

中图法分类号: TP391 文献标识码: A 文章编号: 1006-8961(2025)04-1183-12

论文引用格式: Liu Y L, Cha Y Q, Huang Y S and Gao M. 2025. Classification algorithm for anterior cruciate ligament injury embedded with slice sequence association mode. Journal of Image and Graphics, 30(4):1183-1194(刘英莉, 茶寅秋, 黄一山, 高明. 2025. 嵌入切片序列关联模式的前交叉韧带损伤分类. 中国图象图形学报, 30(4):1183-1194)[DOI:10.11834/jig.240302]

嵌入切片序列关联模式的前交叉韧带损伤分类

刘英莉^{1,2}, 茶寅秋^{1,2}, 黄一山^{1,2}, 高明^{3*}

1. 昆明理工大学信息工程与自动化学院, 昆明 650500; 2. 昆明理工大学云南省计算机技术应用重点实验室, 昆明 650500;
3. 云南省第一人民医院骨科, 昆明 650032

摘要: 目的 膝关节前交叉韧带(anterior cruciate ligament, ACL)损伤的及时诊断是降低骨关节炎等风险重要途径。作为 ACL 损伤判别的常用影像学方法, 三维磁共振成像(magnetic resonance imaging, MRI)相较于二维自然图像细节更繁杂, 医生诊断所需时间长, 智能辅助医疗是有益的。目前应用三维卷积神经网络(3D convolutional neural network, 3D CNN)的 ACL 损伤分类算法存在计算成本大且数据使用不充分的问题, 而基于二维卷积神经网络(2D convolutional neural network, 2D CNN)的方法则忽略了第三维度的相关性和 ACL 的形态多样性。**方法** 提出一种嵌入切片序列关联模式的网络(slice sequence association mode network, SSAMNet)。SSAMNet 通过设计切片序列信息融合(slice sequence information fusion, SFS)模块从 MRI 数据的邻近、全量切片中并行学习序列性质, 建立切片特征映射中的关联模式, 整合切片信息。借助多级尺度特征自适应关注(multi-level scale feature adaptive attention, MSFAA)模块, 对不同关联尺度权重再分配, 以适应 ACL 区域形状和位置表征多变的特性。**结果** 在 MRNet 数据集的实验中, 模型的 AUC(area under curve)值达到 98.4%, 相较于其他的 ACL 损伤分类算法性能最优, 准确性及特异性指标也分别达到了 91.4% 和 97% 的最优值。在 kneeMRI 数据集上进行 5 折交叉验证实验后, SSAMNet 的 AUC 平均值最高, ROC(receiver operating characteristic curve)曲线始终保持着平稳的趋势。可视化结果也表明提出的方法可以有效关注病变区域。**结论** 针对 ACL 损伤分类任务, 所提模型性能优异且具有鲁棒性, 存在临床应用价值。代码开源地址: <https://github.com/wabk/SSAMNet>。

关键词: 磁共振成像(MRI); 3D 图像分类; 切片特征聚合; 自适应尺度注意力; 二维卷积神经网络(2D CNN)

Classification algorithm for anterior cruciate ligament injury embedded with slice sequence association mode

Liu Yingli^{1,2}, Cha Yinqiu^{1,2}, Huang Yishan^{1,2}, Gao Ming^{3*}

1. Faculty of Information Engineering and Automation, Kunming University of Science and Technology, Kunming 650500, China;
2. Yunnan Key Laboratory of Computer Technologies Application, Kunming University of Science and Technology, Kunming 650500, China;
3. Orthopedics Department, the First People's Hospital of Yunnan Province, Kunming 650032, China

Abstract: Objective The prompt diagnosis of anterior cruciate ligament (ACL) injuries in the knee has been shown to reduce the risk of osteoarthritis, knee injuries, and more. As a common imaging method for identifying ACL injuries, magnetic resonance imaging (MRI) has been characterized by low acquisition time cost and non-invasiveness. MRI can be

收稿日期: 2024-06-06; 修回日期: 2024-10-08; 预印本日期: 2024-10-15

* 通信作者: 高明 drgaoming@pku.org.cn

基金项目: 云南省科技厅基础研究项目昆医联合专项(202001AY070001-306)

Supported by: Joint Research Program of Yunnan Science and Technology Agency and Kunming Medical University (202001AY070001-306)

regarded as a three-dimensional data. Compared with two-dimensional images, it has more details. Doctors need to combine multiple images within a dataset to make a comprehensive judgment and draw a conclusion, The diagnosis takes a long time, so intelligent assisted medical treatment is necessary. Currently, deep learning-based methods are mainly used for ACL injury classification, which can be divided into algorithms based on 3D convolutional neural networks (3D CNNs) and 2D convolutional neural networks (2D CNNs). ACL injury classification algorithms using 3D CNNs often suffer from high computational costs and insufficient data usage, while algorithms based on 2D CNNs ignore the third-dimensional correlation (correlation of slice dimensions) and the morphological diversity of ACL. To solve the problems, an ACL injury classification algorithm embedded with slice sequence association mode is proposed (SSAMNet). **Method** SSAMNet uses the classic AlexNet model based on 2D CNNs as the backbone network to acquire discriminative features for ACL injury classification. By designing the slice sequence information fusion (SFS) module, sequence properties are learned in parallel from adjacent and full slices of MRI data. First, the initial slice features extracted through the backbone network are divided into slice groups according to the channel dimension, and a shift operation is performed between the groups to achieve the purpose of merging adjacent slice features. Then, key information is extracted with the help of the initial slice features, the slice sequence relationship is modeled by SSAMNet, global slice information is merged and processed, and finally, the association pattern in the slice feature mapping is established. Secondly, the multi-level scale feature adaptive attention (MSFAA) module is used to splice the multiple levels features processed by the backbone network into multi-scale feature groups. The horizontal and vertical direction features are processed through mean expansion to obtain the directional scale weight coefficients, the goal of redistributing weights at different correlation scales is accomplished to accommodate the variable nature of ACL region shape and location representations. Class imbalance is a common problem among the knee MRI datasets selected for this task, where the number of positive samples representing ACL tears exceeds the number of negative samples representing intact ACLs. This problem is not only reflected in the selected datasets but also exists in the real world, where the number of individuals with torn ACLs surpasses the number of those with intact ACLs. Therefore, we choose to minimize the weighted binary cross-entropy loss to reduce the impact of class imbalance on algorithm performance. During the training process, the learning rate is initialized to $1E-5$, and it is updated to 0.95 times the original value every 10 epochs of training. All experiments are performed on PyTorch using an NVIDIA RTX 3090 GPU. To ensure the fairness of comparative experiments, we report the mean and standard of three experimental results.

