

深度学习超声心动图评估心力衰竭患者左室射血分数：端到端模型与分步流程的比较研究

韩 钰¹, 温盼盼², 周 青¹, 曹 省¹

1. 武汉大学人民医院, 湖北 武汉 430060; 2. 武汉大学遥感信息工程学院, 湖北 武汉 430000

【摘要】目的 构建和比较基于深度回归的视频预测模型(端到端模型)及基于多模型分步的自动计算左心室射血分数(LVEF)的流程方法,并与超声下标准 Simpson 法进行比较,评估其在整体精度、心功能分型一致性及射血分数误差分布。**方法** 构建两种代表性的 LVEF 计算流程:端到端模型使用卷积神经网络与时空卷积结构对超声视频进行特征提取与回归预测;多模型分步流程方法则采用“帧识别+分割掩码+几何运算”三步流程,其中包括通过自监督学习进行左室分割,并利用 Simpson 法与面积长度法计算 LVEF。本研究采用 EchoNet-Dynamic 数据集(用于训练)和 CAMUS 数据库,并在武汉大学人民医院的临床影像数据上进行外部验证。模型性能通过偏差、平均绝对误差、相关性和一致性等指标进行评估。**结果** 在 CAMUS 数据库上,基于多模型分步流程中分割模型的面积长度法计算的 LVEF 比 Simpson 法更接近人工测量,且精度略优($r=0.74$ vs $r=0.72$, $MAE=6.54$ vs 8.39)。在外部验证中,端到端模型在四腔心与两腔心切面的预测性能上相似($r=0.76$ vs 0.77 , $MAE=5.46$ vs 5.30)。与多模型分步流程中 Simpson 方法相比,端到端模型结合双平面的均值表现更为准确($r=0.82$ vs 0.59 , $MAE=4.65$ vs 9.69)。多模型分步流程中 Simpson 法相较于面积长度法仍然更为低估 LVEF (Bias: -6.64 vs -3.85),但相关性较好($r: 0.59$ vs 0.56)。**结论** 本研究构建的端到端模型在 LVEF 预测中的准确性与超声下标准 Simpson 法人工测量相当,相比多步计算的优秀模型,在精度与简化流程方面具有优势,但临床可解释性仍是需要解决的问题。

【关键词】 心力衰竭;人工智能;超声心动图;端到端模型;多模型分步流程

【中图分类号】 R541.6;R445.1

【文献标志码】 A

【文章编号】 1672-6170(2026)04-0029-07

Deep learning echocardiography for assessing left ventricular ejection fraction in patients with heart failure: a comparative study of end-to-end models and step-by-step models HAN Yu¹,

WEN Pan-pan², ZHOU Qing¹, CAO Sheng¹ 1. Renmin Hospital of Wuhan University, Wuhan 430060, China; 2. School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430000, China

