

中图法分类号: TP399 文献标识码: A 文章编号: 1006-8961(2025)04-1170-13

论文引用格式: Wei R, Qi X M, He Y T, Jiang S, Qian W, Xu Y, Zhu Y S, Pascal H, Shu H Z and Yang G Y. 2025. Knowledge distillation and mutual information of multimodal MRI for disease prognosis. *Journal of Image and Graphics*, 30(4):1170-1182(魏然, 戚晓明, 何宇霆, 江升, 钱雯, 徐怡, 祝因苏, Pascal Haigrón, 舒华忠, 杨冠羽. 2025. 结合知识蒸馏与互信息的多模态MRI疾病预后. *中国图象图形学报*, 30(4):1170-1182)[DOI:10.11834/jig.240349]

结合知识蒸馏与互信息的多模态MRI疾病预后

魏然¹, 戚晓明¹, 何宇霆¹, 江升¹, 钱雯², 徐怡², 祝因苏³,
Pascal Haigrón⁴, 舒华忠¹, 杨冠羽^{1*}

1. 新一代人工智能技术与交叉应用教育部重点实验室(东南大学), 南京 210096;
2. 南京医科大学第一附属医院放射科, 南京 210029; 3. 江苏省肿瘤医院影像中心CT室, 南京 210009;
4. 法国雷恩大学信号与图像处理实验室, 雷恩 F-35000, 法国

摘要: 目的 多模态心脏磁共振(cardiac magnetic resonance, CMR)图像预测非缺血性扩张型心肌病(non-ischemic dilated cardio myopathy, NIDCM)的预后在对临床中心力衰竭或心源性猝死等不同应用中发挥着重要作用。针对多模态CMR图像预后表征困难和标注困难两个挑战,提出了一种基于混合匹配蒸馏与对比互信息估计的模型,用于小样本上的多模态CMR图像对NIDCM预后。方法 本文的预后模型有两种不同的设计,解决深度学习网络中多模态CMR图像的表征困难和模型容易陷入局部最优的问题。首先将不同模态CMR图像组合为不同的模态对,并提取对应的图像特征。由于不同模态对的预后目标一致而图像特征分布之间存在差异,因此设计一种混合匹配蒸馏网络,利用逻辑分布一致性将不同图像特征分布关联匹配,以此约束深度学习网络中多模态特征的提取和引导联合表征。然后在不同模态对之间设计一种互信息的对比学习策略,从而估计出多模态分布上的潜在的分类边界,以此作为预后模型的正则化项,避免模型在有限的数据上陷入局部最优。结果 实验在一个NIDCM临床数据集上分别与6种方法进行了比较。F1值和准确率值达到81.25%和85.61%;为了验证模型的泛化性,在一个脑肿瘤公共数据集上也分别与4种方法进行了比较,F1值和准确率值达到85.07%和87.72%。结论 本文所提出的基于混合匹配蒸馏与对比互信息估计的预后网络模型对多模态CMR图像进行了有效表征,同时利用多模态之间的潜在互信息增强深度学习模型在小样本场景下的模型优化,最终使得多模态CMR图像对的NIDCM预后结果更加准确。

关键词: 对比学习;混合蒸馏;多模态心脏磁共振(CMR)图像;互信息估计;预后

Knowledge distillation and mutual information of multimodal MRI for disease prognosis

Wei Ran¹, Qi Xiaoming¹, He Yuting¹, Jiang Sheng¹, Qian Wen², Xu Yi², Zhu Yinsu³,
Pascal Haigrón⁴, Shu Huazhong¹, Yang Guanyu^{1*}

1. The Key Laboratory of New Generation Artificial Intelligence Technology and Its Interdisciplinary Applications (Southeast University), Ministry of Education, Nanjing 210096, China; 2. Department of Radiology, the First Affiliated Hospital of Nanjing Medical University, Nanjing 210029, China; 3. Department of Imaging CT Room, Jiangsu Cancer Hospital, Nanjing 210009, China;
4. The University Rennes, Inserm, LTSI-UMR1099, Rennes F-35000, France

收稿日期:2024-06-27;修回日期:2024-10-11;预印本日期:2024-10-18

* 通信作者:杨冠羽 yang.list@seu.edu.cn

基金项目:国家重点研发计划政府间合作项目(2022YFE0116700)

Supported by: Intergovernmental Cooperation Project of National Key R&D Program of China (2022YFE0116700)

Abstract: Objective Nonischemic dilated cardio myopathy (NIDCM) is a heart condition that can lead to severe outcomes, such as heart failure or sudden cardiac death. Accurate prognosis plays a crucial role in the early diagnosis and effective treatment of this disease. Multimodal cardiac magnetic resonance (CMR) imaging captures heart data from different perspectives and is essential for prognosis. Each CMR modality provides unique and complementary information about the heart's structure and function. However, leveraging these multimodal data sources to predict prognosis poses two main challenges. First, the regions of interest (RoIs) differ among CMR modalities even when imaging the same disease. Hence, combining these modalities into a cohesive predictive model becomes challenging. The distinct characteristics and distribution of data from different modalities also create difficulties in capturing comprehensive information about the prognosis of NIDCM. Second, the limited labeled training data exacerbate the problem. Owing to the difficulty in labeling such data, the available dataset is small. Hence, the risk of a deep learning model falling into local optima increases, hindering its ability to generalize and achieve good predictive performance. Therefore, a specialized approach that can address the complexity of multimodal data representations and the limitations of small sample sizes is warranted. **Method** To overcome these challenges, we propose a novel model based on hybrid matching distillation and contrastive mutual information estimation. Its design focuses on two aspects: improving the representation of multimodal CMR images and preventing the model from falling into local optima due to the limited training data. The first component of our method involves combining different CMR modalities into pairs. Each pair is treated as a unique data source, and the image features corresponding to these modality pairs are extracted. Given that the prognosis objective is consistent across modalities but their feature distributions vary, a hybrid matching distillation network that enforces logical distribution consistency between the modalities is employed. It learns to associate and match different image feature distributions across modalities by leveraging the inherent consistency in prognosis objectives. This matching constrains the extraction of features from each modality, ensuring that the deep learning network can jointly represent multimodal features. As a result, the network effectively captures the complementary information from various modalities, leading to a good predictive performance. The second component is a mutual information contrastive learning strategy applied to estimate potential classification boundaries across the multimodal feature distribution. This step introduces a regularization term into the prognosis model to prevent it from falling into a local optimum during training with a small sample size. The contrastive learning strategy aims to maximize the mutual information between modalities while learning meaningful feature representations. By estimating the classification boundaries, the model can discern the subtle differences in the feature space and consequently enhance its ability to generalize from limited data. This strategy regularizes the learning process and ensures that the model captures the most informative aspects of the multimodal data. The two components of hybrid matching distillation and contrastive mutual information estimation work together to build a robust prognosis model for NIDCM. By utilizing the logical consistency across modalities and the mutual information between them, the model achieves improved feature representation and avoids overfitting to the small sample size. **Result** Experiments were conducted using a clinical dataset of patients with NIDCM to evaluate the performance of the proposed model. The results were compared with those of six state-of-the-art methods to assess the model's effectiveness. The performance was evaluated using two key metrics: F1 score and accuracy (Acc). F1 score is particularly useful in assessing the balance between precision and recall, and Acc is important in assessing the overall correctness of predictions. The proposed model achieved an F1 score of 81.25% and an accuracy of 85.61% on the NIDCM dataset. These results demonstrated significant improvement over the baseline models, highlighting the effectiveness of the hybrid matching distillation and contrastive mutual information estimation techniques in handling multimodal CMR images for prognosis prediction. These two complementary approaches allow the model to comprehensively utilize the limited training data and capture the complex correlations between the different modalities. An additional experiment was conducted on a public dataset related to brain tumors to further validate the generalization capability of the model. This dataset also featured multimodal medical images and allowed us to verify whether the proposed method could be applied beyond the NIDCM domain. The model achieved an F1 score of 85.07% and an accuracy of 87.72% on this dataset, outperforming the four baseline methods once again. These results demonstrated that the proposed model can be applied to other medical imaging tasks beyond NIDCM, making it a versatile and effective tool for prognosis prediction. **Conclusion** A prognosis network model based on hybrid matching distillation and contrastive mutual information estimation was proposed to effectively