Result We use accuracy, sensitivity, specificity, and area under the curve (AUC) as evaluation metrics for this algorithm, calculated using data such as true positives, true negatives, false positives, and false negatives. The experiments are conducted on the MRNet dataset and the knee MRI dataset, and their data distributions are different. The MRNet dataset is currently the largest public dataset on the knee joint, including 1 370 sets of MRI data, and the knee MRI dataset contains 917 sets of data. Because the knee MRI dataset is not divided into training sets and test sets, we use stratified random sampling to separate knee MRI, and we use the fivefold cross-validation method to complete the external validation experiment. The final experimental results on the MRNet dataset show that the AUC value of SSAMNet reaches 98.4%, which has the best performance compared with other comparison networks based on 2D CNNs and 3D CNNs, with a specificity value of 97% and an accuracy value of 91.4%. Moreover, the ablation experimental results on this dataset also prove that SFS and MSFAA have a gain value for the ACL injury discrimination task, especially the significant performance improvement brought by adding the SFS module (AUC increased by more than 2.5%), which also illustrates the effectiveness of the embedded slice sequence correlation mode in improving model performance. The joint use of SFS and MSFAA also increases AUC by more than 4%. Then, a fivefold cross-validation experiment is conducted on the knee MRI dataset. The AUC value of the proposed model reaches 88.7%, and the accuracy value exceeds other models by 3%. The ROC curve tends to be stable on the fivefold datasets of different divisions, proving that the model is stable and generalizable. Finally, the visualization results also show that SSAMNet can effectively focus on the ACL area and the indicators that affect the identification of ACL injuries, proving that the model's conclusion of identifying tear injuries is consistent with the actual basis, further proving that the model can effectively identify ACL injuries. **Conclusion** The proposed SSAMNet has excellent performance in ACL injury discrimination, and its robustness is confirmed in external validation experiments with different data distributions, further revealing the potential clinical application value of the proposed model. However,

the current dataset also has the issue of single-label data. Considering the comprehensive data required for clinical usage, we will utilize patients' past medical records and other information as prior knowledge to construct an auxiliary diagnostic framework for knee joint diseases in the subsequent steps. The code of this paper will be open source soon: <https://github.com/wabk/SSAMNet>.

Key words: magnetic resonance imaging(MRI); 3D image classification; slice feature aggregation; adaptive scale attention; 2D convolutional neural network(2D CNN)

0 引言

作为运动中最常见的损伤之一的前交叉韧带(anterior cruciate ligament, ACL)损伤,可能导致膝关节不稳定、骨关节炎等继发性问题(Yu 和 Garrett, 2007)。因此,ACL损伤的早期筛查至关重要。磁共振成像(magnetic resonance imaging, MRI)具备无创性、可用性及成像良好等优势,它结合深度学习的方法广泛应用于医疗智能化领域,例如判定乳腺肿瘤良恶性、识别阿尔茨海默症发病阶段等(赵小明等, 2024; 黄帅辉和王金凤, 2022)。MRI也是判别ACL损伤分类的主要影像学方法,对患者来说, MRI相比关节镜检查具备非侵入性和采集时间成本低的优势,已被多次证明能保障诊断的准确性(Mackenzie等, 1996; Oei等, 2003; Crawford等, 2007)。借助MRI还可评估膝关节是否伴随半月板、关节软骨及其他韧带的损伤病变(Shea和Carey, 2015)。考虑到膝关节MRI成像中的图像繁杂及细节丰富使得非肌肉骨骼(musculoskeletal, MSK)放射科医生及偏远地区的临床医生准确诊断难问题(Bien等, 2018; Chang等, 2019),所以,采用人工智能辅助医疗是有益的。

与一般二维自然图像的分类不同, ACL损伤分类是利用三维MRI数据进行图像分类,对损伤的精准判别更复杂困难。目前一些公认的卷积神经网络框架在ACL损伤判别任务中得到了广泛且成功的应用,例如二维卷积神经网络(2D convolutional neural network, 2D CNN)、三维卷积神经网络(3D convolutional neural network, 3D CNN)等。Bien等人(2018)通过2D CNN中的骨干网络 AlexNet(Krizhevsky等, 2017)联合多视图信息提取深层特征表示,并将其应用于ACL损伤分类,其性能与3名MSK放射科医生的评估结果相当。Irmakci等人(2019)探索了不同CNN架构对ACL损伤分类实验结果的

影响,研究表明基于多视图的深度学习方法具有应用价值。Tsai等人(2020)改进了2D CNN,借助多切片归一化和模糊池化形成可有效提取切片关系信息的高效分层网络。Belton等人(2021)分析比较单视图和多视图CNN结合空间注意力块的模型在膝关节损伤检测任务中的性能,证明了联合多视图的网络结构的优越性。Dunnhofer等人(2022)搭建了融合多尺度图像细节单元的特征金字塔网络,并且保证了该单元即插即用,为解决ACL判别过程中需提取图像中的ACL区域且需减少手动标注工作量的问题提供了可行的方案。

Zhang等人(2020)对基于2D CNN的VGG16(Visual Geometry Group)(Simonyan和Zisserman, 2015)及ResNet(residual network)(He等, 2016)和基于3D CNN的DenseNet(dense convolutional network)(Huang等, 2017)的架构进行全面比较,验证了3D CNN在ACL损伤诊断问题中的可行性。Li等人(2023)建立三维加权多视图卷积神经网络对不同MRI视图进行融合,可有效检测ACL损伤。Jeon等人(2021)利用注意力模块和高斯位置编码加强对局部特征的关注,借助基于3D CNN的压缩模块降低了网络设计的复杂性,保障了模型的可解释性,成为目前最先进的ACL损伤分类模型。

尽管这些基础CNN结构及其变体已经能够获得较好的分类效果,但其沿切片维度挖掘序列关系及灵活处理不同病例形态表征的能力仍然不足,可能触及ACL损伤分类任务中的性能瓶颈。具体问题可概括如下:1)2D CNN主要关注二维空间内容信息,对第三维度(切片维度)关注不足;2)3D CNN计算成本大且存在固定输入维度后切片数据使用不充分的问题;3)ACL区域在不同病例中形态各异, CNN在不同空间位置具有共享权重,缺乏处理灵活形状结构的感知能力。

本文提出一种新颖的基于2D CNN的嵌入切片序列关联模式的网络(slice sequence association

mode network, SSAMNet), 高性能地完成 ACL 损伤分类任务。SSAMNet 针对上述 3 个不足提供了点对点的解决方案。本文的主要贡献可概括如下: 1) 通过切片序列信息融合 (slice sequence information fusion, SFS) 模块, 在 2D CNN 基础上提高切片序列维度的关注度, 充分利用切片依赖关系数据; 2) 通过多级尺度特征自适应关注 (multi-level scale feature adaptive attention, MSFAA) 模块, 针对不同病例数据自适应提升关键尺寸关注度, 达到灵活处理不同 ACL 形态结构的目标。

1 方法

1.1 SSAMNet 模型结构

本文主要针对膝关节 MRI 切片序列特性、ACL 区域成像表征研发 SSAMNet 模型, 以适用前交叉韧带损伤的判别分类。该模型设计了两个关键模块:

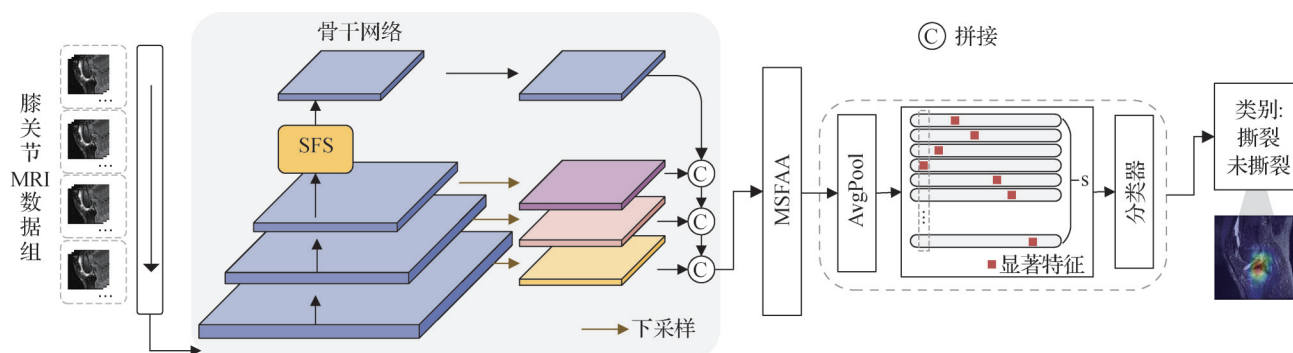


图1 SSAMNet 结构图

Fig. 1 Structure diagram of SSAMNet

1.2 切片序列信息融合模块(SFS)

MRI 数据是通过细微的间隔从组织采样而产生的数十个近似连续的图像切片, 可以视为大小为 $C \times H \times W$ 的 s 张连续切片序列。不同切片承载着与 ACL 损伤判别相关的不同成像特征。仅通过单一切片内的有限信息辅助医生做出论断是不严谨的。结合数张膝关节切片成像信息是进行损伤精细判别的关键因素, 这也与专业医生阅读多张切片判断损伤是否存在的流程相符。2D CNN 对获取单一切片图像内的特征具有优势, 但缺乏切片之间的信息交互过程。如果模型经过 2D CNN 提取的各切片特征能实现邻近切片信息互换, 就可以保证一张切片内具备相邻切片的特征, 这满足联合相邻切片进

切片序列信息融合(SFS)模块和多级尺度特征自适应关注(MSFAA)模块, 融入基础特征提取 CNN 框架, 增强膝关节图像切片序列间的信息交互能力, 提升损伤判别过程中捕获 ACL 尺度细节特征的灵活性。

SSAMNet 整体架构如图 1 所示, 输入膝关节 MRI 切片序列, 输出病例 ACL 损伤判别结果。首先借助特征提取器处理输入的 MRI 切片序列, 获取初始特征。在特征提取器的选择上, 将删除了最后一个分类层的 AlexNet (Krizhevsky 等, 2017) 作为骨干网络, 并加载其在 ImageNet 数据集上的预训练权重。通过 SFS 模块, 得到一系列结合邻近切片信息和全局切片序列关系信息的特征。将该特征与骨干网络生成的不同分辨率特征图拼接后输入 MSFAA 模块, 聚合重要尺度特征后进行分类。由于 MRI 连续切片序列表示一个预测结果, 所以仅需要切片序列的显著特征引导分类器进行最终判断。

行诊断的现实需求。医生通过专业知识可挑选出影响判别 ACL 损伤的关键切片指征, 这要求模型能识别切片组中各切片的重要性和捕获切片序列的相关关系。为此, 设计切片序列信息融合模块(SFS), 其结构如图 2 所示。

给定 MRI 数据 $S = [s_1, s_2, \dots, s_s] \in \mathbf{R}^{C \times H \times W}$, 由骨干网络对 S 处理获得的特征输出为 $S^{(y)} = [s_1^{(y)}, s_2^{(y)}, \dots, s_m^{(y)}, \dots, s_s^{(y)}] \in \mathbf{R}^{C_y \times H_y \times W_y}$, 其中 H_y 和 W_y 是骨干网络第 y 层输出特征图的高度和宽度, C_y 是第 y 层输出特征通道数, y 为骨干网络的第 8 层。

首先将 $S^{(y)}$ 的各切片沿通道维度划分为 d 组不重叠的特征图, 对于第 m 张切片特征 $s_m^{(y)}$ 即存在着 $s_m^{(y)} = [s_{m1}^{(y)}, \dots, s_{m(j-1)}^{(y)}, s_{mj}^{(y)}, s_{m(j+1)}^{(y)}, \dots, s_{md}^{(y)}] \in \mathbf{R}^{(C_y/d) \times H_y \times W_y}$,

函数 $LCC(\cdot)$ 表示关于 $S^{(y)}$ 的邻近切片信息融合, 即

$$S^{(y)'} = LCC(S^{(y)}) \quad (1)$$

具体来说, 对第 m 张切片 $s_m^{(y)}$ 应用邻近切片信息融合后的输出特征为

$$s_m^{(y)'} = LCC(s_m^{(y)}) = [s_{m1}^{(y)}, \dots, s_{(m+1)(j-1)}^{(y)}, s_{(m-1)j}^{(y)}, s_{m(j+1)}^{(y)}, \dots, s_{md}^{(y)}] \quad (2)$$

式中, $s_{(m+1)(j-1)}^{(y)}$ 和 $s_{(m-1)j}^{(y)}$ 分别代表第 $m+1$ 张切片 $j-1$ 组通道、第 $m-1$ 张切片 j 组通道特征。如图2中局部切片信息融合模块所示, 邻近切片信息融合后第 m 张切片将同时包含第 $m-1$ 张、第 $m+1$ 张切片内

的部分信息。最后关于全量切片 $S^{(y)}$ 得到结合邻近切片信息的特征 $S^{(y)'}$ 。

值得注意的是, 即使函数 $LCC(\cdot)$ 代表的邻近切片信息融合可以加强模型学习过程中邻近切片间的特征交互, 但是聚合的切片信息仍然有限, 对切片序列关系的捕捉是局部不充分的。为此, 在SFS中针对全量切片设计了全局关键切片序列关系聚合模块来获取显著切片特征, 学习跨切片的相关性, 如图2中全局关键切片序列关系聚合模块所示。

