

基于时空特征与动态融合的多序列MRI肝脏肿瘤分类

黄龙光*, 邵虹

沈阳工业大学信息科学与工程学院, 辽宁 沈阳

收稿日期: 2026年2月27日; 录用日期: 2026年3月19日; 发布日期: 2026年3月30日

摘要

多序列磁共振成像在肝脏肿瘤诊断中提供了丰富的时空信息, 但现有分析方法存在跨序列信息交互不足和时空特征整合不充分的问题。针对该问题, 本文提出了一种基于时空特征与动态融合的Transformer模型。该模型根据多序列MRI的成像特性对强相关序列进行分组处理, 通过动态瓶颈桥接模块在不同序列间建立信息交互通道, 结合动态权重机制调节序列重要性; 同时采用时空特征协同融合模块利用双向注意力机制实现空间与时间特征的语义对齐, 并通过融合策略实现时空特征的有效整合。在LLD-MRI 2023数据集上的实验表明, 该方法取得了84.62%的准确率、96.58%的曲线下面积、84.04%的F1分数和81.13%的Kappa系数, 在多数指标上优于现有方法。

关键词

肝脏肿瘤分类, 多序列MRI, 动态特征融合, Transformer, 瓶颈桥接

Multi-Sequence MRI Liver Tumor Classification Based on Spatio-Temporal Features and Dynamic Fusion

Longguang Huang*, Hong Shao

School of Information Science and Engineering, Shenyang University of Technology, Shenyang Liaoning

Received: February 27, 2026; accepted: March 19, 2026; published: March 30, 2026

Abstract

Multi-sequence magnetic resonance imaging provides rich spatio-temporal information for liver

*通讯作者。

tumor diagnosis, but existing analysis methods suffer from insufficient cross-sequence information interaction and inadequate spatio-temporal feature integration. To address these problems, a Transformer model based on spatio-temporal features and dynamic fusion is proposed in this paper. The model groups strongly correlated sequences according to the imaging characteristics of multi-sequence MRI, establishes information interaction channels between different sequences through the Dynamic Bottleneck Bridge module, and adjusts sequence importance by combining a dynamic weighting mechanism. Meanwhile, the Spatio-Temporal Fusion module is employed to achieve semantic alignment of spatial and temporal features using a bidirectional attention mechanism and realizes effective integration of spatio-temporal features through a fusion strategy. Experiments on the LLD-MRI 2023 dataset demonstrate that the method achieves 84.62% accuracy, 96.58% area under the curve, 84.04% F1-score, and 81.13% Kappa coefficient, outperforming existing methods on most metrics.

Keywords

Liver Tumor Classification, Multi-Sequence MRI, Dynamic Feature Fusion, Transformer, Bottleneck Bridge

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

肝癌是目前我国第四位常见恶性肿瘤及第二位肿瘤致死病因, 严重威胁我国人民的生命和健康[1]。在肝脏肿瘤诊断中, 磁共振成像(Magnetic Resonance Imaging, MRI)可以提供丰富的病变信息, 各个序列具有不同的成像特点和诊断价值[2] [3]。

早期多序列 MRI 分析方法主要处理单一序列类型。针对非增强序列, Wojciechowska 等[4]、Chlebus 等[5]和 Trivizakis 等[6]采用传统深度学习网络提取特征。例如, Wojciechowska 等仅利用 T1 序列区分积液囊肿与其他局灶性肝脏病变, 然而, 单一序列的信息量有限, 缺乏多模态信息的协同作用, 影响了诊断的准确性。Mo 等[7]和 Xiao 等[8]引入了额外序列信息并增强了边缘感知能力, 但无法扩展到多序列分析, 存在序列信息利用不充分的问题。针对对比增强序列, Hu 等[9]、Qiu 等[10]、Kumar 等[11]和 Phan 等[12]通过图像级融合或并行卷积核提取增强特征, 但忽略了序列间的时序变化信息, 未能评估不同肝脏肿瘤的动态病理特征。其中, Hu 等提出的双输入网络分别处理静脉期和动脉期影像, 采用稀疏连接提升计算效率, 但在多期相特征融合时采用简单拼接方式, 未能筛选特征重要性, 导致保留了大量无关的正常组织信息。Kumar 等引入空间 - 通道联合注意力模块, 结合位置编码与跨相位交互突出关键区域特征, 但这些方法仍忽略了序列间的时序变化信息, 未能充分评估不同肝脏肿瘤的动态病理特征。

随着技术发展, 研究者开始整合增强与非增强序列信息。Wang 等[13]使用 3D ResNet 处理七个 MRI 序列, 但重采样策略可能丢失细粒度结构信息, 且独立编码器架构导致参数冗余和融合效率低下。近期方法尝试改进这些问题: Lou 等[14]提出了 SDR-Former 模型, 采用双分辨率 CNN-Transformer 架构, 通过专用分支处理不同分辨率序列, 但缺乏不同序列间的信息交互, 未充分利用不同序列之间的互补信息; Wang 等[15]提出 LCA-DB 模型, 基于多实例学习的空间信息建模, 但仅依赖局部注意力权重融合, 难以有效处理特征冗余问题; Huang 等[16]提出 STM-Former 模型, 结合瓶颈桥接与双向 LSTM 建模时序模式, 但依赖手工特征且采用简单平均融合, 无法充分发挥强相关序列的互补优势, 同时难以减轻多类别

病变信息混叠造成的干扰; Hu 等[17]提出了 Auto-Classification 模型, 通过注意力引导数据增强提升多级信息利用, 但主要关注多分辨率特征, 忽略了序列间时序关联的重要性。

现有方法主要存在两方面不足: 一是缺乏有效的跨序列信息交互机制, 多数方法采用独立处理或简单拼接策略, 未能充分利用不同序列间的互补信息, 同时采用静态权重分配难以根据样本特征进行自适应调节; 其次, 时空特征整合不充分, 未能建立有效的序列间语义对应关系, 特别是忽略了 MRI 序列中丰富的时间动态信息。

为解决上述问题, 本文提出基于时空特征和动态特征融合的 Transformer 模型(Dynamic Spatio-Temporal Fusion Former, DSTF-Former)。该模型首先根据多期相 MRI 的成像特性将强相关序列进行分组处理, 通过设计两个核心模块实现有效的特征交互和融合: 设计了动态瓶颈特征桥接(Dynamic Adaptive Bottleneck Bridge, DABB)模块, 通过少量可学习的瓶颈特征在不同序列间建立信息交互通道, 并结合动态权重机制自适应调节序列的重要性, 有效缓解了多序列 MRI 特征冗余问题; 同时设计时空特征融合(Spatio-Temporal Fusion, STF)模块, 利用双向注意力机制实现空间与时间特征的语义对齐, 并通过自适应融合策略充分捕捉动态增强序列的时序模式, 从而提升模型对相似病灶的区分能力。