addresses the two major challenges in using multimodal CMR images for NIDCM prognosis. Hybrid matching distillation ensures that the model learns to represent multimodal data by leveraging the logical consistency across different CMR modalities, thereby improving the model's ability to capture the complementarity between modalities. Contrastive mutual information estimation provides regularization by estimating classification boundaries, preventing the model from overfitting to small sample sizes. Experiments demonstrated that this approach significantly improves the accuracy and F1 score of the prognosis model, outperforming several state-of-the-art methods. The model's generalization capabilities were further confirmed through its successful application to a brain tumor dataset, proving its versatility across various medical imaging tasks. In addition to its high predictive performance, this method demonstrates the potential of deep learning in handling complex medical imaging problems where data scarcity is a concern. By combining hybrid matching distillation with contrastive mutual information estimation, this model can handle the intricate relationships between different modalities and produce robust, reliable predictions even with limited labeled data. Future work should explore the extension of this method to other forms of multimodal medical imaging beyond CMR, such as combining magnetic resonance imaging (MRI) and computed tomography (CT) scans for comprehensive diagnosis. Further improving the efficiency of training and reducing the computational complexity of the model would make it accessible for widespread clinical use. This research opens up new avenues for multimodal prognosis models and sets a foundation for future innovations in medical image analysis.

Key words: contrastive learning; hybrid distillation; multi-modal cardiac magnetic resonance (CMR); mutual estimation; prognosis

0 引言

基于多模态心血管磁共振成像 (cardiac magnetic resonance, CMR) 图像预后研究具有重要意义。多模态 CMR 图像的预后在对临床心力衰竭、心源性猝死和脑肿瘤分类等不同应用中发挥着重要作用。非缺血性扩张型心肌病 (non-ischemic dilated cardio myopathy, NIDCM) 是最常见的心肌病形式 (Arenja 等, 2017), 其 10 年心血管死亡率高达 40% (Maron 等, 2006)。非缺血性扩张型心肌病 (NIDCM) 的预后对于患者出现晚期心力衰竭或心源性猝死的情况至关重要。心脏磁共振成像 (CMR) 被广泛认为是评估非缺血性扩张型心肌病患者的一种无创模式选择 (Quarta 等, 2011)。具体来说, CMR 的 T1 映射 (T1-native (T1 原生), T1-enhanced (T1 增强)) 和 LGE (延迟钆增强) 序列可以详细描述 NIDCM 的心脏特征 (Nwabuo 和 Vasan, 2020)。T1 映射和 LGE 序列在 CMR 中的结合应用为了解 NIDCM 的心脏特征提供了全面的图像学信息, 从而有助于更准确地进行预后和治疗规划。

已有大量工作通过深度学习模型对多模态 CMR 进行分析, 从而辅助临床诊断。随着深度学习在医学领域的逐步发展与应用, 其在医学影像采集、疾病分析诊断、术前规划和术中指导等方面提供准

确高效的辅助。然而, 现有的深度学习模型在多模态磁共振图像预后方面存在局限。依据多模态融合过程是否是可学习的, 可以将已有工作分为两类: 1) 固定的融合方式。由于医学图像多模态数据样本少, 导致多模态融合模型往往容易陷入局部最优, 此类方法将依据临床经验定义好多模态信息的融合方式, 指导模型融合。Li 等人 (2020) 通过观察生物标记的变化探讨肺动脉扩张与结缔组织病相关肺动脉高压患者预后之间的关联。Uthoff 等人 (2020) 不仅利用心脏, 还利用其周围组织的生物标记来表征疾病严重程度, 从而提高预后精度。然而, 这些生物标记的提取取决于临床医生的经验和偏好, 这可能会由于不同操作者之间的差异而引入额外的误差。Sun 等人 (2019) 提出了一种通过整合多维数据的多模态深度神经网络来预测乳腺癌的预后。这些方法度量不同序列的相似性, 增强预测中的相关性。Havaei 等人 (2016) 将输入图像嵌入到一个有良好定义的算术运算的单个潜在向量空间中进行平均, 然后进一步处理以产生所需的分割。Qiu 等人 (2023) 提出了一个模块来规范心肌一致性和定位病理, 然后引入包容性损失来利用心肌疤痕和水肿之间的关系引导多模态融合。但此类方法难以解决在临床中无法完全清晰定义的疾病诊断, 如 NIDCM 等。2) 动态的融合方式。这些方法将不同模态的特征拼接成一个整体, 依靠任务标签驱动网络自动聚合特

征。Baltrušaitis 等人(2019)训练神经网络提取与任务相关的特征,从而进行预后。余力等人(2022)提出一种带有变分自编码器分支的两阶段级联 U-Net 算法,旨在分割肿瘤的同时提取鲁棒的特征预测患者生存期。Zhou 等人(2020)提出了一种基于注意力机制的多模态融合网络,用于脑肿瘤分割。Li 等人(2022)提出了一种 VGG (Visual Geometry Group) 网络的 3D 卷积神经网络模型,通过 DiffGrad 优化器来对模型进行加速并通过 Focalloss 函数提高模型对多个预测目标的精度。丁熠等人(2023)提出多层次并行的多模态脑肿瘤分割框架,可以有效提高脑肿瘤分割的精度并缩短分割时间。Chen 等人(2019)提出了一种基于注意力的多模态数据乳腺癌预后模型用于乳腺癌的预后。Qi 等人(2022)提出了一种多视角空间感知对抗网络,研究对细粒度心脏结构的感知,以便通过空间感知多视图 CMR 融合进行准确分割。Xiao 等人(2022)利用模态感知 Transformer 进行多模态 MRI (magnetic resonance imaging) 特征融合与选择,通过捕获多模态 MRI 之间的相关

性,解决了多模态信息多样性的挑战。Shankar 等人(2024)引入了向后连接,使后期融合表示可用于早期层,提高了这些阶段表示的表达性,同时保留了后期融合设计的优势。由于数据量的样本量缺少,导致模型对于不同模态间相互依赖关系和信息流缺乏特定的优化目标,使得聚合是隐式的和不稳定的。

综上所述,基于深度学习的多模态 CMR 图像 NIDCM 预后存在以下挑战:1)表征困难。多模态 CMR 图像预后与各个模态中不同 RoI (regions of interest) 的各种功能评估有关(图 1(a))。序列间的互补性和相关性不明确,使得联合协调表征 (Baltrušaitis 等, 2019) 在多模态 CMR 的预后表征能力较弱。2)易陷入到局部最优。由于心力衰竭、心源性猝死等的精确标记需要预后检查,因此临床扫描的多序列 CMR 样本量较小,导致模型优化的可训练数据有限。因此,从小样本数据中确定的分类边界与真实的分类边界相差过大,导致训练的模型易陷入局部最优(图 1(b))。

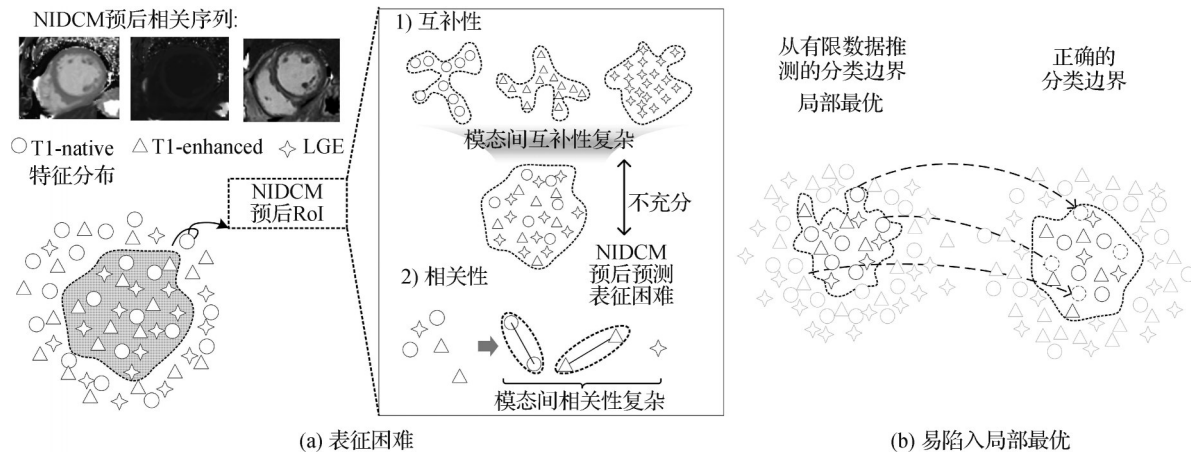


图1 多模态CMR预后NIDCM的挑战

Fig. 1 The challenge of predicting the prognosis of NIDCM using multi modal CMR
((a) difficulty in representation; (b) easy to trap in local optimum)