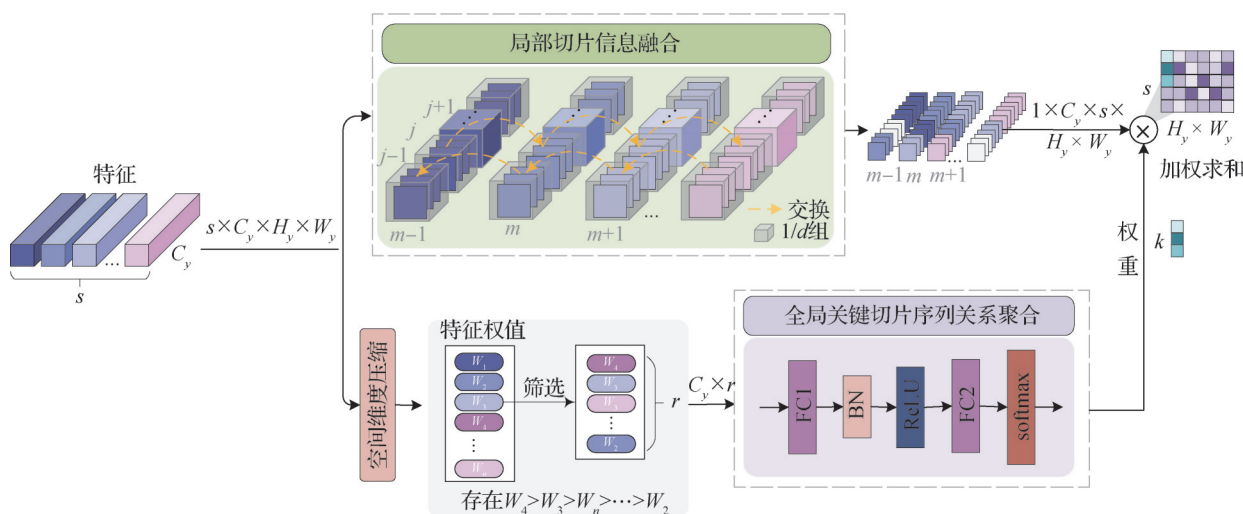


图2 SFS结构图

Fig. 2 Structure diagram of SFS

为了高效建模切片序列关系, 首先压缩输入特征 $S^{(y)}$ 的空间维度, 保持该过程的通道数不变, 借助全连接层建模全局关键切片序列关系, 实现自适应地学习每组患者MRI数据生成序列关系卷积核。最后以类似的通道卷积方式(Liu等, 2021)聚合切片序列关系。详细地说, 对于第 C_s 个通道, 模型的学习过程为

$$\delta_{C_s} = \phi(S_{C_s}^{(y)}) = f_{\text{softmax}} \left(FC_2 \left(ReLU \left(BN \left(FC_1 \left(S_{C_s}^{(y)} \right) \right) \right) \right) \right) \quad (3)$$

式中, 函数 $\phi(\cdot)$ 表示关于 $S^{(y)}$ 的全局关键切片序列关系聚合, $S_{C_s}^{(y)}$ 表示经由骨干网络第 y 层处理后第 C_s 个通道的特征, $\delta_{C_s} \in \mathbf{R}^k$ 代表网络自适应学习的通道权重, k 为卷积核大小。FC₁ (fully connected layer)、FC₂、BN (batch normalization) 和 ReLU (rectified linear unit) 分别代表两个全连接层、批归一化及修正线性

单元。

特别地, SFS在构建全局切片关系前需先在 $S^{(y)}$ 中筛选出对结果判别影响力较大的 r 张切片, 即

$$S^{(y)'} = SRT_r(S^{(y)}) \quad (4)$$

式中, $S^{(y)'} \in \mathbf{R}^{r \times C_y \times H_y \times W_y}$, 函数 $SRT(\cdot)$ 代表在仅关注切片维度的基础上, 根据全量切片特征权值由高到低排序后挑选前 r 张切片特征。

这样做有两个原因。一方面, 全量切片中包含着部分冗余信息, 如图3所示, 膝关节MRI中切片序列前端及末尾部分几乎不包含ACL区域图像, 对指导判别是否存在前交叉韧带损伤作用甚微。相较于 $LCC(\cdot)$ 中仅融合部分邻近切片特征来说, 虽然对全量切片进行特征聚合可以囊括更大范围的切片信息, 但容易引入影响判别的多余切片, 导致不必要的计算量。另一方面, 切片特征的预先筛选机制能提高模型对影响判别结果的关键切片

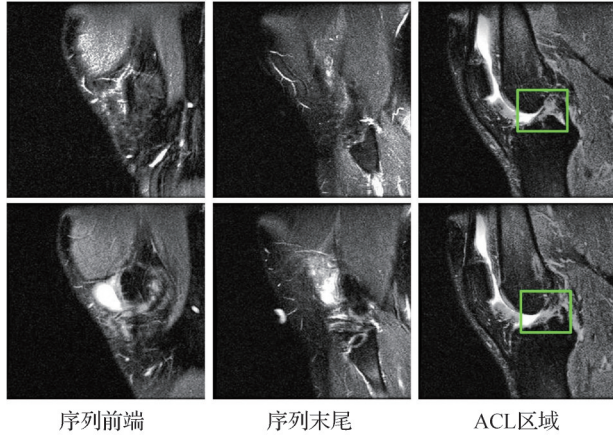


图3 MRI中序列前端、末尾图像与ACL区域图像对比
Fig. 3 Comparison of anterior end, posterior end, and the ACL region of slice sequence in MRI

的关注度。这在某种程度上可以解释为什么提出在全局切片信息聚合前仅挑选权值分数靠前的 r 张切片进行信息融合。此方式在去除冗余信息影响的同时,可捕获特征映射中较大范围切片之间的相关性,弥补先前融合小范围切片信息的不足。

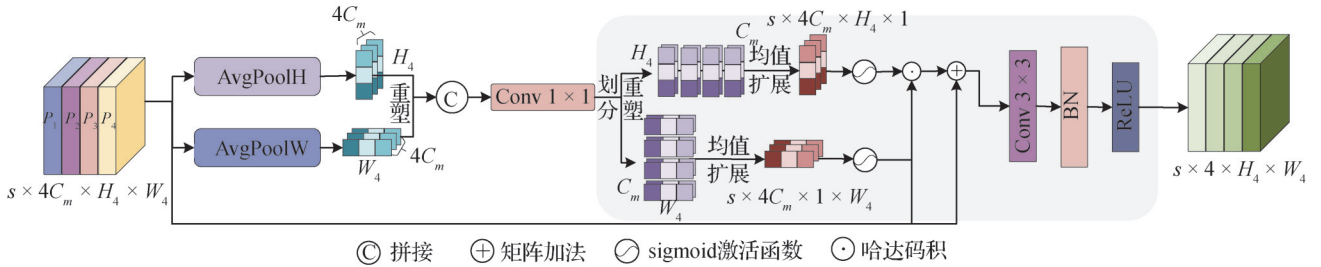


图4 MSFAA 结构图

Fig. 4 Structure diagram of MSFAA

首先将骨干网络中第1个、第2个、第3个最大池化层及第3个ReLU层对MRI数据集 $S \in \mathbb{R}^{s \times C \times H \times W}$ 处理后的输出作为多尺度特征,简写为 B_1, B_2, B_3, B_4 。通过4个 1×1 卷积层及双线性插值法采样降低模型计算复杂度,生成 P_1, P_2, P_3, P_4 。其中, $\{P_1, P_2, P_3, P_4\} \in \mathbb{R}^{s \times C_m \times H_4 \times W_4}$, 处理后特征图通道数和分辨率被压缩至 $C_m, H_4 \times W_4$ 。拼接4个尺度的特征 $P = f_{\text{Concat}}(P_1, P_2, P_3, P_4)$ 作为后续输入向量。其次,与仅建模通道关系来权衡重要性的SE(squeeze and excitation)(Hu等,2018)相比,借助坐标注意力机制(Hou等,2021)可以同时捕捉位置信息和通道之间的关系,丰富ACL区域特征。处理后得到水平