【Corresponding author】 CAO Sheng

【Abstract】Objective To construct and compare a video prediction model (end-to-end model) based on deep regression and an automated calculation procedure for left ventricular ejection fraction (LVEF) based on multi-model stepwise, and compare the standard Simpson's method under ultrasound to evaluate their overall accuracy and consistency in cardiac function classification and LVEF error distribution. **Methods** Two representative LVEF calculation processes were constructed. One was end-to-end model using convolutional neural networks and spatiotemporal convolutional structures for feature extraction and regression prediction from echocardiographic videos. Another was multi-model step-by-step process method adopting a three-step process of "frame recognition + segmentation mask + geometric operation". This included self-supervised learning for left ventricular segmentation and LVEF calculation using both Simpson's method and the area-length method. This study used the EchoNet-Dynamic dataset (for training) and the CAMUS database. External validation was performed using clinical imaging data from our hospital. Model performance was assessed using bias, mean absolute error, correlation and agreement metrics. **Results** On the CAMUS database, the LVEF calculated based on the area-length method of segmented models in a multi-model step-by-step process was closer to human measurement than the Simpson method. Its accuracy was slightly better ($r=0.74$ vs. $r=0.72$, $MAE=6.54$ vs. 8.39). In external validation, the end-to-end model showed similar predictive performance in the four-chamber and two-chamber views ($r=0.76$ vs. 0.77 , $MAE=5.46$ vs. 5.30). Compared to the Simpson method in the multi-model step-by-step process, the end-to-end model combined with the biplane means performance was more accurate ($r=0.82$ vs. 0.59 , $MAE=4.65$ vs. 9.69). Compared with Simpson's method in the multi-model stepwise pipeline, the end-to-end model combined with biplane averaging provided more accurate results ($r=0.82$ vs. 0.59 , $MAE=4.65$ vs. 9.69). Within the multi-model stepwise pipeline, Simpson's method underestimated LVEF more than the area-length method (bias: -6.64 vs. -3.85). However, it showed better correlation ($r: 0.59$ vs. 0.56). **Conclusions** The end-to-end model constructed in this study has comparable accuracy in LVEF prediction to manual measurement using the standard Simpson method under ultrasound. Compared to excellent models with multi-step calculations, it has advantages in terms of accuracy and simplified procedures. However, clinical interpretability remains an issue that needs to be addressed.

【基金项目】湖北省自然科学基金项目(编号:2024AFB883); 武汉大学大学生创新训练项目(国家级)(编号:202510486148); 湖北省卫生健康科技项目(编号:WJ2025M122); 武汉大学医学部教学研究项目/武汉大学医学部医学发展基金(编号:2025YB19)

【通讯作者简介】曹省,男,副教授,副主任医师,博士。中国超声心动图学会理事,中国超声心动图学会冠心病与血管组委员,中国工业与应用数学学会数学与医学交叉分会委员,中国超声医学工程学会分子影像专委会青年委员,中国超声医学工程学会超声治疗与生物效应专委会青年委员,湖北省超声质控中心委员兼办公室主任,湖北省中西医结合学会超声医学专委会常务委员,武汉超声医学工程学会理事。主要研究方向:人工智能与超声新技术应用研究;超声生物效应与诊疗一体化研究。

【Key words】 Heart failure; Artificial intelligence; Echocardiography; End-to-end model; Multi-process model

近年来,基于人工智能(artificial intelligence, AI)的超声心动图左室射血分数(left ventricular ejection fraction, LVEF)自动测量取得了快速进展。Ouyang 等构建了包含 10030 条经胸四腔心视频的 EchoNet-Dynamic 公共数据集,并提出联合左室分割和视频回归的深度学习算法,在 LVEF 预测上实现约 4~6 个百分点的平均绝对误差,其性能已接近专家读片间的一致性,被视为该领域的里程碑工作^[1]。在此基础上,后续研究利用多视图卷积网络和 Transformer 融合多标准切面信息构建模型以提升在复杂运动模式下的预测稳健性^[2]。针对临床标注成本高的问题,半监督与自监督方法通过先学习左室分割或心脏结构表征,再进行 LVEF 回归,仅依赖一半的标注数据即可获得接近完全监督的精度,显著改善了跨设备、跨中心的泛化能力^[3]。

综合现有证据,AI 不仅在 LVEF 测量精度上已接近经验丰富的超声医生,更在重复性和标准化方面具有优势。但不同的自动化计算路径或流程存在不同程度的误差,目前对于不同计算 LVEF 的工作流的比较研究还存在不足。因此本研究分别构建了两条具有代表性的 LVEF 计算流程:一条采用基于深度回归的视频预测模型,直接从心尖四腔、两腔超声序列输出连续 LVEF,即端到端输出;另一条则是多种模型的分步组合计算 LVEF。以经验丰富超声专家使用超声机采用的标准 Simpson 法获得的 LVEF 作为参考真值,系统比较“端到端模型预测”和“多模型分步流程”两类结果与专家结果之间的差距,分析其在整体精度、心功能分型一致性及 LVEF 误差分布。