为了克服以上挑战,本文提出了一种新的多序列表征的对比学习框架——基于配对池的混合匹配蒸馏和对比互估计网络(HCPNet),促进了多模态CMR对NIDCM的预后。1)HCPNet提出了一种混合匹配蒸馏(hybrid matching distillation, HMD)模块,从配对池的不同配对组合中挖掘互补和相似匹配,从而促进多序列CMR表征(图2(a))实现多序列CMR对NIDCM预后的有效表征。通过挖掘不同配

对组合中的互补匹配,HMD能够整合来自多个序列的CMR图像的独特信息。不同序列的CMR图像往往提供了关于心脏结构和功能的不同视角,互补匹配可以将这些不同视角的信息融合起来,为NIDCM预后提供更全面的特征表示。其次,相似匹配的挖掘有助于强化多序列CMR图像之间的共同特征。这些共同特征可能与NIDCM的特定病理生理过程相关,通过HMD模块的相似匹配,可以更加

突出这些关键特征,提高预后的准确性。2)HCPNet提出了一种对比互估计(contrastive mutual estimation, CME)策略以保证模型的优化,使模型不易陷入局部最优。通过比较正负样本对的相对距离并结合互信息,模型能够更加全面地学习到不同样本之间的差异和共性。这种对比促使模型去探索更丰富的特征空间,而不是仅仅局限于局部的特征模式。多样化的特征表示使得模型在优化过程中有更多的方向可以探索,从而降低了陷入局部最优解的可能性。同时增加互信息对比后,模型的优化目标不仅仅是最小化特定的损失值,还包括最大化正负样本

对之间的互信息(图2(b))。

本文提出的 HCPNet 的贡献如下:1)所提出的 HCPNet 模型是第 1 个基于小样本的多模态 CMR 对 NIDCM 进行预后的深度学习框架。HCPNet 在诊断晚期心力衰竭、心源性猝死等临床场景上具有潜在的诊断辅助作用。2)提出了一种新的多模态 CMR 表征方式,以保证了不同序列 RoI 的表征在不同的两两匹配对中的互补性和有相似性。3)提出了对比互估计策略(CME)引导多序列小样本的优化避免陷入局部最优。CME 估计出多模态分布上的潜在的分类边界,并以此作为预后模型的正则化项。

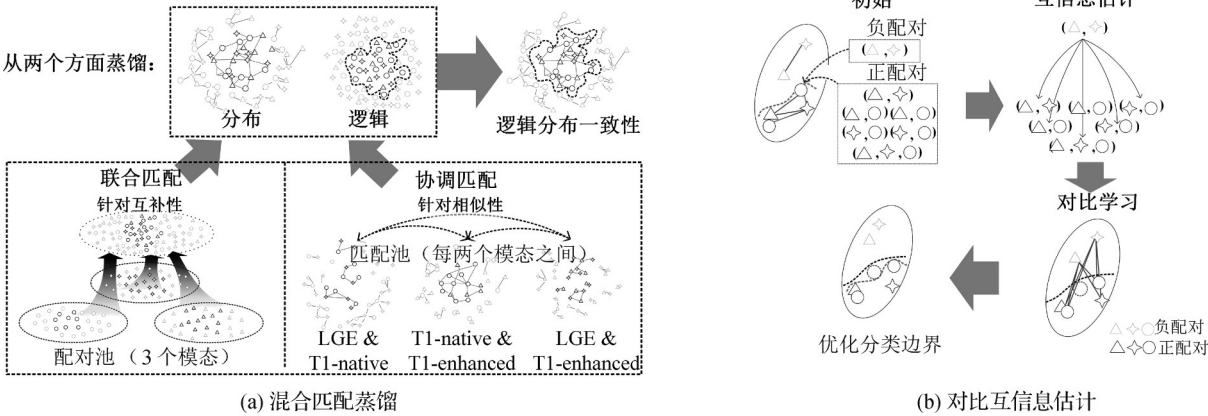


图2 多模态CMR预后NIDCM的方法
Fig. 2 Method for predicting the prognosis of NIDCM using multi modal CMR
(a) hybrid matching distillation; (b) contrastive mutual estimation)

1 基于混合匹配蒸馏与对比互信息估计的预后网络模型

HCPNet提出了一种在小样本数据上实现多模态CMR的高效表征进行预后的方法(图3),它包括两个部分:1)混合匹配蒸馏(HMD)模块高效挖掘利用多模态CMR之间的互补性和相似性(第1.1节)。通过挖掘不同配对组合中的互补匹配,HMD能够整合来自多个序列的CMR图像的独特信息。2)对比互估计(CME)模块通过对输入序列间各种组合配对进行对比和互信息估计纠正分类边界,从而精细化模型的优化(第1.2节)。通过比较正负样本对的相对距离并结合互信息,模型能够更加全面地学习到不同样本之间的差异和共性。这种对比促使模型去探索更丰富的特征空间,而不是仅仅局限于局部的

特征模式。多样化的特征表示使得模型在优化过程中有更多的方向可以探索,从而降低了陷入局部最优解的可能性。

1.1 配对池中多模态高效表征的HMD

HMD通过两个步骤实现了多序列CMR(T1-native, T1-enhanced, LGE)的有效表征:1)配对池匹配;2)混合蒸馏。HMD在不同的序列组合中构造成对池,再分别以特征分布引导相似性提取和预测logit引导互补性提取两种方式进行蒸馏,引导多模态的混合表征。

1.1.1 配对池匹配

考虑到不同序列的组合序列之间存在相似性和互补性(LGE-T1 native, T1 native-T1 enhanced, LGE-T1 enhanced, LGE-T1 native-T1 enhanced),将不同序列的CMR(T1-native, T1-enhanced, LGE)输入到每个特征提取器(E_{lge} 、 E_{tna} 和 E_{ten})。

HMD从两种角度来构造匹配对池:

1)从分布相似性角度。通过交叉注意力机制从表征(R_{lge} 、 R_{tna} 、 R_{ten})中捕获相似性。以T1-native序列和T1-enhanced序列为例,交叉注意力机制取 R_{tna} 和 R_{ten} 作为输入。 R_{tna} 和 R_{ten} (大小相同(b, c, d, w, h))串联为 K 和 V 的输入。 Q 、 K 、 V 将分布转换为 $1 \times (c \times d \times w \times h)$ 和 $(c \times d \times w \times h) \times 1$ 。这两个模态前馈融合过程表示为

$$C_{ten-tna} = f_{Concat}(R_{tna}, R_{ten}) + f_{softmax}\left(\frac{QK^T}{\sqrt{d_k}}V\right) \quad (1)$$

通过类似的方法获得相似性 $C_{lge-tna}$ 和 $C_{ten-lge}$ 。

2)从logits互补性角度。HMD引入了多模态CMR对NIDCM的预后来约束互补性的提取。除此之外,不同序列对的概率($L_{lge-tna}$ 、 $L_{ten-tna}$ 和 $L_{ten-lge}$)被视为联合推理的软logits标签,从而引导不同模态中不同信息的提取,优化联合表征。

