

基于多序列MRI的3D关系注意力网络预测HLA-B27阴性中轴性脊柱关节病

邹青青¹, 王梦虹², 陆紫箫², 赵英华², 冯前进¹

¹南方医科大学生物医学工程学院//广东省医学图像处理重点实验室, 广东 广州 510515; ²南方医科大学第三附属医院(广东省骨科医院)放射科, 广东 广州 510630

摘要:目的 建立一种新的3D多序列关系注意力网络,通过探索不同磁共振成像(MRI)序列图像的互补和相关信息,提升对人类白细胞抗原(HLA)-B27阴性中轴性脊柱关节病(axSpA)的诊断性能。**方法** 回顾性收集2010年1月~2021年8月南方医科大学第三附属医院(TAH组)的375例和南海医院(NHH组)的49例HLA-B27阴性参与者(TAH组:164例axSpA,211例非axSpA;NHH组:27例axSpA,22例非axSpA)的两种参数MRI,包括T1加权图像(T1WI)和压脂序列MRI(FS-MRI),以及相关临床数据。提出一个基于多序列MRI的3D关系注意力网络MSFANet,实现对HLA-B27阴性axSpA与非axSpA的自动鉴别诊断。MSFANet由一个浅层共享特征模块和一个类感知特征学习模块组成,其中类感知特征学习模块采用3D多序列关系注意力机制对多序列MRI特征进行细化和融合。提出一种混合损失函数,通过学习不同支路的损失权重系数来提升MSFANet对序列特征的识别能力,从而增强分类性能。**结果** 实验结果表明,MSFANet优于其它几种最先进的多序列融合算法,其中内部验证集上的AUC、准确度、敏感度和特异度分别达到了0.840,77.93%,83.70%和70.29%,独立外部验证集(NHH)上的上述性能分别达到了0.783,74.47%,82.43%和70.40%。各项差异均具有统计学意义($P<0.05$)。此外,消融实验显示,相同框架下,MSFANet的性能优于基于单序列MRI的模型,证实了融合多序列MRI的有效性和必要性。深度可视化技术显示MSFANet在分类过程中集中于学习图像异常区域的信息。**结论** 本研究成功构建基于多序列MRI的3D深度神经网络对HLA-B27阴性axSpA和非axSpA进行鉴别诊断,并验证了采用多序列关系注意力机制对提升网络分类性能的有效性。

关键词:中轴性脊柱关节病诊断;HLA-B27阴性;磁共振成像;3D多序列关系注意力机制;混合损失

Multi-sequence relation attention network for diagnosing HLA-B27-negative axial spondyloarthritis

ZOU Qingqing¹, WANG Menghong², LU Zixiao², ZHAO Yinghua², FENG Qianjin¹

¹School of Biomedical Engineering, Southern Medical University//Guangdong Provincial Key Laboratory of Medical Image Processing, Guangzhou 510515, China; ²Academy of Orthopedics, Third Affiliated Hospital of Southern Medical University, Guangzhou 510630, China

Abstract: Objective To develop a new 3D multi-sequence relation attention network for exploring the complementary and correlation information of different magnetic resonance imaging (MRI) modalities and improving the diagnostic performance of HLA-B27-negative axial spondyloarthritis (axSpA). **Methods** We retrospectively collected T1-weighted imaging (T1WI) and fat suppression MRI (FS-MRI) data and clinical data of 375 HLA-B27-negative patients from the Third Affiliated Hospital of Southern Medical University (including 164 axSpA and 211 non-axSpA patients) and 49 patients from Nanhai Hospital (including 27 axSpA and 22 non-axSpA patients) between January, 2010 and August, 2021. A 3D relation attention network MSFANet based on multi-sequence MRI was used for automatic diagnosis of axSpA against non-axSpA in these patients. MSFANet consisted of a shallow shared feature learning module and a class-aware feature learning module, and latter module used a 3D multi-sequence relation attention mechanism to refine and fuse multi-sequence MRI features. A hybrid loss function was used to enhance the recognition ability of MSFANet by learning the loss weight coefficients of different branches to improve the classification performance. **Results** The experimental results demonstrated that MSFANet outperformed several state-of-the-art fusion algorithms ($P<0.05$) with AUC, accuracy, sensitivity, and specificity of 0.840, 77.93%, 83.70%, and 70.29% in the internal validation set, and of 0.783, 74.47%, 82.43% and 70.40% in the independent external validation set, respectively. The ablation studies showed that under the same architecture, the fusion model was superior to single-sequence models, which confirmed the effectiveness and necessity of fusing multi-sequence MRI. The visualization results demonstrated that MSFANet could focus on learning information from abnormal areas on MRI during the classification. **Conclusion** We

successfully constructed a 3D deep neural network based on multi-sequence MRI for differential diagnosis of HLA-B27-negative axSpA against non-axSpA and verified the effectiveness of the multi-sequence relation attention mechanism for promoting classification performance of the network.

Keywords: axSpA diagnosis; HLA-B27 negative; magnetic resonance imaging; 3D multi-sequence relation attention; hybrid loss

收稿日期:2023-08-22

基金项目:国家自然科学基金(81871510,82203200)

Supported by National Natural Science Foundation of China (81871510, 82203200).

作者简介:邹青青,在读硕士研究生, E-mail: danica_zou@163.com

通信作者:冯前进,博士,教授,博士生导师, E-mail: fengqj99@smu.edu.cn

中轴性脊柱关节病(axSpA)是一种炎症性慢性疾病,其发病率为0.3%~1.4%,可导致慢性背痛和残疾^[1,2]。根据国际脊柱关节炎学会(ASAS)分类标准(2009),人类白细胞抗原(HLA)-B27阳性或在MRI上可观察到骶髂关节的异常影像学改变,如软骨下骨髓水肿、骨侵蚀、骨硬化和僵硬是确诊axSpA的必要条件^[1,3,4]。然而,HLA-B27阴性的发生率在axSpA人群中可能高达42%~57%^[5],且HLA-B27阴性axSpA的MRI表现比HLA-B27阳性患者具有非典型的对称性病损、较少的放射学损伤以及边缘联合韧带,导致HLA-B27阴性axSpA容易被延迟诊断^[6]。此外,放射科医生对骶髂关节MRI的阅读和解释与个人经验有关,这可能进一步延误对HLA-B27阴性axSpA患者的最佳治疗时间^[7]。如果患者错过最佳的早期干预,持续的疾病活动可能会导致不可逆转的结构异常、功能受损、严重残疾、心血管风险增加和精神障碍等^[8-10]。因此,迫切需要开发有效工具来自动解释HLA-B27阴性患者的MRI数据并协助识别疑似axSpA。