综上所述,提出的SFS可以写为

$$SFS = \phi(SRT_r(S^{(y)})) \otimes LCC(S^{(y)}) = \delta \otimes LCC(S^{(y)}) \quad (5)$$

式中, $\delta = [\delta_{C_1}, \delta_{C_2}, \dots, \delta_{C_s}, \dots, \delta_{C_y}]$ 表示 C_y 个通道的全局关键切片序列关系聚合权值, $\otimes$ 表示加权求和。

1.3 多级尺度特征自适应关注模块(MSFAA)

前交叉韧带的形态、位置在不同患者影像上是变化的,骨干网络在不同空间位置共享相同权重,使得模型不能很好地捕捉ACL可变形形状特性,造成损伤判别效果不佳。考虑到多尺度特征组合既能保留模型浅层架构产生的高分辨率空间细节特征,又能捕获图像中的高级语义信息,满足了膝关节前交叉韧带不同形状大小的识别需求。如果模型借助多尺度信息作为输入特征达到动态调整对不同尺度关注度的目标,就能实现对不同患者影像自适应分配与损伤指征相匹配的尺度权重,灵活挖掘面向不同病例需要着重关注的尺度信息。因此,提出的MSFAA模块可动态学习不同病例对应的尺度权重,匹配不同形态表征,模块结构如图4所示。

方向特征 $F_{\text{CaH}} \in \mathbb{R}^{s \times 4C_m \times 1 \times W_4}$ 及垂直方向特征 $F_{\text{CaV}} \in \mathbb{R}^{s \times 4C_m \times H_4 \times 1}$ 。为从 P_1, P_2, P_3, P_4 中动态分配与各病例强关联的尺度权重,将水平、垂直方向特征重塑为 $F_{\text{CaH}} \in \mathbb{R}^{s \times 4 \times C_m \times 1 \times W_4}$, $F_{\text{CaV}} \in \mathbb{R}^{s \times 4 \times C_m \times H_4 \times 1}$, 对通道维度求均值后扩展维度得出水平、垂直方向系数 θ_H 和 θ_V , 即

$$\theta_H = f_{\text{sigmoid}}(\text{Mean}(F_{\text{CaH}})) \quad (6)$$

$$\theta_V = f_{\text{sigmoid}}(\text{Mean}(F_{\text{CaV}})) \quad (7)$$

式中, $\theta_H \in [0, 1]^{s \times 4C_m \times 1 \times W_4}$, $\theta_V \in [0, 1]^{s \times 4C_m \times H_4 \times 1}$, Mean代表通道维度取均值。

最后将系数 θ_H 和 θ_V 用做尺度注意权重。同时利用 3×3 卷积层提升病变区域细节特征挖掘的能

力。提出的MSFAA模块的输出计算为

$$MSFAA = ReLU\left(BN\left(Conv_{3 \times 3}\left((P \odot \theta_H \odot \theta_V) + P\right)\right)\right) \quad (8)$$

式中, $Conv_{3 \times 3}$ 表示卷积核大小为 3×3 的卷积层, $\odot$ 表示哈达玛积。

2 实验

2.1 数据集描述

2.1.1 MRNet数据集

MRNet(convolutional neural network for classifying MRI series)数据集(Bien等, 2018)是目前最大的公共可用的膝关节MRI数据集。其中包括在美国斯坦福大学医学中心进行的1370次膝关节MRI检查, 随机划分为1130组训练集、120组验证集和120组测试集。最终根据ACL撕裂、ACL未撕裂以及与前两种不同的一般异常, 为每组MRI数据分配一个标签。由于此数据集中的测试集是非公开的, 所以采用与Jeon等人(2021)相似的方式, 先在训练集上完成模型训练, 再使用验证集进行性能评估。

2.1.2 kneeMRI数据集

kneeMRI数据集是由Štajduhar等人(2017)在克罗地亚临床医院中心获得的917组膝关节MRI数据。依据ACL损伤的级别将数据集分为690组无损伤、172组部分损伤和55组完全破裂。采用分层随机抽样以8:2的比例分成训练集与测试集, 最终使用5折交叉验证进行实验。

与Bien等人(2018)的实验类似, 对于上述两个数据集, 考虑区分ACL未撕裂和ACL撕裂的分类任务。

2.2 实验设置

2.2.1 损失函数

由于ACL撕裂数据只占整体数据集的一部分, 即阳性样本的数量明显少于阴性样本的数量, 采用最小化加权二元交叉熵损失的方式缓解类不平衡, 定义为

$$Loss = \sum -w(y_i \log y_p - (1 - y_i) \log(1 - y_p)) \quad (9)$$

式中, y_i 为病例样本的真实类别, y_p 为该样本的预测类别。参数 w 为样本权值, 以增加对阳性样本分类错误的惩罚权重, 减少类不平衡的影响。

2.2.2 实验细节

SSAMNet模型在PyTorch平台上使用NVIDIA RTX 3090Ti显卡进行训练, 采用适应性矩估计(adaptive moment estimation, Adam)优化器(Kingma和Ba, 2015)在数据集MRNet上训练。训练时学习率初始化为 $1E-5$, 权重衰减设置为 $1E-4$, 每训练10轮, 学习率更新为原来的0.95倍。采用受试者工作特征曲线(receiver operating characteristic curve, ROC)的曲线下面积(area under curve, AUC)(Bradley, 1997)、准确性(accuracy)、敏感性(sensitivity)和特异性(specificity)作为评价指标。为保证实验公平性, 最终记录3次实验结果的平均值及其误差范围。

2.3 模型分析

2.3.1 对比实验

为了验证算法的有效性, 选取13个先进的网络模型在数据集MRNet上进行对比实验。分别是基于2D CNN的MRNet(Bien等, 2018)、ELNet(efficiently-layered network)(Tsai等, 2020)、AlexNet(Krizhevsky等, 2017)、ResNet18(He等, 2016)、VGG16(Simonyan和Zisserman, 2015)、GoogLeNet(Szegedy等, 2015)、MPFuseNet_S(Belton等, 2021)、MPFuseNet_M(Belton等, 2021)、Irmakci等人(2019)提出的模型、MRPyrNet_E(Dunnhofer等, 2022)、MRPyrNet_M(Dunnhofer等, 2022)以及Li等人(2023)和Jeon等人(2021)提出的基于3D CNN的方法。其中, MPFuseNet_S(Belton等, 2021)及MPFuseNet_M(Belton等, 2021)分别代表输入为单视图和多视图的MPFuseNet(Belton等, 2021)模型。MRPyrNet_E(Dunnhofer等, 2022)、MRPyrNet_M(Dunnhofer等, 2022)分别代表以ELNet(Tsai等, 2020)、MRNet(Bien等, 2018)为基础模型的MRPyrNet(Dunnhofer等, 2022)结构。实验结果如表1所示。虽然2D CNN能够考虑空间内容, 但对第三维度(切片维度)的建模能力较弱, 阻碍了性能提升。为了克服这个问题, 提出的SSAMNet在切片序列特征提取过程中综合考虑局部、全局切片关系信息, 其性能指标AUC明显优于其他基于2D CNN的方法。特别地, 即使与考虑切片维度的基于3D CNN的方法(Jeon等, 2021)相比, SSAMNet也有与之相当的性能, Jeon等人(2021)借助高斯位置编码辅助聚焦ACL区域, 可以保证较高的识别阳性样本的准确率, 但对阴性样本判别度不足, 在临床应用中极易导致