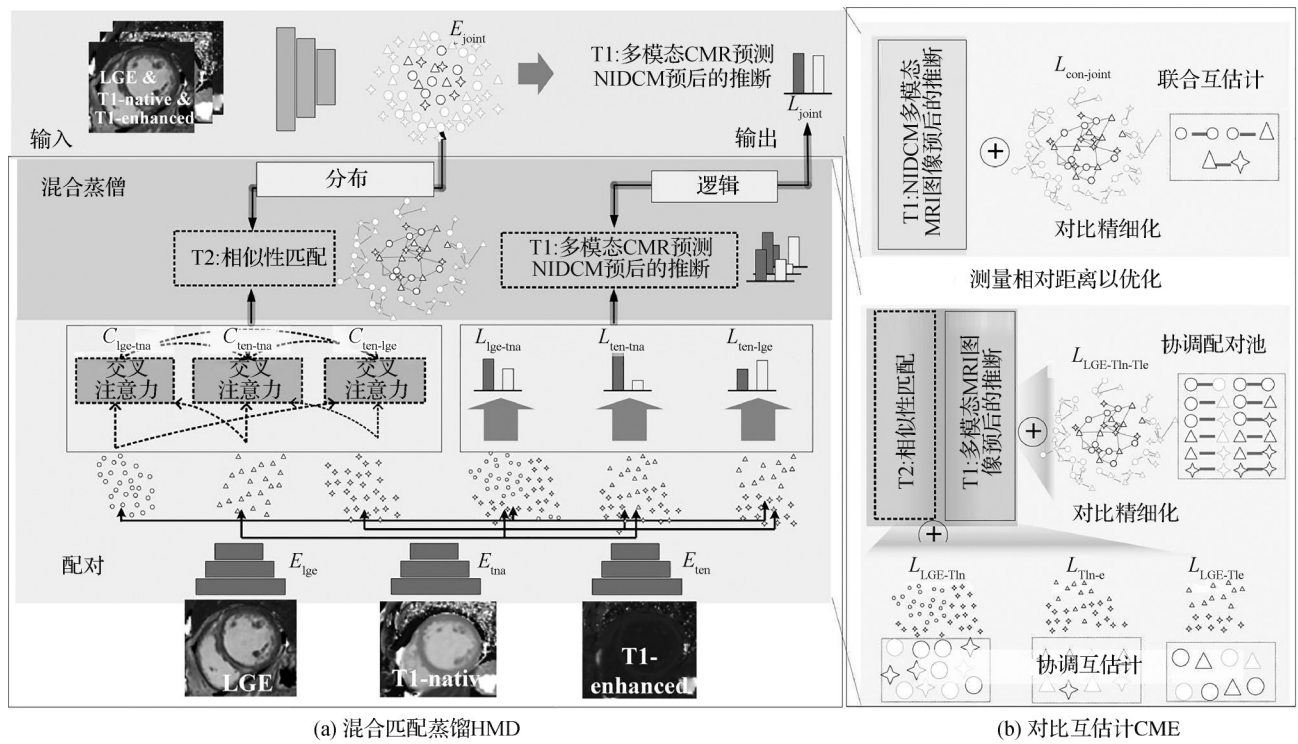


图3 HCPNet的框架

Fig. 3 The framework of HCPNet ((a) hybrid matching distillation; (b) contrastive mutual estimation)

1.1.2 混合蒸馏

将配对池匹配的分布相似性和logits互补性蒸馏提取到联合表征(R_{joint} , L_{joint})中:

1)对于分布相似性,HMD通过引入以Kullback-Leibler(KL)散度的形式测量互信息来最小化 $C_{ten-tna}$, $C_{lge-tna}$, $C_{ten-lge}$ 与 E_{joint} 的差异,表示为

$$L_{dis} = \frac{1}{12} (KL(q||p) + KL(p||q)) \quad (2)$$

式中, q 和 p 是 $C_{ten-tna}$, $C_{lge-tna}$, $C_{ten-lge}$ 与 E_{joint} 的不同组合。通过最大化互信息,将序列对在分布上的相似性嵌入到联合表示 E_{joint} 中。分布蒸馏利用最小化差异任务从分布中提取相关性并优化编码器(E_{lge} 、 E_{tna} 和 E_{ten})。

2)对于logits互补性,使用软logits($L_{lge-tna}$ 、 $L_{ten-tna}$

和 $L_{ten-lge}$)通过下式来约束 L_{joint} ,具体为

$$L_{log} = CE(L_{joint}, L_{lge-tna}) + CE(L_{joint}, L_{ten-tna}) + CE(L_{joint}, L_{ten-lge}) \quad (3)$$

式中, $CE(\cdot)$ 是交叉熵损失(cross entropy loss)。利用 L_{joint} 来约束不同序列对互补性的提取。

1.2 将配对优化细化的CME

CME通过对比不同序列对的相互差异来估计潜在的预测边界,从而修正局部最优边界。CME对不同的配对池中的互补性表征和相似性表征进行相互估计,按类别分为正负对。在相似性表征中,多模态CMR表示通过不同的组合构造负对和正对。

在相似性表征中,将患者的多序列CMR整体作为一个样本 $P_j = \{P_j^1, P_j^2, \dots, P_j^n\}$ 。式中, n 为患者数量, P_j^k 为患者 k 的联合表征。在对比优化过程中,需要扩大负样本与正样本之间的距离、缩小同类别

样本之间的距离。受 Xing 等人 (2021) 的启发, 定义互信息损失为

$$L_{\text{con-joint}} = \sum_{i=1}^{K_p} \left(\log \frac{e^{P_i \times P_{j \neq i}}}{e^{P_i \times P_{j \neq i}} + \frac{k_N}{n}} + \sum_{i=1}^{K_N} \log \left(1 - \frac{e^{P_i \times P_{j \neq i}}}{e^{P_i \times P_{j \neq i}} + \frac{k_N}{n}} \right) \right) \quad (4)$$

式中, k_p 和 k_N 分别代表 n 例患者正样本数和负样本数。 P_i 是 P_j 中一个联合表征。通过最小化 $L_{\text{con-joint}}$, 优化了联合表征, 使其与正样本相似, 而与负样本不同。

在互补性表征方面, 利用 T1-native、T1-enhanced、LGE 不同组合的相关性来建立配对池。有 4 种不同的配对池: T1-native&T1-enhanced、LGE&T1-enhanced、LGE&T1-native 和 LGE&T1-native &T1-enhanced。在构造的配对池上使用相同的策略进行优化, 以 T1-native 和 T1-enhanced 为例, 配对的表征为 $P_{\text{c-T1n-T1e}} = \{P^1_{\text{T1n-T1e}}, \dots, P^z_{\text{T1n-T1e}}\}$, 其中 $z = C^2_{k_N} + C^2_{k_p} + k_p \times k_N$ 。 k_p 和 k_N 表示 n 例患者的正样本数和负样本数。 C_i^2 是组合运算。互信息损失的估计定义为

$$L_{\text{T1n-e}} = \sum_{i=1}^{C^2_{k_p}} \left(\log \frac{e^{P_i \times P_{j \neq i}}}{e^{P_i \times P_{j \neq i}} + \frac{k_N \times k_p}{C^2_{k_p} + k_N \times k_p}} + \sum_{i=1}^{k_N \times k_p} \log \left(1 - \frac{e^{P_i \times P_{j \neq i}}}{e^{P_i \times P_{j \neq i}} + \frac{k_N \times k_p}{C^2_{k_p} + k_N \times k_p}} \right) \right) + \sum_{i=1}^{C^2_{k_N}} \left(\log \frac{e^{P_i \times P_{j \neq i}}}{e^{P_i \times P_{j \neq i}} + \frac{k_N \times k_p}{C^2_{k_N} + k_N \times k_p}} + \sum_{i=1}^{k_N \times k_p} \log \left(1 - \frac{e^{P_i \times P_{j \neq i}}}{e^{P_i \times P_{j \neq i}} + \frac{k_N \times k_p}{C^2_{k_N} + k_N \times k_p}} \right) \right) \quad (5)$$

式中, $C^2_{k_p}$ 和 $C^2_{k_N}$ 是正样本对的数目, $k_N \times k_p$ 是负样本对的数目。 $L_{\text{T1n-e}}$ 的最小化使相似的匹配对接近、不相似的配对远离。 $L_{\text{LGE-T1e}}$ 、 $L_{\text{LGE-T1n}}$ 和 $L_{\text{LGE-T1n-T1e}}$ 的计算方法与 $L_{\text{T1n-e}}$ 类似。最终, HCPNet 的损失函数定义为

$$LOSS = L_{\text{dis}} + L_{\text{log}} + L_{\text{con-joint}} + L_{\text{T1n-e}} + L_{\text{LGE-T1e}} + L_{\text{LGE-T1n}} + L_{\text{LGE-T1n-T1e}} + L_{\text{lab}} \quad (6)$$