近年来,放射组学作为一种很有前途的图像异质性研究定量方法,允许从感兴趣区域提取高维且有价值的图像特征信息^[11]。放射组学目前已经在肺结节^[12]、消化道癌^[13]和乳腺癌^[14]等疾病研究中得到广泛应用,这也引起了人们对axSpA分类的研究兴趣。例如,Tenorio等^[15]基于MRI图像进行了放射组学分析,将基于纹理的定量生物标志物与中轴型和外周亚型的骶髂关节炎联系起来,取得了一定效果。Ye等^[16]使用放射组学来研究中轴性脊柱炎和非中轴性脊柱炎的影像学生物标志物,证明了临床放射组学列线图可以提高鉴别axSpA的性能,从而有助于临床决策过程。但标准放射组学研究包含数据预处理、ROI勾画、特征提取、特征降维和模型构建等多个步骤,过程比较繁琐;而且放射组学对特征提取参数较敏感,且无法根据标签自动学习和更新特征,因此其预测能力往往有限,仍具有较大的改进空间。

随着深度学习在医学图像分析领域取得的巨大成功和快速发展,基于卷积神经网络的分类方法也被应用于中轴性脊柱炎的分类^[17-19]。Bordner等^[20]基于MRI使用Mask-RCNN进行骶髂关节的检测和骨髓水肿的分类,与专家相比有良好的相关性;Bressem等^[21]应用3D ResNet101网络从T1WI和STIR图像中提取特征并进行连接,用于对中轴性脊柱炎的炎症变化和结构变化检测,也获得了较高的准确率。但是,目前还没有采用深度学习方法对HLA-B27阴性axSpA和非axSpA进行鉴别诊断的相关研究报道。

联合采用多序列图像有助于提高深度学习任务性能,也与临床医生诊断习惯一致。目前深度学习技术可以在原始数据级、特征级和决策级上进行多序列融合^[22,23]。例如:Ge等^[24]提出了TripleNet融合基于VGG16网络提

取的两种不同MRI序列的特征,用于皮肤镜图像和临床图像中的黑色素瘤分类。Vielzeuf等^[25]则提出了一种新的特征融合方法,针对每种MRI序列构建一个独立的深度卷积网络进行特征学习,同时使用可学习参数加权的方法融合多序列特征。但这些方法对于充分利用和融合多序列信息,仍存在如下两个主要挑战:(1)独立的特征提取过程意味着网络仍然只能基于单序列输入进行学习,因此提取特征的表征能力有限;(2)现有多序列特征融合一般通过拼接层完成,这种简单的连接无法充分解析不同序列特征之间的异质性,这可能导致融合后的特征内部存在较大语义差距,从而降低整体表现。为了解决这两个问题,本研究提出了一种多序列融合注意力网络MSFANet。该网络由共享特征学习模块和类感知特征学习模块两部分组成。共享特征学习模块将输入的多序列MRI转换为图像空间特征。类感知特征学习模块由两个独立分支和一个融合分支组成,其中独立分支用于分别提取特定序列的鉴别特征,融合分支采用多序列融合注意力模块MSFA学习鉴别不同序列特征之间的相关性,并自适应地进行特征融合。此外,混合损失函数利用可学习的损失权重系数平衡不同分支对分类任务的性能。

本研究提出的结合交叉注意力和混合损失函数的多序列融合框架有助于对HLA-B27阴性axSpA和非axSpA的鉴别诊断。本研究的模型有潜力帮助临床放射医生更早、更准确地检测到HLA-B27阴性axSpA,减少临床中的误诊或漏诊,从而辅助提升axSpA患者的治疗、预后和疾病管理。

1 资料和方法

1.1 实验数据

该研究得到了南方医科大学第三附属医院区域伦理审查委员会的批准(IRB:2020-伦理审查-002)。所有访问的患者数据均提供了书面知情同意书,明确说明所有收集的信息,包括MRI和临床数据,将用于研究者发表。

数据收集:本研究为回顾性研究。分别收集了南方医科大学第三附属医院(TAH)和南海医院(NHH)招募的375例和49例HLA-B27阴性的axSpA(符合2009年版本的ASAS分类标准)和非axSpA患者的双侧骶髂关节MRI数据,其中TAH包括164例axSpA患者和211例非axSpA患者,NHH包括27例axSpA患者和22例非axSpA患者(表1)。数据排除标准如下:缺乏两个以上的临床特征;缺乏T1WI序列或FS-MRI序列;MRI质量差,严重阻碍了双侧骶髂关节的观察。

1.2 数据获取及预处理

所有患者的MRI检查均在1.5T和3.0T MR的磁共振扫描仪下进行采集。两位具有3年和2年经验的肌肉

表1 Non-axSpA的疾病亚型

Tab.1 Detailed disease subtypes in the non-axSpA population [n (%)]

Disease subtypes	TAH (n=211)	NHH (n=22)
Low back pain populations	60 (28.44%)	6 (27.27%)
Non-specific sacroiliitis	46 (21.80%)	11 (50%)
Rheumatoid arthritis	5 (2.37%)	0 (0%)
Osteitis condensans ilii	13 (6.16%)	0 (0%)
Gouty arthritis	4 (1.90%)	2 (9.09%)
Scoliosis	1 (0.47%)	0 (0%)
Degenerative arthritis	43 (20.38%)	0 (0%)
Undifferentiated connective tissue disease	3 (1.42%)	0 (0%)
Diabetes	2 (0.95%)	0 (0%)
Juvenile idiopathic arthritis	4 (1.90%)	0 (0%)
Infection	7 (3.32%)	0 (0%)
Bone metastasis	3 (1.42%)	0 (0%)
Spinal tuberculosis	1 (0.47%)	0 (0%)
Prolapse of lumbar intervertebral disc	11 (5.21%)	1 (4.55%)
Osteonecrosis of the femoral head	1 (0.47%)	0 (0%)
Bone marrow fibrosis	1 (0.47%)	0 (0%)
Lumbar muscle strain	0 (0%)	1 (4.55%)
Osteoarthritis	6 (0.28%)	1 (4.55%)

Note: Data in parentheses are percentage. TAH=the Third Affiliated Hospital of Southern Medical University; NHH=Nanhai Hospital; n=Number.

骨骼放射科医生使用ITK-SNAP软件对MRI进行了标注,标注区域覆盖了双侧骶髂关节的整个区域(表2)。然后1名有31年诊断经验的资深医生回顾了有争议的标注案例,并确定最终的ROI。T1WI和FS-MRI图像如图1A所示。

图像预理由数据重采样、图像归一化和ROI提取组成。首先将所有病例样本重采样到统一图像分辨率,然后将图像灰度值统一归一化到[0, 1]范围,最后利用放射科医生勾画的骶髂关节的ROI坐标,裁剪出以ROI为中心,尺寸为 $24 \times 320 \times 224$ 的三维图像,将这个三维图像作为网络的输入。图1B展示了T1WI和FS-MRI图像的ROI。

1.3 深度学习模型建立

本研究提出了一种名为MSFANet的3D多序列融合网络,旨在通过注意力机制建模序列间的语义关联来增强特征表示,从而促进对HLA-B27阴性axSpA与非axSpA的分类性能。MSFANet以3D的T1WI和FS-MRI两种序列图像作为输入,由浅层共享特征学习模块和类感知特征学习模块两部分组成(图2)。对于MSFANet的训练,本文提出了一种混合损失函数,对3个分支进行监督。