1 资料与方法

1.1 数据来源 以 EchoNet-Dynamic 公共数据集(约 1 万份超声视频)作为端到端模型和多模型分步流程方法的训练集,CAMUS 数据库的图像(500

例患者,每例各 2 张)作为多模型分步流程方法的验证集。回顾性纳入 2024 年 1~12 月在武汉大学人民医院超声科收集的接受经胸超声心动图检查的成人患者心动图数据,共计 200 例患者对应的两腔心(apical two-chamber view, A2C)和四腔心(apical four-chamber view, A4C)超声心动图视频(合计 400 份超声视频),单个视频图像 70~1000 帧,作为端到端模型与多模型分步流程方法的外部验证集。纳入标准:①符合《中国心力衰竭诊断和治疗指南 2024》^[4]诊断标准;存在心力衰竭症状和体征,且 BNP 或 NT-proBNP 水平升高;②经胸超声心动图检查存在心脏结构或功能异常;③心尖四腔心和两腔心切面动态图像存储完整、清晰,满足 AI 分析要求。排除标准:①急性心肌梗死;②严重瓣膜病、先天性心脏病或肥厚型梗阻性心肌病;③严重肾功能不全;③经胸超声心动图图像质量不满足 AI 分析要求(心内膜边界不清、伪影严重、心动周期不完整等)。

1.2 模型介绍 端到端模型利用卷积神经网络与时空卷积结构对输入的超声视频进行特征提取与建模,并结合弱监督学习策略对左室进行语义分割与运动特征编码,从而实现对 LVEF 的直接回归预测。多模型分步流程方法采用了“帧识别+分割掩码+几何运算”三步流程模型,如图 1 所示。其中帧识别模型采用的是一种双向时序网络,通过对心脏收缩期与舒张期标注的训练数据进行学习,能够有效地捕捉并识别超声视频中对应的心脏运动模式,有效定位超声视频中收缩和舒张末期的帧;分割掩码模型则基于自监督策略,通过综合考虑心动周期的整体运动模式以及少量掩码标注,来进行掩码的学习与生成,提高了数据稀缺下的分割精度;几何运算则分别利用了双平面 Simpson 法和 AL 法的原理进行程序运算。

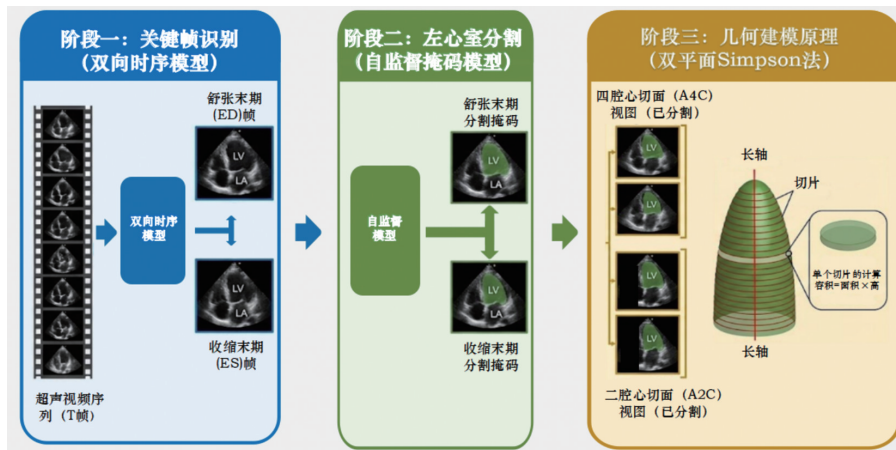


图 1 多模型分步流程方法框架

1.3 数据处理

1.3.1 训练与内部验证 针对基于视频回归的算法模型,本研究选用 EchoNet-Dynamic 数据集开展模型训练与内部验证,按照 7:3 的比例将数据集随机划分为训练集与内部测试集,完成模型的参数优化与初步性能评估。