式中, L_{lab} 和 L_{joint} 是标签计算的交叉熵损失。

2 实验设置

2.1 数据集

本文的主要任务是预测患有 NIDCM 病人 5 年

周期是否能够生存的预后任务。本文使用了江苏省人民医院从临床扫描中采集的多模态 CMR 作为临床数据集, 来验证 HCPNet 在这一任务上的性能。此外, 还采用具有多序列脑 MRI 的公开可用数据集 BraTS2018 (Menze 等, 2015) 来评估 HCPNet 的泛化性。1) 临床数据集 NIDCM 中包含了 3 个 CMR 序列: T1-native、T1-enhanced 和 LGE。在进行心脏 T1 mapping 扫描中, 在打对比剂前进行一次 T1 mapping 的扫描, 称为 T1-native 图像。T1-enhanced 表示增强后 T1 值, 指心肌组织给予对比剂以后的纵向弛豫时间。LGE 是经静脉注射钆对比剂一段时间后, 正常心肌因对比剂廓清而无强化, 异常心肌因侧支循环形成或对比剂弥散和滞留而出现强化所呈现的图像。数据集共有 135 例 (阳性 97: 阴性 38), 其中阳性代表病人 5 年内可以生存, 阴性代表病人 5 年内无法生存。训练集与测试集的比值为 3:1。为了保证模型的稳健性, 采用了 4 次交叉验证, 并选择最优的一次交叉验证进行测试。2) BraTS2018 数据集来自国际医学图像计算和计算机辅助干预协会 (Medical Image Computing and Computer Assisted Intervention Society, MICCAI) 2018 年举办的多模态脑肿瘤分割挑战赛。所有成像数据均由 1~4 名评分者按照相同的标注方案手动分割。BraTS2018 中, 有 285 例患者有 4 种不同的序列 T1 原生 (T1-native)、T1 加权 (T1-weighted)、T2 加权 (T2-weighted)、T2 液体衰减反转恢复 (T2 fluid attenuated inversion recovery) 用于低级别胶质瘤 (low-grade glioma, LGG) 和高级别胶质瘤 (high-grade glioma, HGG) 的分类。需要分割 3 个部分: WT (whole tumor)、ET (enhance tumor) 和 TC (tumor core)。HGG 有 210 名患者, LGG 有 75 名患者。训练集与测试集的比值为 4:1, 通过 5 折交叉验证以获得最佳测试折。

2.2 实验环境

在 HCPNet 训练中, 网络是基于 Pytorch 实现的, 学习速率为 1×10^{-4} , 训练 500 轮。使用 Adam 优化器对网络进行优化, 临床数据集和 BraTS2018 的 batch size 分别为 4 和 2。HCPNet 编码器是 ResNet50 (residual network-50) (He 等, 2016)。在模型训练中, 随机对每个患者进行数据增强 (旋转、平移、缩放和镜像)。所有的实验都是在单个 NVIDIA TITAN RTX GPU 上进行。

2.3 评价指标

本文主要采用F1、准确率(accuracy)、精确率(precision)和召回率(recall)作为评价指标来评估模型对多模态CMR预后NIDCM的性能。

准确率代表了整体的预测准确程度,包括正样本和负样本;精确率代表对正样本结果中的预测准确程度;召回率又称为查全率,评估所有实际正例是否被预测出来的覆盖率占比;单独用精确率或召回率都不能很好地评估模型好坏,F1值是中和了精确率和召回率的指标。

考虑到实验所涉及的数据由于隐私安全,数据量较少,因此对于不同模态组合预后的对比实验,使用统计方法来计算不同方法F1值的 p 值。通过计算 p 值,可以确定不同实验条件下F1值结果的显著性差异。

3 实验结果与分析

3.1 临床数据集多模态CMR预后NIDCM对比实验

在临床数据集NIDCM上,本文模型HCPNet在预后结果中展现了优异的性能。与目前先进的医学图像多模态方法IMDM(integrate multimodal data method)(Yan等,2019),PMNet(PET-MRI Net)(Li等,2022),MMRACR(multi modality relation attention network that employs consistency regularization)(Yang等,2022),Ensemble ResNet(Zhou等,2020),MyoPS-Net(Qiu等,2023)和PFMI(progressive fusion for mul-

timodal integration)(Shankar等,2024)相比,本文HCPNet在多序列CMR图像预后上具有更好的性能,如表1所示,在F1和准确率这两个评价指标上都取得了最好的结果。

HCPNet通过HMD的有效表征和CME的优化细化获得了最高的F1评分为0.81。相较IMDM、PMNet、MMRACR、Ensemble ResNet、MyoPS-Net和PFMI分别提升14%、7%、5%、2%、1%和2%。本文模型的准确率达到0.86,与以上模型相比,分别提升了9%、6%、8%、3%、0.8%和1%。精确率达到0.85,相较于IMDM、PMNet、Ensemble ResNet、MyoPS-Net和PFMI也有一定提升,分别提升15%、11%、6%、3%和4%。召回率达到0.81,相较于IMDM、PMNet、MMRACR和PFMI也有一定提升,分别提升14%、5%、10%和0.2%。现有的模型表示方法主要集中于多模态CMR的联合或协调表示。然而,多模态CMR的复杂相关性使其难以实现对NIDCM预后的有效表征。在4折交叉验证的每一折中,HCPNet均获得了较高的F1值,这表明所提算法的优越性不仅仅是在某一折的结果上,而是具有一定的普遍性。除此之外, p 值检验小于0.01,表明所提出的方法在NIDCM数据上的结果具有统计显著性。综上,在NIDCM数据集上,HCPNet在多模态CMR预后NIDCM上具有更好的性能。

3.2 BraTS数据集多模态MRI脑肿瘤预后的对比实验

通过在脑肿瘤数据集BraTS上的优越性,展现了本文模型HCPNet的适应性和泛化性,如表2所

表1 不同模型在NIDCM数据集上的预后结果对比
Table 1 Comparison of prognostic results of different models on the NIDCM dataset

方法	准确率				精确率				召回率				F1			
	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
IMDM	0.76	0.74	0.70	0.68	0.69	0.69	0.66	0.61	0.66	0.64	0.64	0.59	0.67*	0.65	0.62	0.62
PMNet	0.79	0.75	0.77	0.77	0.73	0.70	0.73	0.71	0.75	0.71	0.75	0.73	0.74*	0.74	0.71	0.70
MMRACR	0.77	0.77	0.72	0.75	0.91	0.87	0.85	0.81	0.70	0.65	0.68	0.69	0.76+	0.70	0.73	0.75
Ensemble ResNet	0.83	0.79	0.80	0.81	0.78	0.74	0.75	0.76	0.81	0.76	0.79	0.79	0.79*	0.75	0.77	0.78
MyoPS-Net	0.85	0.84	0.80	0.81	0.81	0.77	0.79	0.79	0.82	0.75	0.79	0.80	0.80*	0.76	0.80	0.79
PFMI	0.84	0.84	0.83	0.80	0.80	0.75	0.79	0.74	0.80	0.77	0.75	0.75	0.79*	0.78	0.77	0.78
本文	0.86	0.84	0.85	0.81	0.85	0.82	0.82	0.80	0.81	0.77	0.78	0.80	0.81*	0.79	0.79	0.78

注:加粗字体表示各评价指标在所有折次上的最优结果,*表示结果 $p < 0.05$,+表示结果 $p < 0.01$,“1,2,3,4”为折数。

示。HCPNet在多序列脑肿瘤MRI图像预后上具有更好的性能,在F1、准确率和召回率这3个评价指标上都取得了最好的结果。与IMDM、PMNet、MMRACR、Ensemble ResNet、MyoPS-Net和PFMI 6个模型方法相比,HCPNet获得了最高的F1评分,为0.85。相较以上6个模型分别提升了6%、9%、10%、5%、1%和1%;准确率达到0.88,与以上6种模型相比,分别提升了4%、6%、6%、2%、1%和1%;召回率达到0.87,与以上6种模型相比,分别提升了8%、12%、14%、9%、6%和7%;精确率达到0.84,相较于