表 1 不同算法在 MRNet 数据集上的分类指标对比结果

Table 1 Comparison of classification metrics for different algorithms on the MRNet dataset

类别	模型	AUC	accuracy	sensitivity	specificity
3D CNN	Li 等人(2023)	0.920	0.833	0.870	0.808
	Jeon 等人(2021)	0.983	–	0.981	0.924
2D CNN	ResNet18	0.935	0.858	0.889	0.833
	AlexNet	0.942	0.825	0.673	0.950
	GoogLeNet	0.813	0.733	0.722	0.742
	VGG16	0.907	0.825	0.741	0.894
	MPFuseNet_S	0.949	0.908	0.852	0.955
	MPFuseNet_M	0.955	0.892	0.870	0.909
	Irmakci 等人(2019)	0.954	0.778	0.939	0.867
	MRNet	0.965	0.867	0.759	0.968
	ELNet	0.960	0.904	0.923	0.891
	MRPyrNet_E	0.960	0.881	0.827	0.924
	MRPyrNet_M	0.976	0.886	0.815	0.944
	SSAMNet(本文)	0.984	0.914	0.846	0.970

注:加粗字体表示各列最优结果,“–”代表缺少数据。

误诊,而 SSAMNet 的阴性样本识别准确率为 97%,可极大程度上避免该情况发生。相较于三视图 3D CNN(Li 等,2023),所提方法分类指标 AUC 及 accuracy 显著超过其 6.4%、8.1%。对于同样将多级尺度细节信息纳入考量的 MRPyrNet_E 和 MRPyrNet_M 来说,所提模型也在 AUC 指标上超过其 2.4%、0.8%。

2.3.2 消融实验

为了进一步研究不同模块对 SSAMNet 分类性能的增益,对所提模型进行消融实验,以验证模块有效性,结果如表 2 所示。具体来说,未使用 SFS 和 MSFAA 模块的经典 2D CNN 的骨干网络 AlexNet(Krizhevsky 等,2017)产生最低的分类性能指标,在

某种程度上表明骨干网络架构可能在 ACL 撕裂分类上作用有限。通过将 SFS 或 MSFAA 嵌入 AlexNet 中,发现 SSAMNet 的分类结果优于骨干网络,特别是增加 SFS 模块的 AUC 指标提升超过 2.5%。值得注意的是,SFS 和 MSFAA 的联合利用可以进一步带来显著的性能改善(AUC 超过 4%)。

这可以解释为,在添加 SFS 模块之后,能够更有效地捕捉 ACL 损伤特征与切片序列之间的关系,弥补骨干网络对切片维度序列信息挖掘的不足。SFS 和 MSFAA 的联合使用则同时把握空间尺度、切片维度特性,也就倾向于获得最佳性能。

2.3.3 参数 r 取值分析

SFS 中构建全局序列关系前筛选的切片数量 r

表 2 SSAMNet 在 MRNet 数据集上的消融实验结果

Table 2 Ablation experimental results of SSAMNet on the MRNet dataset

模型	SFS	MSFAA	AUC	accuracy	sensitivity	specificity
骨干网络	×	×	0.942±0.008	0.825±0.014	0.673±0.053	0.950±0.017
SSAMNet	×	√	0.943±0.004	0.864±0.021	0.809±0.011	0.909±0.040
SSAMNet	√	×	0.967±0.004	0.869±0.013	0.772±0.028	0.949±0.009
SSAMNet	√	√	0.984±0.002	0.914±0.013	0.846±0.043	0.970±0.015

注:加粗字体表示各列最优结果。“√”和“×”分别代表使用和不使用,“±”后的数据代表误差范围。

对最终的 ACL 损伤分类性能起着至关重要的作用。因此,必须探索合适的参数取值。图 5 分别列出了性能指标 AUC、accuracy、sensitivity 和 specificity 随 r 值增加呈现出的逐渐增加再递减的变化趋势,实验结果如表 3 所示,当 r 取值为 16 时,模型达到性能最优值。

2.3.4 外部数据集验证

表 4 为单视图 MRNet (Bien 等, 2018)、MRPyr-Net_M (Dunnhofer 等, 2022) 及 SSAMNet 在外部验证数据集 kneeMRI 中进行 5 折交叉验证的性能。即使是在与 MRNet 具有不同数据分布的数据集 kneeMRI

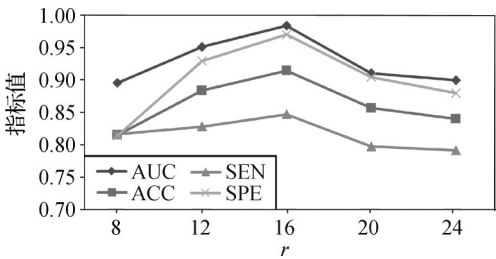


图 5 AUC、accuracy、sensitivity 及 specificity 随切片数量 r 的变化趋势

Fig. 5 The trends of auc, accuracy, sensitivity, and specificity with slice numbers (r)

上,所提出的 SSAMNet 结构分类性能也优于 MRPyr-Net_M,分类精度 accuracy 提升 3%。由于 kneeMRI 仅

表 3 参数 r 的不同取值在 MRNet 数据集上的分类指标对比结果

Table 3 Comparison of classification metrics for different values of parameter r on the MRNet dataset

r	AUC	accuracy	sensitivity	specificity
8	0.895±0.004	0.814±0.013	0.815±0.037	0.813±0.049
12	0.951±0.006	0.883±0.014	0.827±0.011	0.929±0.017
16	0.984±0.002	0.914±0.013	0.846±0.043	0.970±0.015
20	0.910±0.015	0.856±0.010	0.796±0.019	0.904±0.009
24	0.899±0.009	0.839±0.021	0.790±0.039	0.879±0.015

注:加粗字体表示各列最优结果,“±”后的数据代表误差范围。

表 4 不同算法在 kneeMRI 数据集上的 5 折交叉验证分类指标对比结果

Table 4 Comparison of classification metrics for different algorithms in fivefold cross-validation experiments on the kneeMRI dataset