1.3.2 模型外部验证 采用数据来源界定的武汉大学人民医院临床影像数据作为外部独立验证集,LVEF 真值来源于超声报告中的测量值,验证模型在真实临床样本中的诊断效能与稳定性。

1.3.3 比较对象 研究中所比较的对象包括了分步流程的 Simpson 法、AL 法;端到端模型的 A2C 视频单独输出的 LVEF 值(表 2~4 中记作 A2C)、A4C 视频单独输出的 LVEF 值(表 2~4 中记作 A4C)、二者的平均值(表 2~4 中记作 Mean);最后,数据集中给出的患者 LVEF 值或医院患者超声报告中 LVEF 值(表 1~4 中记作 Ref)。

1.4 统计学方法 使用 R 软件进行数据分析。对于连续变量,使用均值和标准差(SD)进行描述性统计分析,分类变量则以频数和百分比呈现。模型性能评估通过偏差(Bias)、平均绝对误差(MAE)、均方根误差(RMSE)以及相关系数等指标进行,偏差和相关系数同时给出 95%置信区间(CI)。各对象平均绝对

误差的比较采用 Wilcoxon 秩和检验,相关系数 r 采用 Steiger 检验以明确不同方法之间是否存在差异性。模型与人工测量值之间的一致性采用 Bland-Altman 一致性限(LoA)分析,上下限分别为 LoAupper(上限)和 LoAlower(下限)。另外,使用热力图展示了各方法对于心力衰竭(心衰)类型分组与专家分组的区别,心衰分类依照 2024 年指南^[4],分为射血分数降低型心衰(HFrEF)($LVEF \leq 40\%$),射血分数轻微降低型心衰(HFmrEF)($40\% < LVEF < 50\%$),射血分数保留型心衰(HFpEF)($LVEF \geq 50\%$)。所有统计检验均使用双侧检验, $P < 0.05$ 为差异有统计学意义。

2 结果

2.1 基于 CAMUS 数据集的多模型分步流程结果比较 在 CAMUS 数据集中的 500 例患者的 1000 张 A2C 和 A4C 图像中,采用分割网络重新处理图像,并比对两种 LVEF 计算方法与 CAMUS 数据集中给出的对应患者 LVEF 值的差异。结果显示,双平面 Simpson 法计算的精确度要稍低于 AL 法($r = 0.72$ vs 0.74 , $MAE = 8.39$ vs 6.54)。对 CAMUS 数据集的图像进行了人工测量计算结果显示,测量结果与数据集所给患者的 LVEF 值的相关性最强($r = 0.96$),但存在系统性高估(Bias = 8.11, LoA 95% CI 为 1.34%~14.88%)。见表 1。

表 1 CAMUS 数据集的多模型分步流程方法结果对比

比较对象	Bias (95% CI)	SD	LoAlower	LoAupper	MAE	RMSE	r	95% CI
Simpson vs Ref	-6.73 (-7.50, -5.96)	8.54	-23.46	10	8.39	10.86	0.72	0.67~0.76
人工测量 vs Ref	8.11 (7.81, 8.42)	3.46	1.34	14.88	8.12	8.82	0.96	0.95~0.96
AL vs Ref	0.7 (-0.07, 1.47)	8.6	-16.15	17.55	6.54	8.62	0.74	0.70~0.78

2.2 外部验证结果的比较 使用基于深度回归的视频预测模型,直接从 A2C、A4C 的超声序列分别输出 LVEF 并求均值(Mean)和多模型分步流程计算(AL 法和 Simpson 法)的结果如表 2 所示,A2C 和 A4C 切面预测的准确性相差不大($r = 0.76$ vs 0.77 , $MAE = 5.46$ vs 5.30),当把二者的预测值求平均后,结果更准确($r = 0.82$, $MAE = 4.65$)。多模型分步流程方法预测的准确性低于直接回归预测的视频模型。在多模型分步流程方法中,Simpson 法相对于 AL 法标准差小,相关性更大,但差异无统计学意义。见表 3。