IMDM、PMNet、MMRACR和PFMI也有一定提升,分别提升4%、6%、5%和0.4%。

同时, p 值检验小于0.01,表明本文方法在BraTS数据集上的结果同样具有统计显著性。综上所述,HCPNet能够实现多序列MRI的高效表达和优化,验证了HCPNet对脑肿瘤分类的适应性和泛化性。通过对比NIDCM和BraTS的结果,可以发现相同方法在BraTS上的效果优于在NIDCM上的效果,这证明了在NIDCM数据集上多模态CMR预后NIDCM是更为困难的任务。

表2 不同模型在BraTS数据集上的预后结果对比
Table 2 Comparison of prognostic results of different models on the BraTS dataset

方法	准确率					精确率				
	1	2	3	4	5	1	2	3	4	5
IMDM	0.84	0.80	0.79	0.81	0.81	0.80	0.74	0.79	0.80	0.78
PMNet	0.82	0.78	0.80	0.79	0.81	0.78	0.76	0.73	0.75	0.71
MMRACR	0.82	0.82	0.81	0.76	0.79	0.79	0.75	0.78	0.77	0.75
Ensemble ResNet	0.86	0.80	0.81	0.84	0.81	0.84	0.80	0.81	0.83	0.83
MyoPS-Net	0.87	0.85	0.83	0.87	0.86	0.86	0.85	0.84	0.84	0.86
PFMI	0.87	0.84	0.86	0.83	0.86	0.83	0.81	0.82	0.79	0.82
本文	0.88	0.86	0.86	0.87	0.88	0.84	0.83	0.82	0.80	0.81

方法	召回率					F1				
	1	2	3	4	5	1	2	3	4	5
IMDM	0.79	0.78	0.77	0.74	0.79	0.79 [*]	0.77	0.73	0.75	0.74
PMNet	0.75	0.70	0.74	0.73	0.71	0.76 [*]	0.76	0.75	0.71	0.73
MMRACR	0.73	0.73	0.73	0.70	0.71	0.75 ⁺	0.74	0.75	0.73	0.72
Ensemble ResNet	0.78	0.77	0.78	0.75	0.76	0.80 [*]	0.79	0.75	0.77	0.78
MyoPS-Net	0.81	0.80	0.78	0.79	0.76	0.84 ⁺	0.84	0.83	0.80	0.81
PFMI	0.80	0.76	0.79	0.80	0.78	0.84 ⁺	0.84	0.80	0.83	0.82
本文	0.87	0.85	0.81	0.85	0.83	0.85⁺	0.84	0.83	0.82	0.83

注:加粗字体表示各评价指标在所有折次上的最优结果,*表示结果 $p < 0.05$,+表示结果 $p < 0.01$,“1,2,3,4,5”为折数。

3.3 NIDCM数据集不同模态组合预后的对比实验

为了进一步评估HCPNet中每个部分的有效性以及多模态CMR在NIDCM预后中的相关性,采用不同的模态输入组合在临床数据集NIDCM上进行实验比较。与所使用的基础模型ResNet50(He等,2016)进行了比较,如表3所示,从而说明本文方法中不同模态组合以及特定模块对NIDCM预后的

贡献。

可以发现:1)与LGE序列相比,T1-native和T1-enhanced在NIDCM的预后中更有效,T1-native作为输入的模型,F1达到0.7315,相较于LGE性能提高了1.28%;T1-enhanced作为输入的模型,F1达到0.7416,相较于LGE性能提高2.29%。2)多模态CMR的直接concat联合表征模型F1达到0.7075,这种多序列联合表示方式很难将多个模态CMR图

像的相关性结合起来,甚至本文推测直接 Concat 联合表示的方式可能带来噪声,从结果上可以看出,其 F1 相较于单个 T1-native 序列输入下降 2.40%、较于单个 T1-enhanced 序列输入下降 3.41%、较于单个 LGE 序列输入下降 1.12%,这种联合表示方式很难保证 NIDCM 预后的表示。3)本文提出的 HCPNet 较 ResNet50-T1 native、ResNet50-T1 enhanced、ResNet50-LGE 和 ResNet50-Concat 在 F1 上分别提升了 8.10%、7.09%、9.38% 和 10.50%;在准确率上分别提升 7.58%、7.58%、13.64% 和 8.34%;在精确率上分别提升 11.56%、10.13%、12.95% 和 12.78%;在召回率上分别提升 4.34%、4.34%、7.47% 和 10.07%。综上,HCPNet 有利于对多模态 CMR 输入进行有效表示,从而提升多模态 CMR 预后 NIDCM 的性能。

表 3 不同模态输入的 NIDCM 预后结果对比
Table 3 Comparison of NIDCM prognosis results for different modal inputs

方法	NIDCM			
	准确率	精确率	召回率	F1
ResNet50-T1-native	0.780 3	0.729 4	0.762 2	0.731 5
ResNet50-T1-enhanced	0.780 3	0.743 7	0.762 2	0.741 6
ResNet50-LGE	0.719 7	0.715 5	0.730 9	0.718 7
ResNet50-Concat	0.772 7	0.717 2	0.704 9	0.707 5
本文	0.856 1	0.845 0	0.805 6	0.812 5

注:加粗字体表示各列最优结果。

3.4 消融实验

消融研究(表 4)验证了 HCPNet 中的每个模块在表示和优化中都发挥了重要的作用。

1)-CME 表示 CME 模块被消融。可以看到,在样本容量较小的情况下,F1 下降 1.58%。通过对互信息估计的对比,CME 修正正负表示的边界,精细优化整个数据集,由此证明了对比的相互差异有利于数据集的优化。2)-HMD 表示混合匹配蒸馏模块被消融。不同序列的 CMR 被连接起来表示预后,F1 下降 10.50%,由此证明 HMD 实现了多模态 CMR 对 NIDCM 预后的有效表征。3)在 2-seq 中,只取 T1-native 和 T1-enhanced 作为 HCPNet 的输入,缺少了来自 LGE 序列的信息补充,F1 下降 9.95%,证明了不同序列之间的相关性对 NIDCM 预后的有效性。4)

在 2-seq-HMD 中,T1-native 和 T1-enhanced 连接为输入,没有 HMD 混合匹配蒸馏中的分布和 logit,F1 相较于 2-seq 实验下降 2.73%。这表明提取多模态 CMR 图像相关性在多模态 CMR 图像预后 NIDCM 中起重要作用。以上消融实验表明了每个模块在整个模型中都起到了不可或缺的作用。

模型分析实验:为了进一步评价 HCPNet 中不同模块的作用机制,以 NIDCM 为例将不同部分的可视化表示进行展示。可以看出, E_{lge} 、 E_{tna} 、 E_{ten} 和 E_{joint} 的不同结果,如图 4 所示。1)NIDCM 在 T1-native 和 T1-enhanced 序列中的预后集中在左心室边缘。LGE 主要关注左心室容积。说明不同序列与 NIDCM 的预后有不同的相关性,也证明了 HCPNet 的 HMD 模块能够实现不同模态 CMR 的各种表示。2)将 E_{lge} 、 E_{tna} 、 E_{ten} 与 E_{joint} 进行比较, E_{joint} 不仅包含更多的相关区域(绿色虚线框),而且纠正了部分序列不同的区域(红色虚线框),说明 HMD 将多模态的表示融合为联合表示。3)通过 E_{lge} 、 E_{tna} 和 E_{ten} 对单个序列预后结果和 E_{joint} 对多模态 CMR 预后 NIDCM 结果的比较,可以看出, E_{joint} 在 NIDCM 上多模态 CMR 对 NIDCM 预后更为全面且更加精准,充分表明了 CME 在联合表征中的有效性。

表 4 不同模块对 HCPNet 性能的影响
Table 4 The impact of different modules on the performance of HCPNet

方法	NIDCM			
	准确率	精确率	召回率	F1
2-seq	0.757 6	0.704 6	0.729 2	0.713 0
2-seq-HMD	0.727 3	0.678 6	0.708 3	0.685 7
-HMD	0.848 5	0.717 2	0.722 2	0.707 5
-CME	0.848 5	0.828 1	0.780 1	0.796 7
所有模块	0.856 1	0.845 0	0.805 6	0.812 5