1.3.1 浅层共享特征学习模块 该模块以3D T1WI和FS-MRI作为输入,由两个残差块组成,其中每个残差块包含两个卷积核尺寸为 $3 \times 3 \times 3$ 的卷积层、批处理归一

化层和ReLU激活函数。两个残差块的参数由T1WI和FS-MRI共享,可学习并提取两种序列图像之间共有的中级语义信息(图2)。

1.3.2 类感知特征学习模块 该模块(图2)通过3个不同的分支:T1WI分支、FS-MRI分支和Fuse分支进一步学习序列特征的高级语义信息并通过MSFA模块进行多序列特征融合。其中,T1WI分支和FS-MRI分支各包含2个残差块,Fuse分支包含1个残差块和2个MSFA模块。5个残差块结构相同但参数并不共享。这样的设计可以在减少计算开销的同时学习到更多三维图像多序列信息,提高分类的效率和准确度。

1.3.3 MSFA T1WI和FS-MRI是两种密切相关的序列,它们分别从结构和病变两个不同的方面反映了axSpA的特征,因此放射科医生通常需要结合两种序列来完成axSpA诊断。基于此本研究设计了MSFA(图3)对多序列特征之间的相关性进行建模。首先,MSFA通过 $1 \times 1 \times 1$ 的卷积核将两个输入序列特征 $Z_1, Z_2 \in R^{C \times D \times H \times W}$ (C 、 D 、 H 和 W 分别代表特征的通道数、深度、高度和宽度)投射到查询 $Q \in R^{N \times d}$ 、键值 $K \in R^{N \times d}$ 和值 $V \in R^{N \times C}$ 中,其中 N 代表单个序列中的token数目($N = D \times H \times W$), d 为 Q 和 K 的特征尺寸,表示为:

$$Q_1 = W_q Z_1, \quad K_2 = W_k Z_2, \quad V_2 = W_v Z_2 \quad (1)$$

为了更好地融合上下文信息、构建两个序列 Z_1 和

表2 磁共振采集参数

Tab.2 Acquisition parameters of MRI

Hos.	Scanner	Sequence	TR/TE (ms)	Slice thickness (mm)	Spacing between slices (mm)	FOV (mm)	Matrix size
TAH	Philips 1.5T (Achieva)	Coronal T1WI	500/18	4	0.45	1435/181	3965/181
		Axial T2WI	3000/100	4	0.5	1260/10	4180/10
		Coronal T2WI	5228 /100	4	0.45	1335 /1	4005/1
		Coronal T2WI SPAIR	3000/100	4	0.5	1260/10	4180/10
		Axial PDW SPAIR	2586/30	6	0.7	1686/30	3366/30
		Coronal PDW SPAIR	2898/30	4	0.45	1325/30	3485/30
	Philips 3.0T (Ingenia)	Coronal FS-MRI	2400/60	3	0.33	1503/60	3003/60
		Axial T1WI	664/20	5	0.55	1755/20T	4765/20T
		Coronal T1WI	550/22	3	0.35	1405/22I	3805/22I
		Axial T2WI	4565/90	5	0.55	1335/90	4005/90
		Coronal T2WI	2300/85	3	0.35	1395/85	4005/85
		Axial T2WI SPAIR	4230/70	6	0.7	1750/70	4120/70
		Coronal T2WI SPAIR	2929/70	3	0.35	1395/70	4005/70
		T2WI DIXON-w	2300/85	3	0.33	1393/85	4003/85
Philips 1.5T (Achieva)	Coronal T1WI	468/10	3.5	0.35	2205/10I	2645/10I	
	Axial T2WI	3200/100	4	0.4	2450/10	2700/10	
	Coronal T2WI	5500/95	4	0.35	2205/95	2655/95	
	Coronal T2WI SPAIR	3000/60	3.5	0.35	2205/60	2605/60	
	Axial PDW STIR	2500/30	5	0.5	2600/30	2890/30	
	Coronal PDW SPAIR	2900/30	4	0.4	2200/30	2600/30	
	NHH	Coronal STIR	2500/60	3.5	0.35	2205/60	2635/60
		Axial T1WI	450/8	4	0.5	220/8 T	315/8 T
		Coronal T1WI	506/10	3	0.35	1805/10I	2545/10I
		Axial T2WI	5000/100	4	0.5	1800/10	2350/10
Coronal T2WI		2500/90	3	0.35	1805/90	2505/90	
Axial T2WI SPAIR		5100/80	5	0.5	2200/80	3000/80	
Philips 3.0T (Elition)	Conoral T2WI SPAIR	3000/70	3	0.35	1805/70	2015/70	
	T2WI DIXON-w	2500/90	3	0.35	1805/90	2125/90	

T1WI=T1-weighted image; T2WI=T2-weighted image; SPAIR=spectral attenuated inversion recovery image; FS-MRI=short TI inversion recovery; Hos.=Hospital.

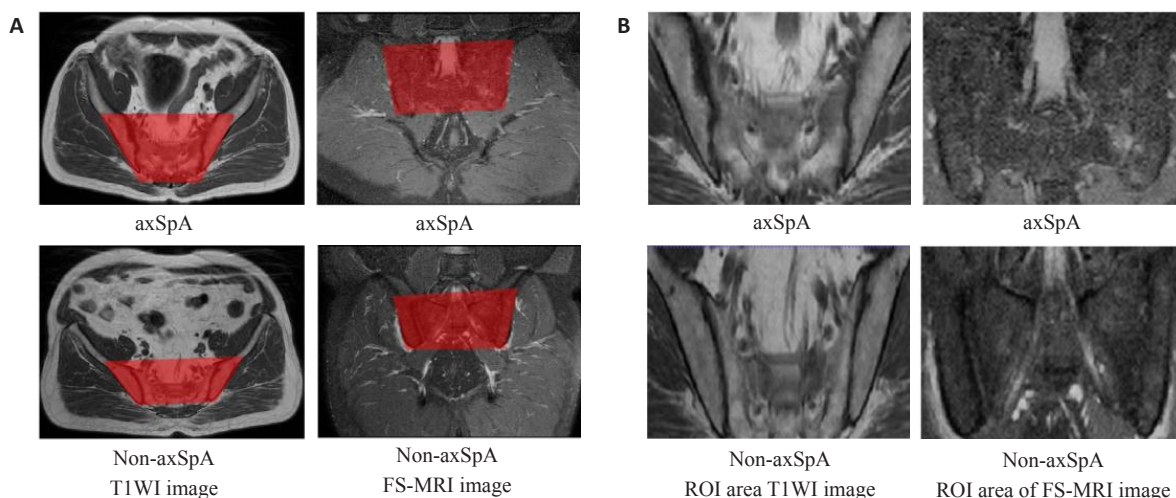


图1 T1WI和FS-MRI图像示例以及 T1WI和FS-MRI图像ROI示例
Fig.1 Representative T1WI and FS-MRI images (A) and the ROIs (B).