2. 本文方法

2.1. 整体网络结构

本文提出一种 DSTF-Former 方法, 用于多序列 MRI 条件下的肝脏肿瘤分类, 其整体框架如图 1 所示。该方法在多序列 MRI 中同时建模空间特征与时间动态信息, 实现时空特征的协同学习。

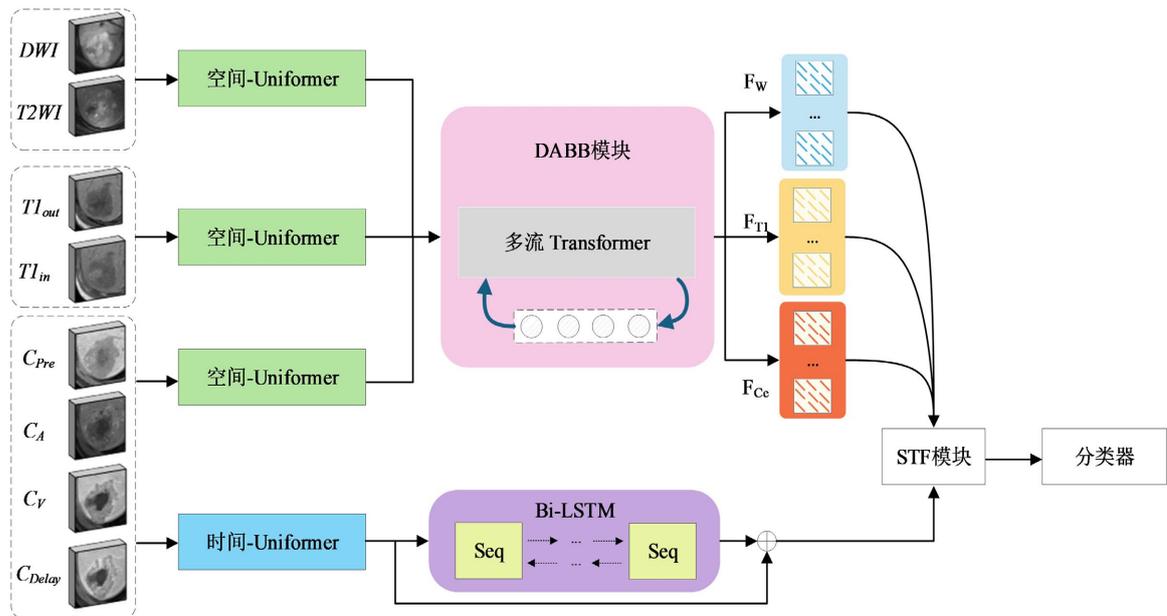


Figure 1. Structural diagram of the DSTF-Former

图 1. DSTF-Former 结构图

根据不同 MRI 序列的成像机制与诊断价值, 本文将多序列影像划分为三个处理流, 以突出互补信息并降低冗余。第一组包括 DWI 与 T2WI 序列, 用于表征肿瘤组织结构及瘤周水肿信息; 第二组为 T1WI 正相位与反相位序列, 主要反映病灶内部成分差异; 第三组为动态增强 MRI 的四个期相, 用于刻画肿瘤血流动力学特征及其随时间的变化模式。通过该分组策略, 模型在保持各序列物理属性差异的同时, 实

现跨序列特征交互。

整体网络主要由多流空间编码器、DABB 模块、时间编码器和 STF 模块构成。多流空间编码器采用三支并行的结构, 分别提取不同序列组的空间特征。DABB 模块在各空间分支之间建立跨序列特征交互, 并通过动态权重机制实现特征融合, 从而增强判别性空间特征表达。

时间编码器对动态增强序列进行时间建模, 捕获肿瘤强化过程中的时序演变特征。STF 模块进一步融合空间与时间特征, 通过注意力机制实现时空特征对齐与加权, 形成统一的特征表示。最终, 融合后的特征经分类器输出肝脏肿瘤类别。

2.2. DABB 模块

在多序列医学影像分类任务中, 不同 MRI 序列通常包含互补信息, 但也不可避免地引入冗余, 从而对模型的判别能力产生干扰。此外, 现有多序列融合方法主要存在两类不足: 一方面, 简单的特征拼接虽然能够整合多序列信息, 却显著增加了计算复杂度; 另一方面, 固定权重的融合策略忽略了不同样本中各序列贡献度的差异性, 难以实现对序列重要性的自适应调节。为解决上述问题, 本文提出了 DABB 模块, 其结构如图 2 所示。

该模块的核心思想是引入少量可学习的瓶颈特征 $B = [b_1, b_2, b_3, \dots, b_{n_b}] \in R^{n_b \times C}$, 其中 n_b 为瓶颈特征的数量, C 为特征维度, B 为跨序列信息交互的桥梁。这些轻量级的特征仅承担信息传递与整合的作用, 从而在保持各分支特征独立建模的同时, 实现跨模态的融合。此时, 输入到多流 Transformer 中的第 i 个流的输入特征表示如公式(1)所示:

$$Z_i = [F_i; B_i] \in R^{(T_i+n_b) \times C} \quad (1)$$

其中, $F_i \in R^{T_i \times C}$ 定义为:

$$F_i = \{F_w, i=1; F_T, i=2; F_C, i=3\}$$

在该模块中, 本文将所有跨通道注意力交互限制在这些瓶颈特征上。对于第 l 层, 更新过程如公式(2)、公式(3)所示:

$$\hat{Z}_i^l = \text{MHA}(\text{LN}(Z_i^l)) + Z_i^l \quad (2)$$

$$Z_i^{l+1} = \text{MLP}(\text{LN}(\hat{Z}_i^l)) + \hat{Z}_i^l \quad (3)$$

其中, LN 表示归一化层, MHA 表示多头注意力机制。 Z_i^l 和 \hat{Z}_i^l 分别表示第 l 层中第 i 个流的特征表示和其更新后的中间表示。

在特征更新后, 首先对每个分支输出的瓶颈特征 B_i^l 进行全局平均池化, 得到各自的全局表示。进一步对其平均得到上下文语义向量, 并通过 MLP 层得到动态调整因子 $\delta \in [-1, 1]$ 。计算过程如公式(4)所示:

$$\delta = \text{MLP}\left(\frac{1}{3} \sum_{i=1}^3 \text{Avg}(B_i^l)\right) \quad (4)$$