注:加粗字体表示各列最优结果,“2-seq”表示只有 T1-native 和 T1-enhanced 作为本文框架的输入。

3.5 超参数对比实验

考虑到多模态数据样本量少的原因,导致模型对于不同超参数敏感,本文通过不同的超参数实验来说明所提出模型的性能。

如图 5 所示,本文设置了多组学习率和 batch 的组合来验证,包括学习率为 1×10^{-3} 、 1×10^{-4} 、 1×10^{-5}

在大小分别为1、2、3、4、5的batch size设置下在心脏多模态NIDCM数据下的实验结果对比。从图5中可以发现:1)学习率过小,如 1×10^{-5} ,导致在相同的训练轮数(500轮)下模型收敛情况弱于学习率为 1×10^{-4} ;而当学习率过大时(1×10^{-3}),模型性能也会下降。因此,本实验中所选择的学习率为 1×10^{-4} ,是适用于本多模态融合场景的。2)随着batch size的

增加(从1到4),模型的性能是不断增加的。这表明通过增加每次输入的模型的样本数量能够提升模型的性能,但当batch size到达5时,模型的性能不再上升,产生了性能的下降。因此,本模型选择batch size为4作为最优的参数设置。3)在不同的超参数组合下,采用的学习率为 1×10^{-4} 和batch size为4的组合带来最佳模型性能。

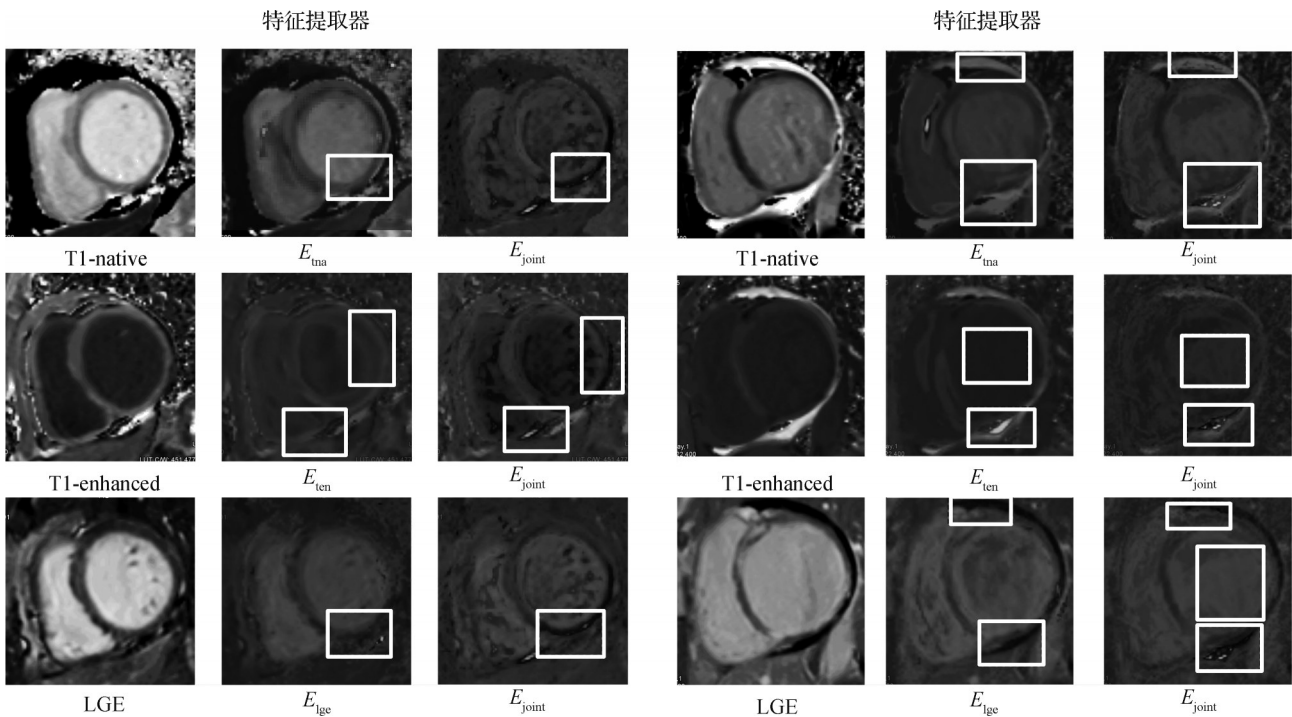


图4 不同序列CMR的表示和联合表示的对比

Fig. 4 Comparison of representation and joint representation of different sequence CMRs

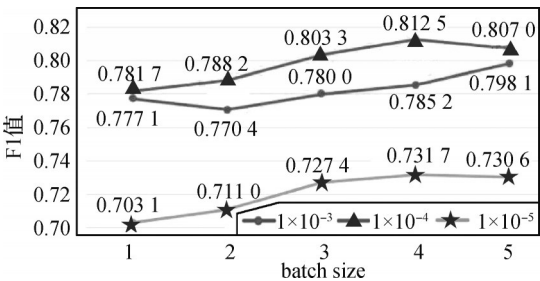


图5 不同超参数的模型性能对比

Fig. 5 Comparison of model performance with different hyperparameters

4 结 论

多模态CMR图像间的信息互补性和相关性复杂,导致多模态CMR图像在对NIDCM的预后上表

征困难;同时由于预后任务标注困难,有标注数据规模小,导致模型容易陷入局部最优。针对这两个挑战,本文提出了一种新的多序列表征和对比学习框架:基于混合匹配蒸馏与对比互信息估计的预后网络模型(HCPNet),用于小样本上的多模态CMR对NIDCM的预后。该方法设计了混合匹配蒸馏模块和对比互估计模块。混合匹配蒸馏模块利用逻辑分布一致性将不同图像特征分布关联匹配,以此约束深度学习网络中多模态特征的提取和引导联合表征;对比互估计模块估计出多模态分布上的潜在的分类边界,以此作为预后模型的正则化项,避免模型在有限的数据上陷入局部最优,克服了多模态CMR预后中难以表示和局部最优的问题,同时利用多模态CMR之间的潜在互信息增强深度学习模型在小样本场景下的模型优化,最终使得多模态CMR对

NIDCM的预后结果更加准确。此外,本文验证了HCPNet在多模态MRI脑肿瘤分类任务上的鲁棒性和泛化性。

参考文献(References)