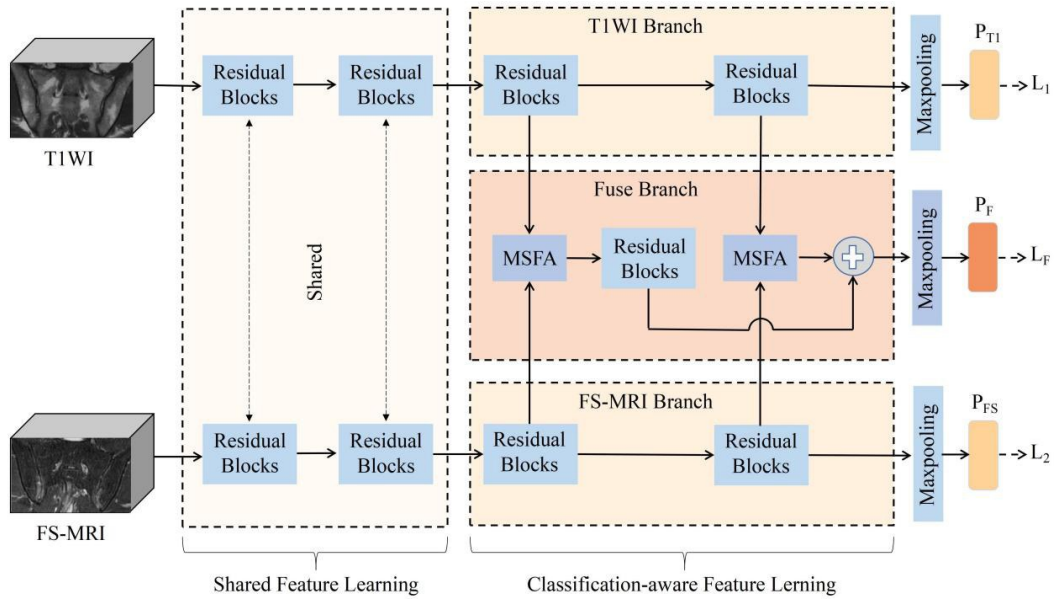


图2 深度学习模型结构图
Fig.2 Overview of the proposed MSFANet.

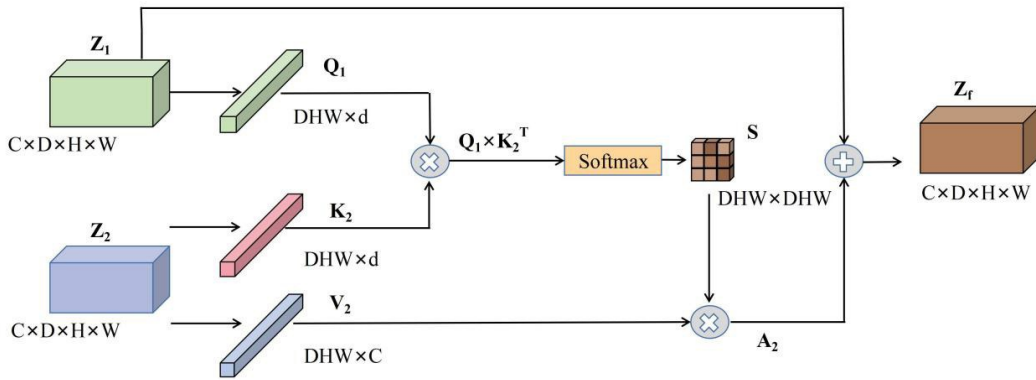


图3 多序列特征融合
Fig.3 Architecture of the proposed MSFA.

Z_2 的相关关系,本研究将 Z_1, Z_2 作为 MSFA 的输入,对 Z_1 的每个子特征进行查询,以获得包含每个 token 和其所对应的 Z_2 的 N 个 token 之间 $1 \times N$ 个相关性的注意力亲和矩阵。具体地,本研究在 Q_1 和 K_2 的转置之间进行矩阵乘法,并应用 softmax 层来计算空间注意图 S ,该图经过归一化后与 Z_2 的关键特征 V_2 中的相应子特征进行点积相乘得到 A_2 。它们的点积将更多的注意力分配给具有高相关得分的区域。然后在 A_2 和 Z_1 之间进行加权求和,最终得到融合后的特征 Z_f :

$$A_2 = \text{soft max} \left(\frac{Q_1 K_2^T}{\sqrt{d_q}} \right) V_2 \quad (2)$$

$$Z_f = \alpha \cdot A_2 + Z_1 \quad (3)$$

值得注意的是,MSFA 模块的输入和输出具有相同的大小,这使得 MSFA 模块的输出可以与其他分支使用

相同大小的卷积块,并允许 MSFA 模块的输出和 Fuse 分支中的残差块构建一个残差连接以加速收敛,同时避免因网络过深导致梯度消失或者爆炸带来的训练困难问题。

1.3.4 损失函数 T1WI、FS-MRI 和 Fuse 三个分支用独立的网络处理不同的序列特征,并用附加的注意力机制整合这些序列特征。Kawahara 等^[26]提出的多序列多任务的损失函数有助于网络学习到更准确的参数,从而提升模型的诊断性能。受此启发,本研究提出了一个深度监督损失的方法,它对每一个支路的预测结果进行监督,旨在获得更鲁棒的诊断结果。提出的深度监督损失定义如下:

$$L = \beta_f L_f + \beta_1 L_1 + \beta_2 L_2 \quad (4)$$

其中, L_f 是根据 Fuse 分支输出计算得到的分类损失, L_1, L_2 分别是 T1WI 分支和 FS-MRI 分支的分类损失

失。 β_f 、 β_1 和 β_2 表示权重系数,通过网络学习而获得。另一方面,公式(4)后两项也可以看作是多序列正则化项,有助于提高网络的泛化能力。 L_f 、 L_1 和 L_2 均采用交叉熵损失函数 L :

$$L(d, d_{gt}) = -d_{gt} \log(d) - (1 - d_{gt}) \log(1 - d) \quad (5)$$

其中 d 为不同分支输出的预测分数, d_{gt} 为真实的病人标签。

1.4 模型训练

本研究将来自TAH的375例数据用于训练和内部验证,并把来自NHH的49例数据用于独立外部验证。本研究使用Adam优化器对网络进行优化,初始学习速率、批大小和权值衰减分别设置为 1×10^{-4} 、6和 1×10^{-4} 。并且采用四折交叉验证方法对所有的模型进行了评估,其中每次训练迭代300个epoch。所有实验在Pytorch环境中采用NVIDIA 1050Ti 8GB GPU硬件实现。