其中, B_i^l 表示第 l 层中第 i 个流输出的瓶颈特征。最后, 动态权重的计算如公式(5)所示:

$$Q_i^l = \text{Softmax}(\alpha_b + \lambda \delta) \quad (5)$$

其中, Q_i^l 为对应 B_i^l 的权重, α_b 是可学习的基础权重参数, 初始取值为[0.33, 0.33, 0.34]。 λ 为调整强度系数, 控制动态调整的幅度, 本文基于经验性分析设置为 0.15, 该取值在保证训练稳定性的同时, 允许适度的动态调整。

瓶颈特征最后进行加权特征融合, 如公式(6)所示:

$$B^{l+1} = \sum_{i=1}^3 w_i^l B_i^l \tag{6}$$

其中, w_i^l 为第 l 层中第 i 个分支对应的动态权重系数。

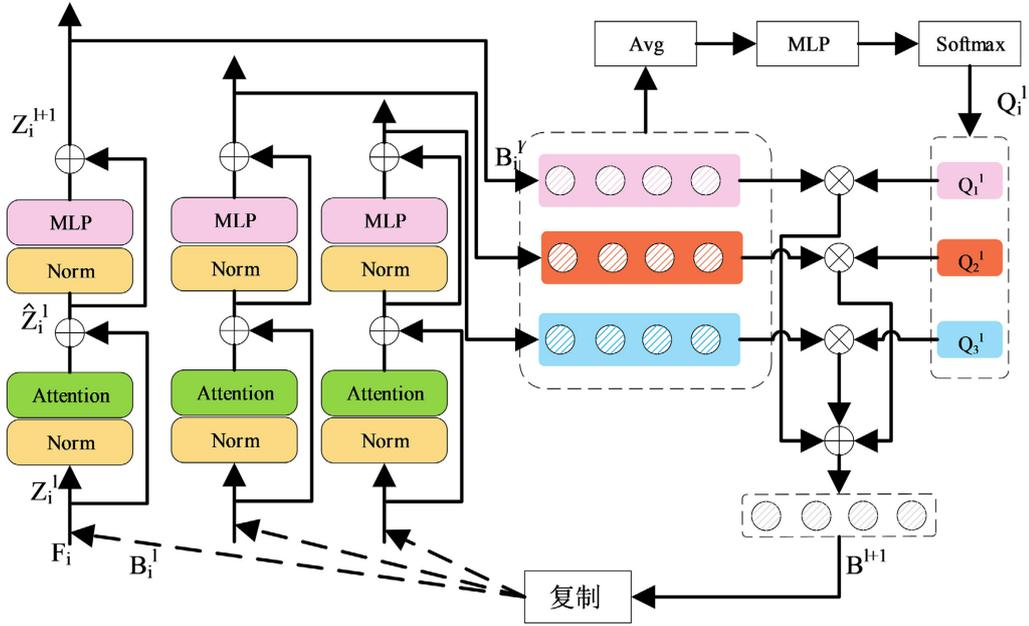


Figure 2. Structural diagram of the DABB module
图 2. DABB 模块结构图

2.3. STF 模块

在多序列医学影像分析中, 空间编码模块提取的多流特征与时间编码模块获得的时序特征在语义层面存在显著差异。空间特征关注病灶的形态学特征和空间分布, 而时间特征则捕获序列间的时序动态变化模式。传统的特征融合方法通常采用简单的拼接或平均策略, 忽略了不同模态间的语义关联性和互补性, 难以充分利用时空特征的协同效应。为解决上述问题, 本文提出了 STF 模块, 其结构如图 3 所示。

该模块包含时空特征对齐和协同融合两个主要部分。时空特征对齐部分通过双向注意力机制建立空间与时间特征的语义对应关系, 协同融合部分实现增强特征的有效整合。设输入的空间特征为 $F_{sp} \in R^{3 \times C}$, 时间特征为 $F_t \in R^{k \times C}$ 。

在时空特征对齐部分, 为建立空间 - 时间特征的双向依赖关系, 模块通过计算不同的查询、键、值矩阵实现时间特征对空间特征和空间特征对时间特征的交互, 分别得到空间特征 $F_{t \rightarrow s}$ 和时间特征 $F_{s \rightarrow t}$, 计算过程如公式(7)、公式(8)所示:

$$F_{t \rightarrow s} = \text{softmax} \left(\frac{Q_t K_s^T}{\sqrt{d_k}} \right) V_s \tag{7}$$

$$F_{s \rightarrow t} = \text{softmax} \left(\frac{Q_s K_t^T}{\sqrt{d_k}} \right) V_t \tag{8}$$

其中, Q_t 、 K_s 、 V_s 分别为时间特征的查询矩阵和空间特征的键、值矩阵, Q_s 、 K_t 、 V_t 为相应的反向投

影, d_k 为注意力头的特征维度。增强的特征表示通过残差连接计算得到, 如公式(9)、公式(10)所示:

$$F_{RS} = F_{sp} + F_{s \rightarrow t} \quad (9)$$

$$F_{RT} = F_E^t + F_{t \rightarrow s} \quad (10)$$

将对齐后的时空特征进行融合。首先对增强的空间特征 F_{RS} 展平为 $F_s^{flat} \in R^{3C}$, 再通过 MLP 网络进行特征整合, 计算公式如公式(11)所示:

$$F_s^{fused} = \text{MLP}(\text{Reshape}(F_{RS})) \in R^C \quad (11)$$

随后, 将融合后的空间特征与增强的时间特征拼接得到特征 F_{Con} , 计算过程如公式(12)所示:

$$F_{Con} = [F_s^{fused}; F_{RT}] \in R^{2C} \quad (12)$$

得到融合输入后, 利用自适应权重网络计算空间与时间模式的相对重要性, 计算过程如公式(13)所示:

$$w = \text{Softmax}(W_2 \text{ReLU}(W_1 F_{Con} + b_1) + b_2) \quad (13)$$

其中, W_1 和 W_2 为权重矩阵, b_1 和 b_2 为偏置向量。最后使用 F_s^{fused} 和 F_{RT} 加权得到输出的时空融合特征 F_{out} , 计算过程如公式(14)所示:

$$F_{out} = w_s \times F_s^{fused} + w_t \times F_{RT} \quad (14)$$

其中, w_s 和 w_t 为融合权重。

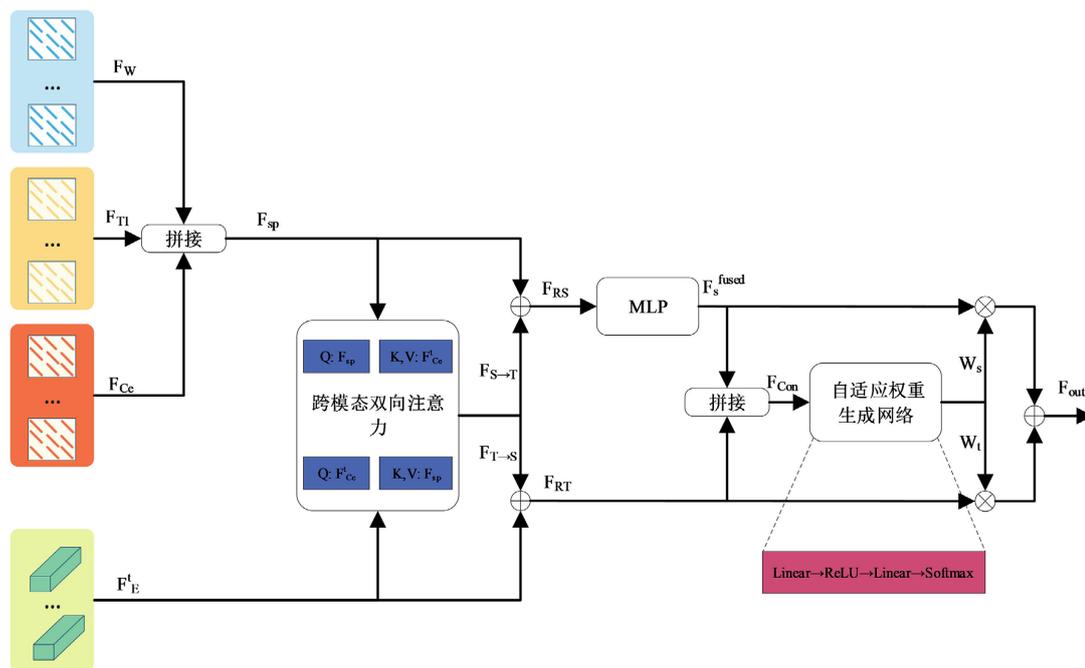


Figure 3. Structural diagram of the STF module

图 3. STF 模块结构图

3. 实验结果分析

3.1. 数据集介绍

LLD-MRI 2023 数据集源自 2023 年 MICCAI 肝脏病灶诊断挑战赛, 旨在推动计算机辅助诊断系统的

发展。数据集包含 498 例带注释的多序列肝脏 MRI 影像, 涵盖七类肝脏病灶: 肝细胞癌(HCC)、肝内胆管癌(ICC)、肝转移瘤(HM)、肝囊肿(HC)、肝血管瘤(HH)、局灶性结节增生(FNH)和肝脓肿(HA)。影像包括八个不同的成像阶段。数据集已预先划分为训练集(316 例)、验证集(78 例)和测试集(104 例)。七类病灶在训练集、验证集和测试集中的病例数如表 1 所示。

Table 1. Sample counts of seven tumor categories in dataset partitions

表 1. 七类病灶在数据集划分中的样本数量

类别	HH	ICC	HA	HM	HC	FNH	HCC
训练集	50	37	34	32	34	29	100
验证集	13	9	8	8	8	7	25
测试集	16	13	12	10	11	10	32

3.2. 训练设置及实验设备

实验环境采用 PyTorch 深度学习框架, 采用两块 NVIDIA RTX 3090 (24 GB) GPU 并行训练。训练中, 采用 AdamW 优化器, 初始学习率为 0.0001, 并使用余弦退火调度动态调整学习率以优化性能。为防止过拟合, 应用了 0.05 的权重衰减。训练过程持续 300 轮, 其中前 5 个周期作为热身阶段, 学习率逐渐提升。使用标准的交叉熵损失函数比较模型的输出与真实标签。由于硬件限制, batch_size 大小设置为 4, 每个病变体积统一调整为 $16 \times 128 \times 128$ 。为了进一步减轻过拟合, 采用了多种数据增强技术, 包括随机旋转、擦除和沿不同解剖轴翻转。在训练阶段, 随机裁剪病变体积为 $14 \times 112 \times 112$, 在评估时提取相同大小的中心裁剪。

为评估模型的临床部署可行性, 在测试集上测试了模型的推理效率。在上述硬件环境下, 模型对 114 例测试样本的总推理耗时为 40.8 秒, 平均单例推理时间约为 357 ms。

3.3. 评价指标

为验证模型的有效性, 本文采用准确率、AUC、F1 分数和 Kappa 系数来评价分类结果。其中, 准确率衡量模型对正类预测的精确程度, 其计算方式如公式(15)所示:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (15)$$

F1 分数为准确率与召回率的调和平均, 用于衡量分类结果在精确性与完整性之间的平衡, 计算方式如公式(16)所示:

$$\text{F1} = \frac{2 \times (\text{Precision} \times \text{Recall})}{(\text{Precision} + \text{Recall})} \quad (16)$$

Kappa 系数用于评估分类结果与真实标签之间的一致性, 其计算方式如公式(17)所示:

$$\text{Kappa} = \frac{P_o - P_e}{1 - P_e} \quad (17)$$

其中, TP 表示真阳性, 即模型将正类别样本正确分类为正类别, FP 表示假阳性, 即模型将负类别样本错误分类为正类别。FN 表示假阴性, 即模型将正类别样本错误分类为负类别。 P_o 是观察到的比例一致性, 即评估者在所有样本上达成一致的比率。 P_e 是预期的随机一致性比例, 即在假设评估者独立且随机进行分类的情况下, 他们达成一致的比率。

3.4. 对比实验

本文提出的 DSTF-Former 模型与多个经典和先进的模型进行了对比, 包括 ResNet2D-50 [18]、DenseNet-121 [19]、MSCSNN [23]、ResNet3D-50 [13]、BoTNet-50 [20]、H2Former [22]、UniFormer [21]、SegMamba [24]、LCA-DB [15]、SDR-Former [14]、STM-Former [16]和 Auto-Classification [17]。其中, H2Former 和 SegMamba 主要用于分割, 因此在实验中只使用了它的编码器, STM-Former 中未使用手工提取的放射组学特征。

各模型指标对比如表 2 所示。从表中可以看出, 本文方法在所有评价指标上均取得最优结果。与传统卷积网络相比, 引入 Transformer 编码结构的模型整体性能更优, 表明其在多期相 MRI 特征建模中的优势。与基于单一编码器的模型相比, 多分支编码结构能够更充分地挖掘不同期相之间的互补信息。

在此基础上, 本文方法通过自适应特征融合策略进一步提升了模型性能, 说明在多序列特征建模过程中, 同时考虑时空信息对齐与动态权重分配, 有助于增强模型对关键判别信息的表达能力。