模型	划分方式	AUC	accuracy	sensitivity	specificity
MRNet_S	K1	0.734	0.734	0.522	0.804
	K2	0.798	0.783	0.631	0.833
	K3	0.822	0.826	0.630	0.891
	K4	0.866	0.848	0.717	0.891
	K5	0.670	0.690	0.435	0.775
	均值	0.778	0.776	0.587	0.839
MRPyrNet_M	K1	0.894	0.870	0.804	0.891
	K2	0.878	0.880	0.630	0.964
	K3	0.877	0.783	0.826	0.768
	K4	0.887	0.810	0.783	0.819
	K5	0.888	0.870	0.783	0.899
	均值	0.885	0.842	0.765	0.868
SSAMNet	K1	0.880	0.886	0.761	0.928
	K2	0.880	0.853	0.739	0.891
	K3	0.870	0.870	0.761	0.906
	K4	0.900	0.875	0.783	0.906
	K5	0.906	0.875	0.739	0.920
	均值	0.887	0.872	0.757	0.910

注:加粗字体为均值部分的各列最优结果。

包含矢状位视图,所以使用单视图 MRNet 结构 (MRNet_S)进行实验。值得注意的是,如图6(b)所示,在不同划分的5次验证实验K1、K2、K3、K4、K5中,单视图 MRNet 的分类指标 AUC 出现较大波动(超过 19%),而图 6(a)所示的 SSAMNet 性能指标 AUC 走势平稳,起伏较小。这表明,SSAMNet 结构在不同数据分布上的稳健性及泛化性,代表着 SSAMNet 潜在的临床应用前景。

2.4 可视化分析

图 7 是在 MRNet 数据集上使用类激活映射 (class activation map, CAM) 技术生成的判别热图

(Zhou 等, 2016)。通过对 SSAMNet 获取的分类特征进行可视化处理,发现模型识别出 A 包含的 ACL 撕裂及胫骨水肿, B 产生的 ACL 撕裂对吻征。SSAMNet 和 MRPyrNet_M 相较于 MRNet_S 增加对切片图像多级尺度细节的挖掘,热力图显示模型对 ACL 撕裂及胫骨水肿区域具有更高的关注度。而在这 3 个模型中,仅 SSAMNet 中包含了切片序列信息融合,热力图可见该模型额外关注到 ACL 撕裂产生的对吻征,说明模型学习切片序列信息的重要性,进一步证实 SSAMNet 对 ACL 损伤分类任务的有效性。

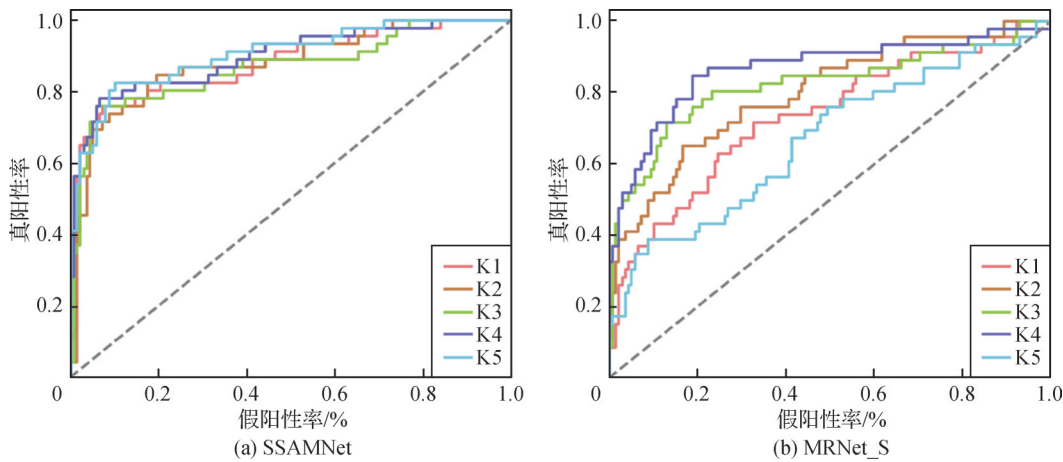


图6 SSAMNet 与 MRNet-S 在 kneeMRI 数据集上的 5 折交叉验证 ROC 曲线对比结果

Fig. 6 Comparison of ROC curves in five-fold cross-validation experiments on the kneeMRI between SSAMNet and MRNet-S ((a)SSAMNet; (b)MRNet_S)

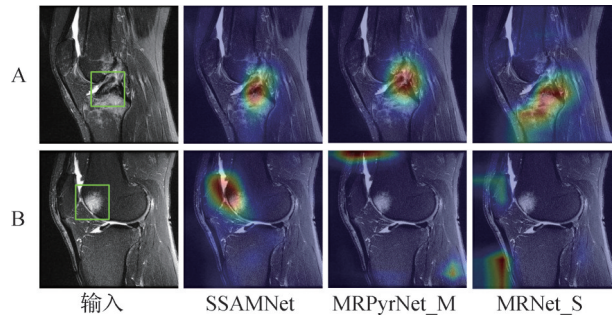


图7 SSAMNet、MRPyrNet_M 及 MRNet_S 可视化热图对比

Fig. 7 Comparison of SSAMNet, MRPyrNet_M, MRNet_S visualization results through CAM

3 结 论

本文基于切片数据整合和尺度重要性动态调整提出了 ACL 损伤分类网络 SSAMNet。通过聚合切片序列关系缓解 2D CNN 捕获三维特征的局限性;关于 ACL 形态存在变动性和不规则性的问题,借助动

态分配尺度权重的方式,增强重点尺度与非重点尺度间的显著差异,提高 ACL 损伤判别精确率。对比实验与消融实验结果证明了所提算法的有效性。在不同数据分布的数据集上的验证实验结果也表明 SSAMNet 具有泛化性及临床应用性。目前该模型存在使用的数据集标签单一的问题,后续将研究融合患者既往病历等先验知识到模型框架中,搭建辅助医疗框架。

参考文献 (References)

Belton N, Welaratne I, Dahlan A, Hearne R T, Hagos M T, Lawlor A and Curran K M. 2021. Optimising knee injury detection with spatial attention and validating localisation ability//Proceedings of the 25th Annual Conference on Medical Image Understanding and Analysis. Oxford, UK: Springer: 71-86 [DOI: 10.1007/978-3-030-80432-9_6]
Bien N, Rajpurkar P, Ball R L, Irvin J, Park A, Jones E, Bereket M, Patel B N, Yeom K W, Shpanskaya K, Halabi S, Zucker E, Fan-