1.5 Grad-CAM可视化

本研究使用了梯度加权类激活图映射(Grad-CAM)进行了可视化分析,突出显示与模型预测相关的图像区域,以提高网络的可解释性^[27,28]。具体地,本研究从每一个序列的单独分支中选择最后一个卷积层,计算全连接层相对于该卷积层的梯度,得到权重向量,将权重向量和卷积层的输出特征图进行加权求和,得到加权后的特征图。再进行类别标签上采样,得到和输入图像相同大小的激活图。对于给定的3D图像,本研究按层输出激活图。其中 $f_k(x, y, z)$ 代表最后一个卷积层中第 k 个通道上第 z 层特征图 (x, y, z) 的激活值,激活值全局平均池化的结果是 $F_k = \sum_{x,y} f_k(x, y, z)$ 。对于类别 c ,输入到softmax里的类别得分表示为 $S_c = \sum_k w_k^c F_k$, w_k^c 为 F_k 对于类别 c 的重要性(权重)。因此我们得到:

$$S_c = \sum_k w_k^c F_k = \sum_{(x,y)} \sum_k w_k^c f_k(x, y, z) \quad (6)$$

定义 M_c 为类别 c 的激活图映射,则

$$M_c(x, y, z) = \sum_k w_k^c f_k(x, y, z) \quad (7)$$

最终, $S_c = \sum_{x,y} M_c(x, y, z)$,其中 $M_c(x, y, z)$ 直接代表了 (x, y, z) 的激活值对图像分类为 c 的重要性。

1.6 评估指标

在本研究中,采用准确性(Accuracy),敏感性(Sensitivity),特异性(Specificity),F1评分(F1-score)和受试者工作特征曲线下面积(AUC)评估所有模型的性能。指标的公式定义如下:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (8)$$

$$Sensitivity = \frac{TP}{TP + FN} \quad (9)$$

$$Specificity = \frac{TN}{TN + FP} \quad (10)$$

$$F1 - score = 2 \times \frac{Sensitivity \times Precision}{Sensitivity + Precision} \quad (11)$$

$$AUC = \sum_{i=1}^{c-1} \frac{(FPR_{i+1} - FPR_i) \times (FPR_{i+1} + FPR_i)}{2} \quad (12)$$

其中,TP, TN, FP, FN分别为真阳性(正确分类的正样本)、真阴性(正确分类的负样本)、假阳性(被错误标记为正样本的负样本)和假阴性(被错误标记为负样本的正样本)的数量。与其它4个指标相比,AUC为受试者工作特性曲线(ROC)与坐标轴围成的面积,其中纵坐标和横坐标分别为TPR和FPR,表示为由不同临界值决定的真阳性率和假阳性率,计算公式如下:

$$TPR = \frac{TP}{TP + FN} \quad (13)$$

$$FPR = \frac{FP}{FP + TN} \quad (14)$$

1.7 统计学方法

采用SPSS22.0软件进行统计学分析。所有的实验都进行了4折交叉验证,以4次实验结果的均数±标准差作为最终的结果。通过配对 t 检验,检测不同模型的差异性。所有统计检验均为双侧检验,当 $P \leq 0.05$ 时被认为差异有统计学意义。

2 结果

2.1 与最先进的方法进行比较

表3比较了MSFANet在内部验证集上与当前最先进的多序列分类方法的性能。从表中可以观察到,MSFANet的AUC、灵敏度、F1分数、召回率排名最高,特异性排名第二高。与其它两种多阶段融合方法CentralNet和Multi-stream相比,MSFANet取得了最高的AUC、准确率、灵敏度和F1指数,分别为0.840, 77.93%, 83.7%和71.24%,比两种多阶段融合方法分别高出0.041和0.040, 2.93%和1.62%, 0.45%和2.22%, 2.67%和2.67%。另外,该表还给出了两个基于双分支结构的基线网络,ShuffleNet和EfficientNet,的分类结果。这些网络只在决策级融合序列特征,无法学习不同分支之间的相互关系。以ShuffleNet为例,它使用同一病人的T1WI和FS-MRI图像作为输入,特征提取器在训练过程中共享权重参数,为了融合多序列信息,将两种序列的特征直接拼接。与之相比,本研究提出的算法MSFANet在序列间的交互使得其在性能上平均提升超过3%。统计学检验结果显示,MSFANet显著优越于其它8种方法($P < 0.05$)。图4A对比展示了不同模型性能的ROC曲线。为了检验模型的可重复性,本研究还引入了独立外部验证集的数据(表4)。本研究发现,MSFANet在绝大部分指标上的预测性能相较于其它模型均有显著提高。具体来说,AUC上升了3%~7%,准确率提升了1%~5%,F1-score增加了2%~10%,且 $P < 0.05$ 。

表 3 在内部验证集中不同模型的对比结果

Tab.3 Classification results of the proposed algorithm and multi-sequence fusion algorithm in the internal validation set

Model	AUC	ACC (%)	SEN (%)	SPEC (%)	F1-score (%)	P
EmbeddingNet ^[29]	0.772±0.015	72.08±0.53	78.69±1.82	67.58±3.40	63.60±2.71	<0.001
CentralNet ^[25]	0.799±0.010	75.00±2.53	83.25±4.04	69.26±2.89	68.57±2.25	0.0053
TripleNet ^[24]	0.788±0.054	76.60±8.77	82.37±4.99	69.23±2.17	65.40±5.20	<0.001
Multi-stream ^[30]	0.800±0.054	76.31±7.22	81.48±4.60	73.68±7.66	68.57±6.46	0.0054
MTLN ^[31]	0.775±0.032	75.27±1.96	79.55±2.97	65.26±3.20	64.60±2.44	<0.001
Bressem et al. ^[21]	0.808±0.049	74.62±6.54	83.58±3.15	68.42±4.05	68.75±5.21	0.0059
ShuffleNet ^[32]	0.790±0.044	73.14±4.87	83.33±3.07	64.71±2.20	66.34±2.36	<0.001
EfficientNet ^[33]	0.767±0.021	74.93±2.48	78.56±0.90	66.78±6.45	60.71±1.05	<0.001
MSFANet (ours)	0.840±0.011	77.93±1.59	83.70±2.38	70.29±0.68	71.24±6.48	

2.2 消融实验

2.2.1 MSFA 的有效性 本研究首先比较了使用MSFA模块进行多序列特征融合的MSFANet和没有使用MSFA模块的基线网络(Baseline)的分类性能。表5中的结果表明,当均采用Hybrid loss作为损失函数时,

MSFANet的分类性能显著优越于Baseline,其中AUC、准确度、和敏感度分别提高0.0721,2.93%,和1.41%($P<0.001$)。当均采用BCE loss时,虽然MSFANet相比Baseline的灵敏度略降低了0.31%,但AUC,准确度和特异度分别提高了0.0691,3.2%和1.99%(图4B)。