- Arenja N, Riffel J H, Fritz T, André F, Aus dem Siepen F, Mueller-Hennessen M, Giannitis E, Katus H A, Friedrich M G and Buss S J. 2017. Diagnostic and prognostic value of long-axis strain and myocardial contraction fraction using standard cardiovascular MR imaging in patients with nonischemic dilated cardiomyopathies. *Radiology*, 283(3): 681-691 [DOI: 10.1148/radiol.2016161184]
- Baltrušaitis T, Ahuja C and Morency L P. 2019. Multimodal machine learning: a survey and taxonomy. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 41(2): 423-443 [DOI: 10.1109/TPAMI.2018.2798607]
- Chen H L, Gao M Y, Zhang Y, Liang W B and Zou X C. 2019. Attention-based multi-NMF deep neural network with multimodality data for breast cancer prognosis model. *BioMed Research International*, 2019: #9523719 [DOI: 10.1155/2019/9523719]
- Ding Y, Zheng W, Geng J, Qiu L Y and Qin Z G. 2023. Multi-level parallel neural networks based multimodal human brain tumor image segmentation framework. *Journal of Image and Graphics*, 28(7): 2182-2194 (丁熠, 郑伟, 耿技, 邱沛谊, 秦志光. 2023. 基于多层级并行神经网络的多模态脑肿瘤图像分割框架. *中国图象图形学报*, 28(7): 2182-2194) [DOI: 10.11834/jig.211197]
- Havaei M, Guizard N, Chapados N and Bengio Y. 2016. HeMIS: hetero-modal image segmentation//*Proceedings of the 19th International Conference on Medical Image Computing and Computer-Assisted Intervention*. Athens, Greece: Springer: 469-477 [DOI: 10.1007/978-3-319-46723-8_54]
- 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*. Las Vegas, USA: IEEE: 770-778 [DOI: 10.1109/CVPR.2016.90]
- Li C, Song L Y, Zhu G P, Hu B L, Liu X B and Wang Q. 2022. Alzheimer's level classification by 3D PMNet using PET/MRI multi-modal images//*Proceedings of 2022 IEEE International Conference on Electrical Engineering, Big Data and Algorithms*. Changchun, China: IEEE: 1068-1073 [DOI: 10.1109/EEBDA53927.2022.9744769]
- Li X D, Zhang C F, Sun X X, Yang X M, Zhang M J, Wang Q and Zhu Y S. 2020. Prognostic factors of pulmonary hypertension associated with connective tissue disease: pulmonary artery size measured by chest CT. *Rheumatology*, 59(11): 3221-3228 [DOI: 10.1093/rheumatology/keaa100]
- Maron B J, Towbin J A, Thiene G, Antzelevitch C, Corrado D, Arnett D, Moss A J, Seidman C E and Young J B. 2006. Contemporary definitions and classification of the cardiomyopathies: an American heart association scientific statement from the council on clinical cardiology, heart failure and transplantation committee; quality of care and outcomes research and functional genomics and translational biology interdisciplinary working groups; and council on epidemiology and prevention. *Circulation*, 113(14): 1807-1816 [DOI: 10.1161/CIRCULATIONAHA.106.174287]
- Menze B H, Jakab A, Bauer S, Kalpathy-Cramer J, Farahani K, Kirby J, Burren Y, Porz N, Slotboom J, Wiest R, Lanczi L, Gerstner E, Weber M A, Arbel T, Avants B B, Ayache N, Buendia P, Collins D L, Cordier N, Corso J J, Criminisi A, Das T, Delingette H, Demiralp C, Durst C R, Dojat M, Doyle S, Festa J, Forbes F, Geremia E, Glocker B, Golland P, Guo X T, Hamamci A, Iftekharuddin K M, Jena R, John N M, Konukoglu E, Lashkari D, Mariz J A, Meier R, Pereira S, Precup D, Price S J, Raviv T R, Reza S M S, Ryan M, Sarikaya D, Schwartz L, Shin H C, Shotton J, Silva C A, Sousa N, Subbanna N K, Szekely G, Taylor T J, Thomas O M, Tustison N J, Unal G, Vasseur F, Wintermark M, Ye D H, Zhao L, Zhao B S, Zikic D, Prastawa M, Reyes M and Van Leemput K. 2015. The multimodal brain tumor image segmentation benchmark (BRATS). *IEEE Transactions on Medical Imaging*, 34(10): 1993-2024 [DOI: 10.1109/TMI.2014.2377694]
- Nwabuo C C and Vasan R S. 2020. Pathophysiology of hypertensive heart disease: beyond left ventricular hypertrophy. *Current Hypertension Reports*, 22(2): #11 [DOI: 10.1007/s11906-020-1017-9]
- Qi X M, He Y T, Yang G Y, Chen Y, Yang J, Liu W, Zhu Y S, Xu Y, Shu H Z and Li S. 2022. MVSCAN: spatial-aware multi-view CMR fusion for accurate 3D left ventricular myocardium segmentation. *IEEE Journal of Biomedical and Health Informatics*, 26(5): 2264-2275 [DOI: 10.1109/JBHI.2021.3122581]
- Qiu J Y, Li L, Wang S H, Zhang K, Chen Y Y, Yang S and Zhuang X H. 2023. MyoPS-Net: myocardial pathology segmentation with flexible combination of multi-sequence CMR images. *Medical Image Analysis*, 84: #102694 [DOI: 10.1016/j.media.2022.102694]
- Quarta G, Sado D M and Moon J C. 2011. Cardiomyopathies: focus on cardiovascular magnetic resonance. *British Journal of Radiology*, 84(special issue 3): S296-S305 [DOI: 10.1259/bjr/67212179]
- Shankar S, Thompson L and Fiterau M. 2024. Progressive fusion for multimodal integration//*Proceedings of the ICLR 2024*. [s.l.]: [s.n.]
- Sun D D, Wang M H and Li A. 2019. A multimodal deep neural network for human breast cancer prognosis prediction by integrating multidimensional data. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 16(3): 841-850 [DOI: 10.1109/TCBB.2018.2806438]
- Uthoff J, Alabed S, Swift A J and Lu H P. 2020. Geodesically smoothed tensor features for pulmonary hypertension prognosis using the heart and surrounding tissues//*Proceedings of the 23rd International Conference on Medical Image Computing and Computer Assisted Intervention*. Lima, Peru: Springer, 2020: 253-262 [DOI: 10.

1007/978-3-030-59713-9_25]

- Xiao X J, Zhao J F and Li S. 2022. Task relevance driven adversarial learning for simultaneous detection, size grading, and quantification of hepatocellular carcinoma via integrating multi-modality MRI. *Medical Image Analysis*, 81: #102554 [DOI: 10.1016/j.media.2022.102554]
- Xing X H, Hou Y N, Li H, Yuan Y X, Li H S and Meng M Q H. 2021. Categorical relation-preserving contrastive knowledge distillation for medical image classification//*Proceedings of the 24th International Conference on Medical Image Computing and Computer Assisted Intervention*. Strasbourg, France: Springer: 163-173 [DOI: 10.1007/978-3-030-87240-3_16]
- Yan R, Ren F, Rao X S, Shi B R, Xiang T G, Zhang L L, Liu Y D, Liang J, Zheng C H and Zhang F. 2019. Integration of multimodal data for breast cancer classification using a hybrid deep learning method//*Proceedings of the 15th International Conference on Intelligent Computing Theories and Application*. Nanchang, China: Springer: 460-469 [DOI: 10.1007/978-3-030-26763-6_44]
- Yang X, Xi X M, Yang L, Xu C Z, Song Z Y, Nie X S, Qiao L S, Li C L, Shi Q L and Yin Y L. 2022. Multi-modality relation attention network for breast tumor classification. *Computers in Biology and Medicine*, 150: #106210 [DOI: 10.1016/j.combiomed. 2022. 106210]
- Yu L, Liu X X, Yan C Y, Li J R, Zhang Z Q, Huang Y Z and Xu J. 2022. Brain glioma tumor segmentation and survival prediction from multi-modality MRIs via cascaded U-Net. *Journal of Image and Graphics*, 27(3): 850-861 (余力, 刘霄雪, 闫朝阳, 李建瑞, 张志强, 黄韞枢, 徐军. 2022. 面向多模态MRI脑胶质瘤区域三维分割与生存期预测的级联U-Net网络. *中国图象图形学报*, 27(3): 850-861) [DOI: 10.11834/jig.210137]

Zhou T X, Ruan S, Guo Y and Canu S. 2020. A multi-modality fusion network based on attention mechanism for brain tumor segmentation//*The 17th IEEE International Symposium on Biomedical Imaging*. Iowa City, USA: IEEE: 377-380 [DOI: 10.1109/ISBI45749. 2020.9098392]

作者简介

魏然,女,硕士研究生,主要研究方向为医学影像分析与机器学习。E-mail: 2821629674@qq.com

杨冠羽,通信作者,男,教授,博士生导师,主要研究方向为医学影像分析与机器学习。E-mail: yang.list@seu.edu.cn

戚晓明,男,博士研究生,主要研究方向为医学影像分析与机器学习。E-mail: qixiaoming12138@163.com

何宇霆,男,博士研究生,主要研究方向为医学影像分析与机器学习。E-mail: csyutingh@ust.hk

江升,男,硕士研究生,主要研究方向为医学影像分析与机器学习。E-mail: 1719646079@qq.com

钱雯,女,博士研究生,主要研究方向为循环影像诊断。E-mail: aleenxu@hotmail.com

徐怡,女,主任医师,主要研究方向为循环影像诊断。E-mail: aleenxu@hotmail.com

祝因苏,男,主任医师,主要研究方向为循环影像诊断。E-mail: zhuyinsu@njmu.edu.cn

Pascal Haigron,男,教授,主要研究方向为信号与数字图像处理。E-mail: pascal.haigron@univ-rennes.fr

舒华忠,男,教授,主要研究方向为医学影像分析。E-mail: shu.list@seu.edu.cn