Table 2. Comparison of performance metrics in comparative experiments

表 2. 对比实验性能指标对比

方法	ACC	AUC	F1	Kappa
ResNet2D-50 [18]	0.6923	0.9298	0.6898	0.6244
DenseNet-121 [19]	0.7404	0.9346	0.7171	0.6797
MSCSNN [23]	0.7115	0.9379	0.7089	0.6536
ResNet3D-50 [13]	0.6730	0.9480	0.6590	0.5980
BoTNet-50 [20]	0.7212	0.9314	0.7139	0.6628
H2Former [22]	0.7212	0.9311	0.7342	0.6660
UniFormer [21]	0.7115	0.9021	0.7115	0.6433
SegMamba [24]	0.6827	0.9078	0.6753	0.6065
LCA-DB [15]	0.6925	0.8856	0.6599	-
SDR-Former [14]	0.7885	0.9536	0.7910	0.7467
STM-Former [16]	0.7788	0.9330	0.7649	0.7312
Auto-Classification [17]	0.8590	0.9710	0.8390	-
本文方法(DSTF-Former)	0.8462	0.9658	0.8404	0.8113

3.5. 消融实验

3.5.1. 消融实验指标分析

为验证各模块的有效性, 在相同实验设置下对模型进行了消融实验, 结果如表 3 所示。基线模型由三支 UniFormer 空间编码器与单分支时间编码器组成。

Table 3. Comparison of performance metrics in ablation experiments

表 3. 消融实验性能指标对比

方法	ACC	AUC	F1	Kappa	Params	FLOPs
Base	0.7692 ± 0.028	0.9312 ± 0.025	0.7576 ± 0.026	0.7183 ± 0.027	90.72 M	172.56 G
Base + STF	0.8077 ± 0.024	0.9491 ± 0.022	0.7689 ± 0.023	0.7669 ± 0.024	102.58 M	194.91 G
Base + DABB	0.8173 ± 0.021	0.9526 ± 0.019	0.7880 ± 0.020	0.7749 ± 0.021	92.38 M	175.52 G
Base + STF + DABB	0.8462 ± 0.015	0.9658 ± 0.013	0.8404 ± 0.014	0.8113 ± 0.015	104.24 M	198.06 G

从表中可以看出, 引入任一模块均可提升模型性能。加入 STF 模块后, 准确率、AUC、F1 分数和 Kappa 系数均有所提高, 说明该模块能够有效融合空间与时间特征, 增强模型整体判别能力。加入 DABB 模块后, 模型性能进一步提升, 表明跨序列特征交互与动态权重机制有助于抑制冗余信息并突出关键序列特征。

当同时引入 STF 和 DABB 模块时, 模型在各项评价指标上均取得最优结果, 验证了两个模块在多序列 MRI 特征建模中的互补性与协同作用。

3.5.2. 混淆矩阵可视化与分析

为进一步分析不同模块对分类性能的影响, 对消融实验的混淆矩阵进行了对比, 如图 4 所示。基线模型在部分病灶类别间存在明显混淆, 尤其是 HCC、HM 与 ICC 之间的误分类较为突出。

引入 STF 模块后, 整体分类性能有所改善, 部分类别的误分情况得到缓解, 但对影像特征高度相似类别的区分能力仍有限。引入 DABB 模块后, HCC 与 ICC 的识别准确率进一步提升, 说明跨序列特征交互有助于增强模型对关键信息的建模能力。

当 STF 与 DABB 模块联合使用时, 各类别的分类性能均得到进一步改善, 混淆现象明显减少, 表明时空特征对齐与跨序列动态融合的协同作用能够有效提升模型的整体判别能力。

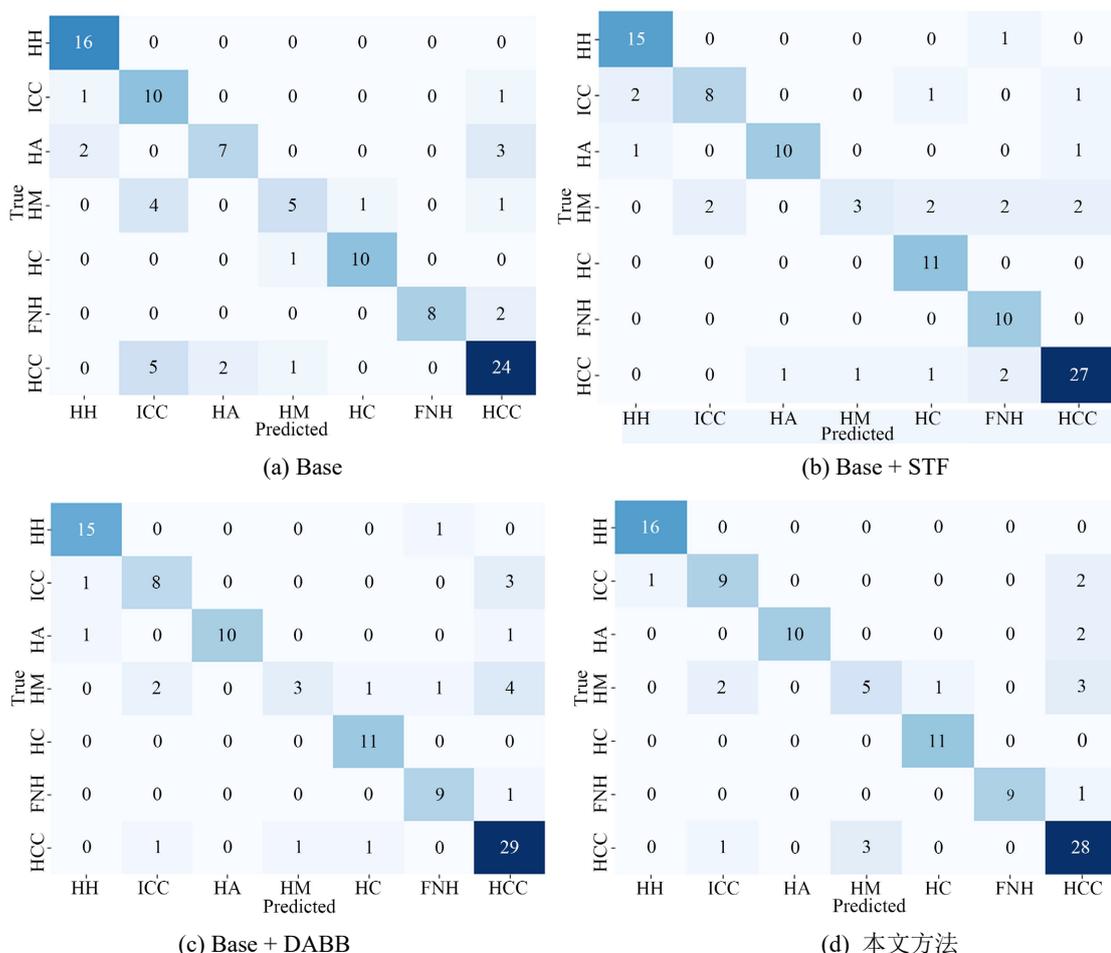


Figure 4. Comparison of confusion matrices of different models on the liver tumor classification task