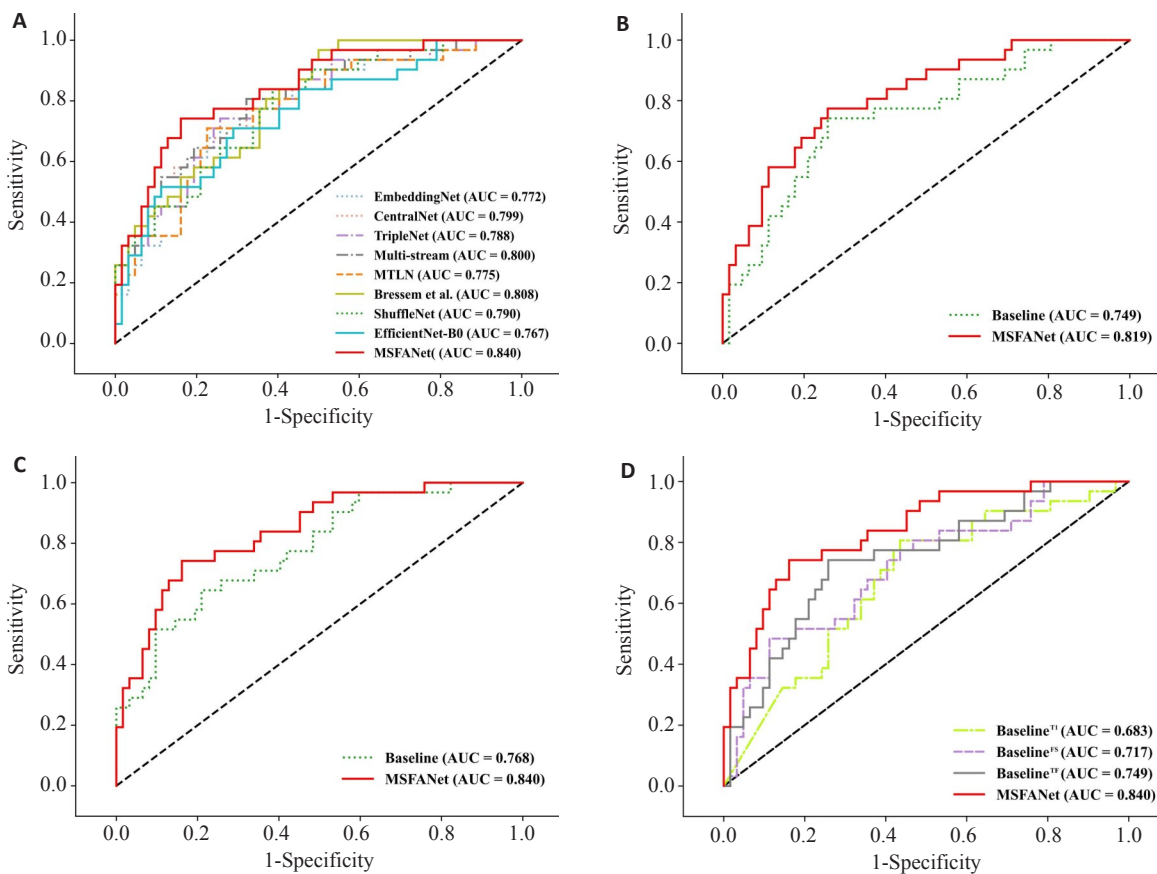


图4 模型间的ROC曲线对比

Fig.4 ROC curves for inter-model comparison. A: Comparison of the results of the proposed algorithm and other multi-sequence fusion algorithm. B: Comparison of the results using BCE loss. C: Comparison of the results using hybrid loss. D: Comparison of the results of different sequence algorithms.

表 4 在独立外部验证集中不同模型的对比结果

Tab.4 Classification results of the proposed algorithm and multi-sequence fusion algorithm in the external validation set

Model	AUC	ACC (%)	SEN (%)	SPEC (%)	F1-score (%)	P
EmbeddingNet ^[29]	0.707±0.056	71.43±2.92	76.25±6.60	67.15±2.92	73.30±2.18	0.032
CentralNet ^[25]	0.741±0.036	72.98±4.25	83.15±2.35	72.18±1.86	72.68±7.00	0.028
TripleNet ^[24]	0.729±0.098	70.40±2.66	75.70±3.22	65.53±3.27	71.75±3.45	<0.001
Multi-stream ^[30]	0.736±0.020	71.40±3.35	79.93±1.04	65.35±0.25	72.03±0.53	<0.001
MTLN ^[31]	0.738±0.031	72.38±2.01	78.83±8.62	69.45±6.43	72.93±4.64	<0.001
Bressem et al. ^[21]	0.717±0.026	69.38±3.74	78.50±1.06	63.80±2.50	69.88±2.35	0.048
ShuffleNet ^[32]	0.727±0.046	69.88±1.96	74.90±2.95	65.08±2.46	71.45±2.34	<0.001
EfficientNet	0.714±0.025	69.38±6.88	82.40±1.87	63.65±9.31	67.18±1.91	0.043
MSFANet (ours)	0.783±0.043	74.47±4.01	82.43±0.92	70.40±8.20	74.65±3.69	-

2.2.2 Hybrid loss 的有效性 本研究比较了采用交叉熵函数和提出的混合损失函数对模型性能的影响。相比只使用单一损失函数,采用混合损失函数的MSFANet在AUC、准确度和灵敏度上分别提高了0.021、1.86%、0.43% (表5);同时,使用混合损失函数的Baseline比采用单一损失函数时的AUC、准确率和特异度分别提高了0.019、2.13%、6.75%(图4C)。

2.3 与单序列分类方法比较

与只采用单序列 T1WI 和 FS-MRI 作为输入的模

型(Baseline^{T1}和Baseline^{FS})相比,采用直接拼接融合模型(Baseline^{TF})的准确率分别提高了5.58%和2.66%,而采用MSFA融合机制模型(MSFANet)的准确率分别提高了10.64%和7.72%(表6,P<0.05,图4D)。

2.4 可视化分析

图5显示了基于T1WI和FS-MRI两个序列图像生成的激活图,其中第1列和第3列是原始图像,第2列和第4列是对应的注意力图。激活图中从红色、黄色到蓝色区域分别表示重要性不断降低。

表 5 使用两种损失函数的不同模型的对比结果

Tab.5 Classification results of all the models using the two loss functions

Network	Loss	AUC	ACC (%)	SEN (%)	SPEC (%)	P
Baseline	BCE loss	0.749±0.096	72.87±4.63	83.58±7.25	62.76±4.24	<0.001
Baseline	Hybrid loss	0.768±0.029	75.00±1.37	82.29±0.78	69.51±3.47	<0.001
MSFANet	BCE loss	0.819±0.018	76.07±4.12	83.27±2.87	71.50±5.10	<0.001
MSFANet	Hybrid loss	0.840±0.011	77.93±1.59	83.70±2.38	70.29±0.68	-

