎19例和未见异常18例,共4个类别,11708张切片。所有样本标注均由2位专业的影像科医师完成。

实验对图像进行了预处理。方法如下。

(1)训练分割模型:在一部分标注骶髂关节缝隙及关节缝隙左右拓宽1cm区域的数据上,以U-Net为基础训练一个自动分割模型。

(2)分割图像:用训练好的U-Net分割图像中的骶髂关节缝隙区域,使模型进一步聚焦关键特征。

(3)图像裁切:裁切U-Net分割后图像存在的部分背景遮罩,将原始大小为512×512像素的图像裁切为224×224像素,减少冗余计算的同时使其适应模型输入。

#### 3.2 实验设置

##### 3.2.1 实验环境

实验基于Linux操作系统,GPU型号为Nvidia GeForce RTX3090,深度学习框架为Pytorch,CUDA

版本为CUDA11.0。优化器为能够自适应调整学习率的AdamW,损失函数为标签平滑的交叉熵损失函数,初始学习率为0.0001,总训练迭代次数为160轮。衡量模型性能的指标包括准确率、召回率和精确率,以及参数数量和计算量(GFLOPs)。

##### 3.2.2 实验参数选择

在余弦相似度判别模块中,滑动窗口大小  $k$  的取值可以作为一个超参数进行调整。在本实验中,分别取  $k = 1, 3, 5, 7, 9$  时对比模型性能,实验结果见表1。

表1 不同  $k$  值下模型性能对比

Table 1 Comparison of model performance under different  $k$  values

$k$	准确率	召回率	精确率	GFLOPs
1	0.740	0.714	0.681	3.62
3	0.734	0.710	0.679	1.35
5	0.733	0.709	0.677	0.94
7	0.713	0.703	0.663	0.68
9	0.704	0.695	0.652	0.48

其中,  $k = 1$  表示不进行冗余筛选的对照组,计算量为GFLOPs。经实验验证,在  $k = 3$  或  $k = 5$  时,模型的计算量显著低于对照组,且相比其他窗口大小,模型的性能下降不明显。 $k = 3$  时,计算量下降62.7%;  $k = 5$  时,计算量下降74.0%。结合实验结果可知,  $k = 5$  时模型能更好地平衡性能与计算量。

##### 3.3 对比实验

为了验证模型的分类性能,本文在骶髂关节数CT扫描数据集上对比不同MIL方法。比较的方法包括:基于最大池化的Maxpooling-MIL,基于注意力机制的ABMIL<sup>[17]</sup>,基于多尺度特征的DSMIL<sup>[18]</sup>,以及对小型数据集敏感但计算复杂度较高的PSA-MIL<sup>[19]</sup>、基于聚类约束注意力的CLAM<sup>[20]</sup>和结合自注意力以及位置编码的TransMIL<sup>[21]</sup>。对比结果见表2。表2中,加粗数据为最好结果。

表2 骶髂关节数据集上各方法分类性能

Table 2 Classification performance of each method on the sacroiliac joint dataset

模型	准确率	召回率	精确率	参数数量/M
Maxpooling-MIL	0.740	0.714	0.681	<b>21.8</b>
ABMIL <sup>[17]</sup>	0.740	0.753	0.724	23.5
DSMIL <sup>[18]</sup>	0.753	0.763	0.773	24.2
PSA-MIL <sup>[19]</sup>	0.766	0.779	0.780	25.1
CLAM <sup>[20]</sup>	0.731	0.774	0.787	24.8
TransMIL <sup>[21]</sup>	0.727	0.736	0.747	30.3
本文	<b>0.818</b>	<b>0.805</b>	<b>0.830</b>	24.5

Maxpooling-MIL 直接取所有实例特征中的最大值聚合分类,而基于注意力的方法可以通过可学习的注意力网络为每个实例分配权重,因此其参数量更少,但性能更差。TransMIL 参数量相对较大,在样本较少时容易出现过拟合,导致泛化性差,实际效果反而不如轻量化模型。本文提出的方法相较 Maxpooling-MIL 准确率提高 7.8%,同时参数量只高出了 2.7 M。其他主要基于注意力的 MIL 方法中,与性能表现最好的 PSA-MIL 相比,提出的方法准确率提高了 5.2%,召回率提高了 2.6%,精确率提高了 5%,参数量少了 0.6 M。

实验结果表明,提出的模型在增加少量参数的条件下实现了较高的性能。

### 3.4 消融实验

为验证余弦相似度判别和高斯加权机制的有效性,本文在骶髂关节 CT 扫描数据集上进行了消融实验。实验以 Max-Pooling 作为基准方法,增加不同的模块组合来验证不同模块的作用,实验结果见表 3。

表 3 骶髂关节 CT 数据集上消融实验结果

Table 3 Ablation experiment results on the sacroiliac joint CT dataset

模型变体	准确率	召回率	精确率	GFLOPs
Baseline	0.740	0.714	0.681	3.62
+余弦相似度	0.751	0.793	0.767	<b>0.94</b>
+高斯加权	0.784	0.815	0.792	3.67
+并行注意力	0.756	0.738	0.724	3.67
+余弦相似度+高斯加权	0.779	0.782	0.795	1.01
+高斯加权+并行注意力	0.805	0.795	0.827	3.67
+余弦相似度+并行注意力	0.753	0.763	0.773	1.01
+余弦相似度+高斯加权+并行注意力	<b>0.818</b>	<b>0.805</b>	<b>0.830</b>	1.01