- ton G, Amanatullah D F, Beaulieu C F, Riley G M, Stewart R J, Blankenberg F G, Larson D B, Jones R H, Langlotz C P, Ng A Y and Lungren M P. 2018. Deep-learning-assisted diagnosis for knee magnetic resonance imaging: development and retrospective validation of MRNet. *PLoS Medicine*, 15 (11): #1002699 [DOI: 10.1371/journal.pmed.1002699]
- Bradley A P. 1997. The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern Recognition*, 30(7): 1145-1159 [DOI: 10.1016/S0031-3203(96)00142-2]
- Chang P D, Wong T T and Rasiej M J. 2019. Deep learning for detection of complete anterior cruciate ligament tear. *Journal of Digital Imaging*, 32: 980-986 [DOI: 10.1007/s10278-019-00193-4]
- Crawford R, Walley G, Bridgman S and Maffulli N. 2007. Magnetic resonance imaging versus arthroscopy in the diagnosis of knee pathology, concentrating on meniscal lesions and ACL tears: a systematic review. *British Medical Bulletin*, 84(1): 5-23 [DOI: 10.1093/bmb/ldm022]
- Dunnhofer M, Martinel N and Micheloni C. 2022. Deep convolutional feature details for better knee disorder diagnoses in magnetic resonance images. *Computerized Medical Imaging and Graphics*, 102: #102142 [DOI: 10.1016/j.compmedimag.2022.102142]
- He K M, Zhang X Y, Ren S Q and Sun J. 2016. Deep residual learning for image recognition//*Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Las Vegas, USA: IEEE: 770-778 [DOI: 10.1109/cvpr.2016.90]
- Hou Q B, Zhou D Q and Feng J S. 2021. Coordinate attention for efficient mobile network design//*Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Nashville, USA: IEEE: 13708-13717 [DOI: 10.1109/cvpr46437.2021.01350]
- Hu J, Shen L and Sun G. 2018. Squeeze-and-excitation networks//*Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Salt Lake City, USA: IEEE: 7132-7141 [DOI: 10.1109/CVPR.2018.00745]
- Huang G, Liu Z, Van Der Maaten L and Weinberger K Q. 2017. Densely connected convolutional networks//*Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Honolulu, USA: IEEE: 2261-2269 [DOI: 10.1109/cvpr.2017.243]
- Huang S H and Wang J F. 2022. Brain MRI image classification incorporating smooth group sparseness. *Journal of Image and Graphics*, 27(3): 885-897 (黄帅辉, 王金凤. 2022. 融入平滑组稀疏化的脑部MRI图像分类. *中国图象图形学报*, 27(3): 885-897) [DOI: 10.11834/jig.210367]
- Irmakci I, Anwar S M, Torigian D A and Bagci U. 2019. Deep learning for musculoskeletal image analysis//*Proceedings of the 53rd Asilomar Conference on Signals, Systems, and Computers*. Pacific Grove, USA: IEEE: 1481-1485 [DOI: 10.1109/IEEECONF44664.2019.9048671]
- Jeon Y S, Yoshino K, Hagiwara S, Watanabe A, Quek S T, Yoshioka H and Feng M L. 2021. Interpretable and lightweight 3-D deep learning model for automated ACL diagnosis. *IEEE Journal of Biomedical and Health Informatics*, 25 (7): 2388-2397 [DOI: 10.1109/JBHI.2021.3081355]
- Kingma D P and Ba J L. 2015. Adam: a method for stochastic optimization//*Proceedings of the 3rd International Conference on Learning Representations*. San Diego, USA: ICLR [DOI: 10.48550/arXiv.1412.6980]
- Krizhevsky A, Sutskever I and Hinton G E. 2017. ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6): 84-90 [DOI: 10.1145/3065386]
- Li F, Zhai P H, Yang C, Feng G, Yang J and Yuan Y. 2023. Automated diagnosis of anterior cruciate ligament via a weighted multi-view network. *Frontiers in Bioengineering and Biotechnology*, 11: #1268543 [DOI: 10.3389/fbioe.2023.1268543]
- Liu Z Y, Wang L M, Wu W, Qian C and Lu T. 2021. TAM: temporal adaptive module for video recognition//*Proceedings of 2021 IEEE/CVF International Conference on Computer Vision (ICCV)*. Montreal, Canada: IEEE: 13688-13698 [DOI: 10.1109/iccv48922.2021.01345]
- Mackenzie R, Palmer C R, Lomas D J and Dixon A K. 1996. Magnetic resonance imaging of the knee: diagnostic performance statistics. *Clinical Radiology*, 51(4): 251-257 [DOI: 10.1016/S0009-9260(96)80341-2]
- Oei E H G, Nikken J J, Verstijnen A C M, Ginai A Z and Myriam Hunink M G. 2003. MR imaging of the menisci and cruciate ligaments: a systematic review. *Radiology*, 226(3): 837-848 [DOI: 10.1148/radiol.2263011892]
- Shea K G and Carey J L. 2015. Management of anterior cruciate ligament injuries: evidence-based guideline. *Journal of the American Academy of Orthopaedic Surgeons*, 23(5): e1-e5 [DOI: 10.5435/jaaos-d-15-00094]
- Simonyan K and Zisserman A. 2015. Very deep convolutional networks for large-scale image recognition//*Proceedings of the 3rd International Conference on Learning Representations*. San Diego, USA: ICLR [DOI: 10.48550/arXiv.1409.1556]
- Štajduhar I, Mamula M, Miletić D and Ūnal G. 2017. Semi-automated detection of anterior cruciate ligament injury from MRI. *Computer Methods and Programs in Biomedicine*, 140: 151-164 [DOI: 10.1016/j.cmpb.2016.12.006]
- Szegedy C, Liu W, Jia Y Q, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V and Rabinovich A. 2015. Going deeper with convolutions//*Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Boston, USA: IEEE: 1-9 [DOI: 10.1109/CVPR.2015.7298594]
- Tsai C H, Kiryati N, Konen E, Eshed I and Mayer A. 2020. Knee injury detection using MRI with efficiently-layered network (ELNet)//*Proceedings of the 3rd Conference on Medical Imaging*

- with Deep Learning. Montreal, Canada: PMLR: 784-794
- Yu B and Garrett W E. 2007. Mechanisms of non-contact ACL injuries. British Journal of Sports Medicine, 41(S1): 47-51 [DOI: 10.1136/bjsm.2007.037192]
- Zhang L Y, Li M F, Zhou Y J, Lu G M and Zhou Q. 2020. Deep learning approach for anterior cruciate ligament lesion detection: evaluation of diagnostic performance using arthroscopy as the reference standard. Journal of Magnetic Resonance Imaging, 52(6): 1745-1752 [DOI: 10.1002/jmri.27266]
- Zhao X M, Liao Y H, Zhang S Q, Fang J X, He X X, Wang G Y and Lu H S. 2024. Method of classifying benign and malignant breast tumors by DCE-MRI incorporating local and global features. Journal of Image and Graphics, 29(1): 256-267 (赵小明, 廖越辉, 张石清, 方江雄, 何遐遐, 汪国余, 卢洪胜. 2024. 融合局部与全局特征的DCE-MRI乳腺肿瘤良恶分类. 中国图象图形学报, 29(1): 256-267) [DOI: 10.11834/jig.230092]
- Zhou B L, Khosla A, Lapedriza A, Oliva A and Torralba A. 2016. Learning deep features for discriminative localization//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, USA: IEEE: 2921-2929 [DOI: 10.1109/cvpr.2016.319]

作者简介

刘英莉,女,副教授,主要研究方向为医学图像智能应用。

E-mail:lyl@kust.edu.cn

高明,通信作者,男,主治医师,主要研究方向为健康数据科学及骨关节退行性变。E-mail:drgaoming@pku.org.cn

茶寅秋,女,硕士研究生,主要研究方向为医学图像处理。

E-mail:chayq@stu.kust.edu.cn

黄一山,男,硕士研究生,主要研究方向为医学图像处理。

E-mail:1624115542@qq.com