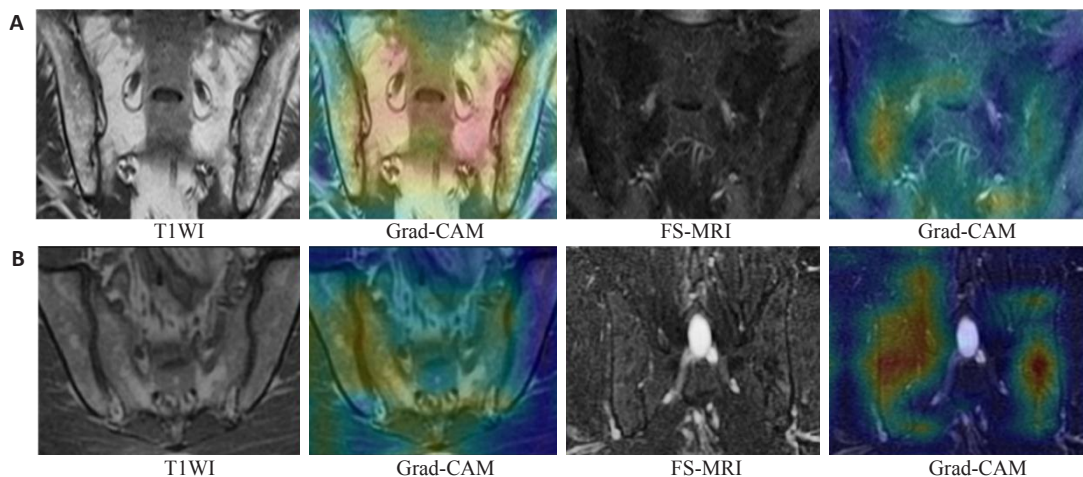


图 5 分类模型的梯度加权类激活图示例

Fig.5 Examples of gradient-weighted class activation mapping (Grad-CAM) for the classification model in a 40-year-old male patient with HLA-B27 negative axSpA (A) and a 35-year-old male with HLA-B27 negative non-axSpA (B).

表6 使用不同序列的不同模型的对比结果

Tab.6 Classification results of different sequence algorithms with different models

Network	Sequence	AUC	ACC (%)	SEN (%)	SPEC (%)	P
Baseline ^{T1}	T1WI	0.683±0.044	67.29±4.11	83.05±5.90	67.25±1.64	<0.001
Baseline ^{FS}	FS-MRI	0.717±0.033	70.21±5.28	81.82±5.82	65.60±2.51	<0.001
Baseline ^{TF}	T1WI+FS-MRI	0.749±0.096	72.87±4.63	83.58±7.25	62.76±4.24	0.0043
MSFANet	T1WI+FS-MRI	0.840±0.011	77.93±1.60	83.70±2.38	70.29±0.68	-

3 讨论

HLA-B27阴性axSpA的早期诊断对提高病人的疾病管理和生活质量具有重要的临床应用价值和意义。MRI是一种广泛用于axSpA早期诊断的成像方式。然而,对于低年资放射科医生来说,正确识别且解释骶髂关节的结构和病理异常,从而准确诊断axSpA尚具有挑战性。本研究模仿临床诊断习惯开发了MSFANet,通过融合T1WI和FS-MRI两种序列图像,来区分HLA-B27阴性群体中的axSpA和非axSpA病人。传统的多序列融合算法平等对待每一个序列数据,没有充分利用对分类有意义的多序列信息,造成算法瓶颈^[34,35]。最近的一些研究在决策阶段进行序列特征融合,但这种方法无法针对中级特征挖掘更多有用的信息^[24,25]。针对以上问题,本研究设计了基于序列间交互机制的多序列网络,通过捕捉T1WI和FS-MRI两个序列之间的全局相关性,以充分挖掘和利用来自不同序列的信息,提高对HLA-B27阴性axSpA的诊断性能。

本研究对MSFANet的有效性进行了充分验证。首先,本研究将MSFANet与其他最先进(SOTA)的融合算法和分类算法进行了比较。MSFANet通过利用序列间的信息交互,取得了比其他SOTA算法更好的分类性能,其中在内部验证集上的AUC、准确度、灵敏度和特异度分别达到0.840,77.93%,83.70%和70.29%,在独立外部验证集分别达到0.783,74.47%,82.43%和70.40%。然后,本研究通过消融实验验证了所提出的MSFANet中每个创新模块对分类的重要性;结果显示,同时使用交互注意力模块和混合损失函数时,模型的所有评价指标都得到了提升。此外,为了验证融合多序列数据对提高诊断性能的有效性,本研究对比了相同框架下,采用单序列T1WI和FS-MRI以及融合双序列图像的模型性能。实验结果显示,双序列协同模型比单序列模型具有更出色的表现,证实了T1WI和FS-MRI在HLA-B27阴性axSpA诊断中的信息互补作用,而且也证明了所提出的MSFA模块对双序列特征进行细化和融合的有效性。

在现有研究的基础上,本文算法具有以下优势:共享特征模块和类感知特征学习的组合可以在保持高性能预测的前提下减小内存和计算需求;本研究提出的交

叉注意力特征融合方法和混合损失函数,能够有效提升模型性能,其中MSFA模块比决策端直接融合方法的精度和鲁棒性有明显提升;通过深度学习可视化技术对模型的感兴趣区进行可视化,提高了模型的临床可解释性;独立外部验证检验了模型的可重复性;实验包括了在不同机器上使用不同参数采集的MRI图像,符合真实的临床场景,测试集结果表明本研究的模型对HLA-B27阴性的axSpA诊断具有良好的临床泛化性。然而我们的研究仍然存在不足。首先,本研究是回顾性分析,不可避免地存在选择偏倚,因此需要前瞻性研究。第二,本研究选择了全局标签训练模型,并没有提供基于骶髂关节的象限分析,这将允许对不同的关节区域进行更准确的评估。

综上所述,本文提出了一种新的多序列融合框架,先通过共享特征学习模块提取浅层通用特征,再用独立的网络分支分别处理每种序列特征,最后使用额外的中央网络融合单序列特征。该框架可以有效地捕获多序列MRI信息,提高对HLA-B27阴性axSpA的诊断准确性。本研究设计了一种具有跨模态注意力机制的多序列融合模块MSFA,从多序列MRI中充分挖掘高质量的特征表示,并探索特征之间的潜在合作关系,从而增强融合特征的表达能力。本研究引入了混合损失函数,通过学习各支路的损失权重系数,平衡各个分支对分类任务的贡献,提高模型对任务的推断准确性。

参考文献:

- [1] Poddubnyy D, Rudwaleit M, Haibel H, et al. Rates and predictors of radiographic sacroiliitis progression over 2 years in patients with axial spondyloarthritis[J]. *Ann Rheum Dis*, 2011, 70(8): 1369-74.
- [2] Bakland G, Alsing R, Singh K, et al. Assessment of SpondyloArthritis International Society criteria for axial spondyloarthritis in chronic back pain patients with a high prevalence of HLA-B27[J]. *Arthritis Care Res*, 2013, 65(3): 448-53.
- [3] Sieper J, Poddubnyy D. Axial spondyloarthritis[J]. *Lancet*, 2017, 390(10089): 73-84.
- [4] Sieper J, Rudwaleit M, Baraliakos X, et al. The Assessment of SpondyloArthritis international Society (ASAS) handbook: a guide to assess spondyloarthritis[J]. *Ann Rheum Dis*, 2009, 68(Suppl 2): ii1-44.
- [5] Miranda VL, der Heijde Désirée V, Alexandre S, et al. Is a positive