图 4. 不同模型在肝脏肿瘤分类任务中的混淆矩阵对比

进一步对 ICC 混淆为 HCC 的案例分析, 其对比增强图像如图 5 所示, 部分被误判为 HCC 的 ICC 病

例在动态增强过程中未表现出典型的渐进性强化模式，而在动脉期呈现较为明显的整体强化表现，并在后续期相中出现一定程度的强化减退趋势。其时间维度上的强化变化更接近“快进快出”的模式，而未体现典型 ICC 的持续强化特征。此外，该类病灶未显示清晰的靶样强化结构，边缘强化不典型，使得其影像学表现与 HCC 存在一定重叠。在此情况下，模型在时间建模过程中更倾向于依据强化强度变化趋势进行判别，从而产生误分类现象。

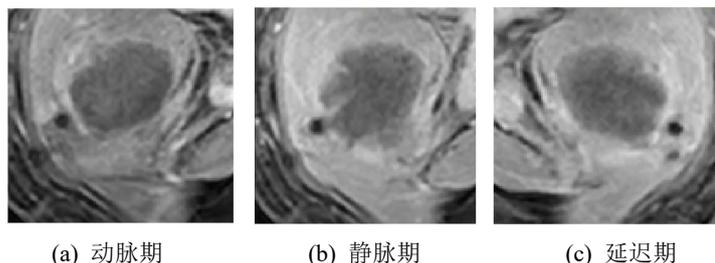


Figure 5. Misclassified ICC on enhanced image of HCC
图 5. ICC 误判为 HCC 的增强图像

3.5.3. t-SNE 可视化与聚类指标分析

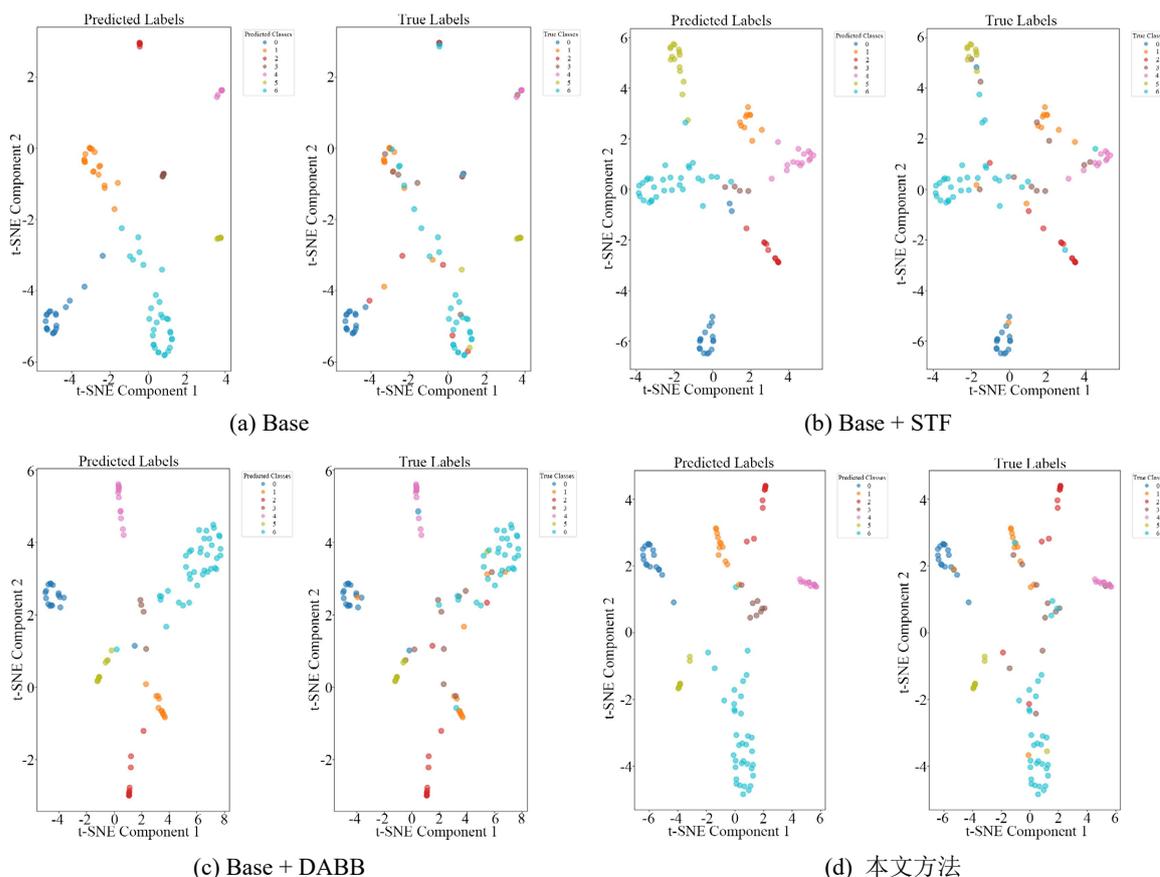


Figure 6. t-SNE visualization results of feature distribution for different models
图 6. 不同模型 t-SNE 特征分布可视化结果

为直观展示不同消融模型在特征空间中的分布情况，采用 t-SNE 对高维特征进行降维可视化，结果

如图 6 所示, 相关聚类指标如表 4 所示。

基线模型在特征空间中类别重叠较为明显, 类间分离度较低。引入 STF 模块后, 类间分离度有所提升, 但类内紧致性仍然不足。引入 DABB 模块后, 特征分布更加紧凑, 类内一致性得到改善。

本文方法在轮廓系数、簇间/簇内距离比和戴维斯 - 鲍丁指数等指标上均取得最佳结果, t-SNE 可视化结果显示其类别边界更加清晰, 验证了 STF 与 DABB 模块在提升特征判别性方面的互补优势。

Table 4. Clustering performance metrics comparison of different models in the truth label space