在引入余弦相似度模块后,模型的准确率下降 0.7%,但计算量大幅降低到基线模型的 26%,提高模型训练效率。加入高斯加权后与基线相比,准确率提高 8.6%,且不增加额外计算量。并行注意力机制在小幅增加计算量的同时提高 1.6%的准确率。实验表明,各模块均能够提高模型的性能或减少计算量,提高模型训练效率。

## 4 结束语

本文提出一种基于多实例学习的骶髂关节疾病

CT 图像分类方法。利用骶髂关节数据集冗余数据多、特征分布不均匀的特点,通过三重优化方案抑制冗余计算,提高了模型训练效率和性能:

(1) 余弦相似度度量实现实例级特征筛选,能够在小幅缩减性能的情况下大幅降低模型计算量。

(2) 结合解剖特征和高斯分布对实例加权,能够合理利用数据特征分布特点增强模型分类性能。

(3) 结合空间-通道注意力的并行注意力机制实现高维特征的跨维度交互增强。提出的方法大幅降低了模型计算量,提升了模型训练效率,避免大量高度相似切片可能导致的过拟合问题。

## 参考文献

- [1] 展晓梅. MRI、CT、X 线联合诊断早期强直性脊柱炎骶髂关节病变的价值分析[J]. 中国社区医师, 2023, 39(30): 107-109.
- [2] SHENKMAN Y, QUTTEINEH B, JOSKOWICZ L, et al. Automatic detection and diagnosis of sacroiliitis in CT scans as incidental findings[J]. Medical Image Analysis, 2019, 57: 165-175.
- [3] BERGHE T V D, BABIN D, CHEN M, et al. Neural network algorithm for detection of erosions and ankylosis on CT of the sacroiliac joints: Multicentre development and validation of diagnostic accuracy[J]. European Radiology, 2023, 33(11): 14.
- [4] ZHANG Ke, LUO Guibo, LI Wenjuan. Automatic image segmentation and grading diagnosis of sacroiliitis associated with AS using a deep convolutional neural network on CT images[J]. Journal of Digital, 2023, 36(5): 2025-2034.
- [5] 杜涛, 闫建红. DI-MobileNet: 基于轻量化网络的骶髂关节炎识别方法[J]. 智能计算机与应用, 2025, 15(2): 138-143.
- [6] 郑金龙, 韩萍, 史河水, 等. 骶髂关节病变的 CT 诊断[J]. 临床放射学杂志, 2001(6): 456-459.
- [7] 庄娘妥, 武兴杰, 曾文彦, 等. 强直性脊柱炎累及骶髂关节的 X 线与 CT 诊断[J]. 实用医技杂志, 2007(23): 3126-3128.
- [8] 谢卓恒, 伊鸣, 黄新瑞. 多实例学习在医学图像分析中的应用进展[J]. 集成技术, 2025, 14(2): 24-32.
- [9] RAHMAN M. Artificial Intelligence (AI) and Machine Learning (ML) in medical imaging informatics towards diagnostic decision making[EB/OL]. (2020-10-22). <https://www.fda.gov/meida/151482/download>.
- [10] TANG Wenhao, HUANG Sheng, FANG Heng, et al. Multiple instance learning framework with masked hard instance mining for whole slide image classification[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). Piscataway, NJ: IEEE, 2023: 4078-4087.
- [11] VENTE C D, GINNEKEN B V, HOYNG C B, et al. Uncertainty-aware multiple-instance learning for reliable classification: Application to optical coherence tomography[J]. Medical Image Analysis, 2024, 97: 103259.
- [12] LIN Tiancheng, XU Hongteng, YANG Canqian, et al. Interventional multiinstance learning with deconfounded instancelevel prediction[C]//Proceedings of the AAAI Conference on Artificial Intelligence. Palo Alto: AAAI Press, 2022: 1601-1609.
- [13] ANAP J, HAMBAR P, PATEL H, et al. Reverse image querying

- [J]. International Journal of Research in Applied Science and Engineering Technology, 2022, 2022:44822.
- [14] SUN Peixia, LIN Huiting, LUO Tao. Learning discriminative CNN features and similarity metrics for image retrieval [C]// Proceedings of 2016 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC). Piscataway, NJ: IEEE, 2016:15.
- [15] EFTEKHARI A, SAMADZADEGAN F, JAVAN F D. Building change detection using the parallel spatial-channel attention block and edge-guided deep network [J]. International Journal of Applied Earth Observation Geoinformation, 2023, 117:18.
- [16] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional block attention module [M]// FERRARI V, HEBERT M, SMINCHISESCU C, et al. Computer Vision - ECCV 2018. Cham: Springer, 2018:3-19.
- [17] ILSE M, TOMCZAK J, WELLING M. Attention-based deep multiple instance learning [C]// Proceedings of the 35<sup>th</sup> International Conference on Machine Learning. Stockholm: PMLR, 2018:2127-2136.
- [18] LI Bin, LI Yin, ELICEIRI K W. Dual-stream multiple instance learning network for whole slide image classification with self-supervised contrastive learning [C]// Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Piscataway, NJ: IEEE, 2021:14318-14328.
- [19] PELED S, MARUVKA Y E, FREIMAN M. PSA-MIL: A probabilistic spatial attention-based multiple instance learning for whole slide image classification [J]. arXiv preprint arXiv, 2503.16284, 2025.
- [20] LU M Y, WILLIAMSON D F K, CHEN T Y, et al. Data-efficient and weakly supervised computational pathology on whole-slide images [J]. Nature Biomedical Engineering, 2021, 5:555-570.
- [21] SHAO Zhucheng, BIAN Hao, CHEN Yang, et al. TransMIL: Transformer based correlated multiple instance learning for whole slide image classification [J]. arXiv preprint arXiv, 2106.00908, 2021.