- family history of spondyloarthritis relevant for diagnosing axial spondyloarthritis once HLA-B27 status is known [J]? *Rheumatol Oxf Engl*, 2019, 58(9): 1649-54.
- [6] Coates LC, Baraliakos X, Blanco FJ, et al. The Phenotype of Axial Spondyloarthritis: Is It Dependent on HLA-B27 Status[J]? *Arthritis Care Res (Hoboken)*, 2021, 73(6): 856-60.
- [7] Imke R, Johanna C, Falk H, et al. Determinants of diagnostic delay in axial spondyloarthritis: an analysis based on linked claims and patient-reported survey data [J]. *Rheumatol Oxf Engl*, 2019, 58(9): 1634-8.
- [8] Dougados M, Baeten D. Spondyloarthritis [J]. *Lancet*, 2011, 377(9783): 2127-37.
- [9] Marco GC, Delgado-Domínguez CJ, Gálvez-Ruiz D, et al. The effect of axial spondyloarthritis on mental health: results from the atlas[J]. *J Rheumatol*, 2019, 46(10): 1284-9.
- [10] Iván F, Javier R, Fernanda G, et al. Potential relation of cardiovascular risk factors to disease activity in patients with axial spondyloarthritis[J]. *Ther Adv Musculoskelet Dis*, 2021, 13.
- [11] Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data[J]. *Radiology*, 2016, 278(2): 563-77.
- [12] Peikert T, Bartholmai BJ, Maldonado F. Radiomics-based management of indeterminate lung nodules are we there yet [J]? *Am J Respir Crit Care Med*, 2020, 202(2): 165-7.
- [13] Wesdorp N, Hellingman T, Jansma E, et al. Advanced analytics and artificial intelligence in gastrointestinal cancer: a systematic review of radiomics predicting response to treatment [J]. *Eur J Nucl Med Mol Imaging*, 2020, 48: 1785-94.
- [14] Conti A, Duggento A, Indovina I, et al. Radiomics in breast cancer classification and prediction [J]. *Semin Cancer Biol*, 2021, 72: 238-50.
- [15] Tenório APM, Faleiros MC, Junior JRF, et al. A study of MRI-based radiomics biomarkers for sacroiliitis and spondyloarthritis [J]. *Int J Comput Assist Radiol Surg*, 2020, 15(10): 1737-48.
- [16] Ye L, Miao SL, Xiao Q, et al. A predictive clinical-radiomics nomogram for diagnosing of axial spondyloarthritis using MRI and clinical risk factors [J]. *Rheumatology (Oxford)*, 2022, 61(4): 1440-7.
- [17] Song CY, Zhu S, Liu YY, et al. DCNAS-Net: deformation convolution and neural architecture search detection network for bone marrow oedema [J]. *BMC Med Imaging*, 2023, 23(1): 45
- [18] Lee KH, Choi ST, Lee GY, et al. Method for diagnosing the bone marrow edema of sacroiliac joint in patients with axial spondyloarthritis using magnetic resonance image analysis based on deep learning [J]. *Diagnostics*, 2021, 11(7): 1156.
- [19] Shenkman Y, Qutteineh B, Joskowicz L, et al. Automatic detection and diagnosis of sacroiliitis in CT scans as incidental findings [J]. *Med Image Anal*, 2019, 57: 165-75.
- [20] Adrien B, Aouad T, Medina CL, et al. A deep learning model for the diagnosis of sacroiliitis according to Assessment of Spondylo-Arthritis International Society classification criteria with magnetic resonance imaging [J]. *Diagn Interv Imaging*, 2023, 104(7/8): 373-83.
- [21] Bressemer KK, Adams LC, Proft F, et al. Deep learning detects changes indicative of axial spondyloarthritis at MRI of sacroiliac joints [J]. *Radiology*, 2022, 305(3): 655-65.
- [22] Zhang YD, Dong ZC, Wang SH, et al. Advances in multimodal data fusion in neuroimaging: overview, challenges, and novel orientation [J]. *Inf Fusion*, 2020, 64: 149-87.
- [23] Wang S, Celebi ME, Zhang YD, et al. Advances in data preprocessing for biomedical data fusion: an overview of the methods, challenges, and prospects [J]. *Inf Fusion*, 2021, 76: 376-421.
- [24] Ge ZY, Demyanov S, Chakravorty R, et al. Skin Disease Recognition Using Deep Saliency Features and Multimodal Learning of Dermoscopy and Clinical Images [C]//International Conference on Medical Image Computing and Computer-Assisted Intervention. Cham: Springer, 2017: 250-258.
- [25] Vielzeuf V, Lechervy A, Pateux S, et al. CentralNet: A Multilayer Approach for Multimodal Fusion [C]//European Conference on Computer Vision. Cham: Springer, 2019: 575-89.
- [26] Kawahara J, Daneshvar S, Argenziano G, et al. Seven-point checklist and skin lesion classification using multitask multimodal neural nets [J]. *IEEE J Biomed Health Inform*, 2019, 23(2): 538-46.
- [27] Desai S, Ramaswamy HG. Ablation-CAM: visual explanations for deep convolutional network *via* gradient-free localization [C]//2020 IEEE Winter Conference on Applications of Computer Vision (WACV). March 1-5, 2020, Snowmass, CO, USA. IEEE, 2020: 972-80.
- [28] Chattopadhyay A, Sarkar A, Howlader P, et al. Grad-CAM: generalized gradient-based visual explanations for deep convolutional networks [C]//2018 IEEE Winter Conference on Applications of Computer Vision (WACV). March 12-15, 2018, Lake Tahoe, NV, USA. IEEE, 2018: 839-47.
- [29] Yap J, Yolland W, Tschandl P. Multimodal skin lesion classification using deep learning [J]. *Exp Dermatol*, 2018, 27(11): 1261-7.
- [30] Fernandez-Quilez A, Eftestøl T, Kjosavik SR, et al. Multi-planar T2W MRI for an improved prostate cancer lesion classification [C]//2022 IEEE 19th International Symposium on Biomedical Imaging (ISBI). March 28-31, 2022, Kolkata, India. IEEE, 2022: 1-5.
- [31] Zhang YT, Li HM, Du J, et al. 3D multi-attention guided multi-task learning network for automatic gastric tumor segmentation and lymph node classification [J]. *IEEE Trans Med Imaging*, 2021, 40(6): 1618-31.
- [32] Zhang X, Zhou X, Lin M, et al. ShuffleNet: An Extremely Efficient Convolutional Neural Network for Mobile Devices [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2017: 6848-56.
- [33] Zhang XY, Zhou XY, Lin MX, et al. ShuffleNet: an extremely efficient convolutional neural network for mobile devices [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. June 18-23, 2018, Salt Lake City, UT, USA. IEEE, 2018: 6848-56.
- [34] Wang Y, Yang YW, Guo X, et al. A novel multimodal MRI analysis for alzheimer's disease based on convolutional neural network [C]//2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). July 18-21, 2018, Honolulu, HI, USA. IEEE, 2018: 754-7.
- [35] Zhuang HX, Liu RH, Wu CW, et al. Multimodal classification of drug-naïve first-episode schizophrenia combining anatomical, diffusion and resting state functional resonance imaging [J]. *Neurosci Lett*, 2019, 705: 87-93.