表 4. 不同模型在真实标签空间的聚类性能指标对比

指标	Base	Base + STF	Base + DABB	本文方法
轮廓系数(\uparrow)	0.3944 \pm 0.025	0.3807 \pm 0.026	0.4123 \pm 0.023	0.4260 \pm 0.018
簇间/簇内距离比(\uparrow)	29.5266 \pm 3.85	43.4693 \pm 3.22	41.2704 \pm 3.48	44.1949 \pm 2.95
戴维斯 - 鲍丁指数(\downarrow)	1.2593 \pm 0.085	1.2903 \pm 0.092	1.3295 \pm 0.098	1.1671 \pm 0.072

3.6. 病灶区域热力图分析

本研究利用梯度加权类激活映射技术生成了可视化热力图, 以展示模型在三类关键 MRI 序列上的聚焦区域, 增强模型的可解释性。如图 7 所示, 热力图按行排列, 自上而下分别对应动态增强 MRI (Dynamic Contrast-Enhanced Magnetic Resonance Imaging, DCE-MRI)、弥散加权成像与 T2 加权成像的融合序列, 以及 T1 加权成像的内外相位序列, 从左至右展示了病灶区域的连续层面切片。

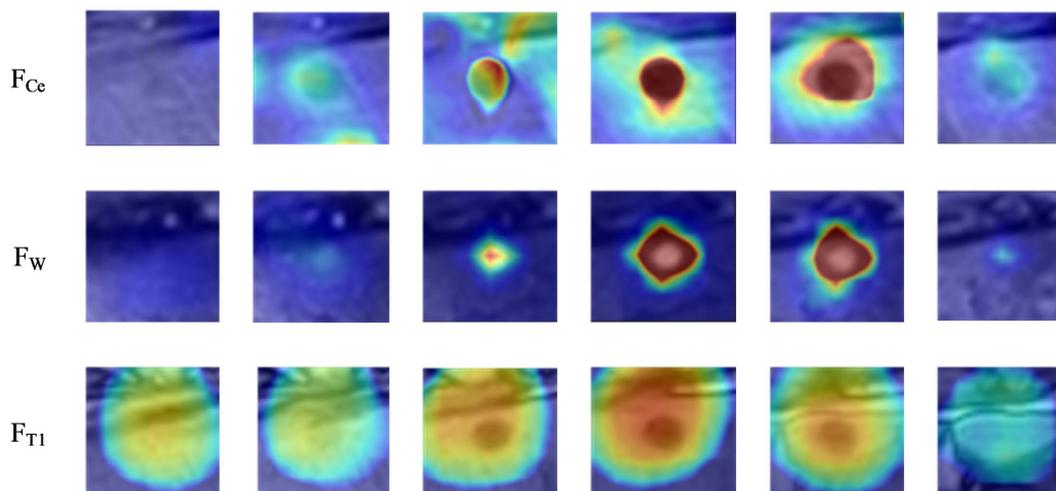


Figure 7. Heatmaps corresponding to 3D MRI of different streams in the spatial branch

图 7. 不同流的 3D MRI 在空间分支中对应的热力图

在 DCE-MRI 中, 高响应区域与病灶位置高度一致, 部分切片呈现边缘强化并随深度逐渐向内部扩展, 表现出由环状至不均匀的强化模式。这说明模型能够捕捉对比剂进入与洗脱过程中的血供差异, 关注肿瘤血流动力学特征及血管生成的空间异质性。

在 DWI 与 T2WI 融合序列中, 热力图高响应主要集中于病灶核心区, 与弥散受限区域高度重叠, 表明模型对细胞密度较高的部位尤为敏感。同时, 部分切片的高响应区也覆盖了 T2WI 的高信号带, 提示模型在判别时兼顾了瘤周水肿及结构改变, 从而更全面地刻画病灶的病理环境。

在 T1WI 正反相位序列中, 热力图响应集中在病灶边界和囊性低信号区, 并未出现明显的信号差异

或局灶性高信号。这表明在该病例中, 模型主要利用解剖边界与液体特征来进行判断, 而相位差特征并未发挥显著作用。这种表现与肝囊肿的影像学性质相符, 说明模型能够根据不同病灶类型灵活调整关注重点。

总体来看, 三类序列的热力图结果显示, 模型在判别过程中能够有针对性地聚焦于不同模态中的关键信息: DCE-MRI 突出血流动力学差异, DWI 与 T2WI 融合反映细胞密度与水肿情况, T1WI 内外相位强调解剖与液体特征。这种多模态互补聚焦与临床判读规律一致, 提升了模型的可解释性。

4. 结束语

本文针对多序列 MRI 肝脏病灶分类中存在的特征异质性和跨序列冗余问题, 提出了一种融合时空特征与动态加权策略的 DSTF-Former 模型。该模型通过 STF 模块实现空间与时间特征的有效对齐与融合, 并结合 DABB 模块在瓶颈桥结构中建立序列间的交互与动态权重分配, 从而充分挖掘多序列特征的互补性。实验结果表明, 该方法在多项评价指标上均优于现有先进模型, 能够更好地区分影像学表现相似的肝脏病灶类别。

尽管如此, 模型在区分影像学特征高度相似的病灶时仍存在一定局限, 部分类别之间的混淆尚未完全消除。未来研究可考虑将模型提取的特征与手工提取的影像学特征相结合, 充分利用临床知识与先验信息, 以进一步提升模型在复杂病例中的鲁棒性和判别力。

参考文献

- [1] Ferlay, J., Colombet, M., Soerjomataram, I., Parkin, D.M., Piñeros, M., Znaor, A., *et al.* (2021) Cancer Statistics for the Year 2020: An Overview. *International Journal of Cancer*, **149**, 778-789. <https://doi.org/10.1002/ijc.33588>
- [2] Lee, S., Kim, Y., Shin, J., Hwang, S.H., Roh, Y.H., Chung, Y.E., *et al.* (2020) CT and MRI Liver Imaging Reporting and Data System Version 2018 for Hepatocellular Carcinoma: A Systematic Review with Meta-Analysis. *Journal of the American College of Radiology*, **17**, 1199-1206. <https://doi.org/10.1016/j.jacr.2020.06.005>
- [3] 赵燕凤, 欧阳汉. 肝脏肿瘤的 MR 检查与诊断思路[J]. 磁共振成像, 2012, 3(6): 456-464.
- [4] Wojciechowska, M., Irving, B., Dennis, A., Wilman, H.R., Banerjee, R., Brady, S.M., *et al.* (2018) Automated Detection of Cystic Lesions in Quantitative T1 Liver Images. In: Nixon, M., Mahmoodi, S. and Zwiggelaar, R., Eds., *Medical Image Understanding and Analysis*, Springer, 51-56. https://doi.org/10.1007/978-3-319-95921-4_6
- [5] Chlebus, G., Meine, H., Abolmaali, N., *et al.* (2018) Automatic Liver and Tumor Segmentation in Late-Phase MRI Using Fully Convolutional Neural Networks. *Proceedings of the Annual Meeting of the German Society of Computer- and Robot-Assisted Surgery*, Leipzig, September 2018, 195-200.
- [6] Trivizakis, E., Manikis, G.C., Nikiforaki, K., Drevelegas, K., Constantinides, M., Drevelegas, A., *et al.* (2019) Extending 2-D Convolutional Neural Networks to 3-D for Advancing Deep Learning Cancer Classification with Application to MRI Liver Tumor Differentiation. *IEEE Journal of Biomedical and Health Informatics*, **23**, 923-930. <https://doi.org/10.1109/jbhi.2018.2886276>
- [7] Mo, S., Cai, M., Lin, L., Tong, R., Chen, Q., Wang, F., *et al.* (2020) Multimodal Priors Guided Segmentation of Liver Lesions in MRI Using Mutual Information Based Graph Co-Attention Networks. In: Martel, A.L., *et al.*, Eds., *Medical Image Computing and Computer Assisted Intervention—MICCAI 2020*, Springer, 429-438. https://doi.org/10.1007/978-3-030-59719-1_42
- [8] Xiao, X., Hu, Q.V. and Wang, G. (2023) Edge-Aware Multi-Task Network for Integrating Quantification Segmentation and Uncertainty Prediction of Liver Tumor on Multi-Modality Non-Contrast MRI. In: Greenspan, H., *et al.*, Eds., *Medical Image Computing and Computer Assisted Intervention—MICCAI 2023*, Springer, 652-661. https://doi.org/10.1007/978-3-031-43901-8_62
- [9] 胡瑶, 王远军. 一种双输入稀疏连接的肝细胞癌和肝内胆管细胞癌分类网络[J]. 小型微型计算机系统, 2024, 45(5): 1129-1135.
- [10] Qiu, Z., Pan, Y., Wei, J., Wu, D., Xia, Y. and Shen, D. (2021) Predicting Symptoms from Multiphase MRI via Multi-Instance Attention Learning for Hepatocellular Carcinoma Grading. In: de Bruijne, M., *et al.*, Eds., *Medical Image Computing and Computer Assisted Intervention—MICCAI 2021*, Springer, 439-448. https://doi.org/10.1007/978-3-030-87240-3_42
- [11] Kumar, V., Li, Y. and Zhang, S. (2025) Uncertainty-Guided and Cross-Modality Attention Network for Liver Tumor

- Segmentation and Classification. *Expert Systems with Applications*, **212**, Article ID: 120069.
- [12] Phan, A., Cao, H., Le, T., Trieu, T. and Phan, T. (2023) Improving Liver Lesions Classification on CT/MRI Images Based on Hounsfield Units Attenuation and Deep Learning. *Gene Expression Patterns*, **47**, Article ID: 119289. <https://doi.org/10.1016/j.gep.2022.119289>
- [13] Wang, S., Han, X., Du, J., Wang, Z., Yuan, C., Chen, Y., et al. (2021) Saliency-Based 3D Convolutional Neural Network for Categorising Common Focal Liver Lesions on Multisequence MRI. *Insights into Imaging*, **12**, Article No. 173. <https://doi.org/10.1186/s13244-021-01117-z>
- [14] Lou, M., Ying, H., Liu, X., Zhou, H., Zhang, Y. and Yu, Y. (2025) SDR-Former: A Siamese Dual-Resolution Transformer for Liver Lesion Classification Using 3D Multi-Phase Imaging. *Neural Networks*, **185**, Article ID: 107228. <https://doi.org/10.1016/j.neunet.2025.107228>
- [15] Wang, R., Shi, X., Pang, S., Chen, Y., Zhu, X., Wang, W., et al. (2025) Cross-Attention Guided Loss-Based Deep Dual-Branch Fusion Network for Liver Tumor Classification. *Information Fusion*, **114**, Article ID: 102713. <https://doi.org/10.1016/j.inffus.2024.102713>
- [16] Huang, S., Hong, Z., Wu, B., Liang, J. and Huang, Q. (2025) Spatio-Temporal Collaborative Multiple-Stream Transformer Network for Liver Lesion Classification on Multiple-Sequence Magnetic Resonance Imaging. *Engineering Applications of Artificial Intelligence*, **142**, Article ID: 109933. <https://doi.org/10.1016/j.engappai.2024.109933>
- [17] Hu, M., Wang, S., Wu, M., et al. (2024) Automatic Classification of Focal Liver Lesions Based on Multi-Sequence MRI. *Journal of Imaging Informatics in Medicine*, **38**, 1986-1998.
- [18] He, K., Zhang, X., Ren, S. and Sun, J. (2016) Deep Residual Learning for Image Recognition. 2016 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, 27-30 June 2016, 770-778. <https://doi.org/10.1109/cvpr.2016.90>
- [19] Huang, G., Liu, Z., Van Der Maaten, L. and Weinberger, K.Q. (2017) Densely Connected Convolutional Networks. 2017 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, 21-26 July 2017, 2261-2269. <https://doi.org/10.1109/cvpr.2017.243>
- [20] Srinivas, A., Lin, T., Parmar, N., Shlens, J., Abbeel, P. and Vaswani, A. (2021) Bottleneck Transformers for Visual Recognition. 2021 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Nashville, 20-25 June 2021, 16514-16524. <https://doi.org/10.1109/cvpr46437.2021.01625>
- [21] Li, K., Wang, Y., Gao, P., et al. (2022) UniFormer: Unified Transformer for Efficient Spatiotemporal Representation Learning. arXiv: 2201.04676.
- [22] He, A., Wang, K., Li, T., Du, C., Xia, S. and Fu, H. (2023) H2Former: An Efficient Hierarchical Hybrid Transformer for Medical Image Segmentation. *IEEE Transactions on Medical Imaging*, **42**, 2763-2775. <https://doi.org/10.1109/tmi.2023.3264513>
- [23] Xu, X., Wang, C., Guo, J., Gan, Y., Wang, J., Bai, H., et al. (2020) MSCS-DeepLN: Evaluating Lung Nodule Malignancy Using Multi-Scale Cost-Sensitive Neural Networks. *Medical Image Analysis*, **65**, Article ID: 101772. <https://doi.org/10.1016/j.media.2020.101772>
- [24] Xing, Z., Ye, T., Yang, Y., Liu, G. and Zhu, L. (2024) SegMamba: Long-Range Sequential Modeling Mamba for 3D Medical Image Segmentation. In: Linguraru, M.G., et al., Eds., *Medical Image Computing and Computer Assisted Intervention—MICCAI 2024*, Springer, 578-588. https://doi.org/10.1007/978-3-031-72111-3_54