端到端模型与多模型分步流程方法的专家报

表 2 外部验证数据集的端到端模型与多模型分步流程结果比较

比较对象	Bias	Bias (95% CI)	SD	LoAlower	LoAupper	MAE	RMSE	r	95% CI
A2C vs Ref	-0.30	(-1.21~0.61)	6.54	-13.12	12.51	5.30	6.53	0.77	0.71~0.82
A4C vs Ref	1.66	(0.74~2.58)	6.59	-11.25	14.57	5.46	6.78	0.76	0.70~0.81
Mean vs Ref	0.68	(-0.08~1.44)	5.48	-10.07	11.43	4.65	5.51	0.82	0.77~0.86
AL vs Ref	-3.85	(-5.44~-2.25)	11.42	-26.24	18.54	9.48	12.03	0.56	0.45~0.65
Simpson vs Ref	-6.64	(-8.13~-5.15)	10.69	-27.60	14.32	9.69	12.56	0.59	0.49~0.67

前三行为端到端模型结果,后两行为多模型分步流程结果

告测量值的差的直方图见图 2、图 3。端到端模型的 0 值两端分布相对于多模型分步流程方法更加密集,尤其在于取均值后整体误差更小。另外,心衰分型的混淆矩阵热力图显示,端到端模型测量 LVEF 对应的心衰类别在三类中与专家一致的占比均最高(图 4),而多模型分步流程方法则都显然低估了部分 HFmrEF 与 HFpEF 患者的 LVEF(图 5)。各类方法的误差范围占比见表 4,总体说来端到端模型控制测量误差在 10% 的人数占比约 60%,多流程则占比约 30%。

表 3 外部验证数据集结果之间的性能差异统计学检验

比较对象	MAE1	MAE2	PMAE	r_1	r_2	P_r
A2C vs A4C	5.30	5.46	0.677	0.774	0.762	0.67
A2C vs Mean	5.30	4.649	0.00342	0.774	0.824	<0.001
A4C vs Mean	5.46	4.649	0.00249	0.762	0.824	<0.001
Simpson vs Mean	9.688	4.649	<0.001	0.587	0.824	<0.001
AL vs Mean	9.481	4.649	<0.001	0.557	0.824	<0.001
Simpson vs AL	9.688	9.481	0.605	0.587	0.557	0.453

MAE1:每列对比的第一个对象与人工测量比对的 MAE 值,MAE2:每列对比的第二个对象与人工测量比对的 MAE 值,PMAE:两个对象与人工测量 MAE 值比较的差异

进一步探究多模型分步流程中 Simpson 法与 AL 法计算 LVEF 产生差异的原因,通过查看一些影像分割结果(图 6),a 图中显示了模型精确依照像素差异进行了掩码划分,但不平滑,b 图则为模型给出了较平滑的掩码结果。不同平滑程度的掩码结果,影响了 LVEF 的测量,也导致了不同方法计算结果的差异。另外,通过提取了外部数据集中分割模型计算的收缩与舒张末期容积并进行了统计分析显示,Simpson 法和 AL 法计算的结果存在一定差异(表 5)。Simpson 法的平均舒张末期容积约为 120 ml,收缩末期容积为 67 ml;中位舒张末期容积约为 114 ml,收缩末期容积为 60 ml。AL 法的平均舒张

末期容积约为 103 ml,收缩末期容积约为 55 ml;中位舒张末期容积约为 96.5 ml,收缩末期容积约为 49.5 ml。

表 4 外部验证数据集中模型给出 LVEF 与人工测量的误差分布 [n(%)]

方法	相差<5%	相差 5%~10%	相差>10%
AL	35(17.5)	30(15.0)	135(67.5)
Simpson	37(18.5)	29(14.5)	134(67.0)
A2C	59(29.5)	47(23.5)	94(47.0)
A4C	56(28.0)	58(29.0)	86(43.0)
Mean	62(31.0)	52(26.0)	86(43.0)

表 5 外部验证数据集采用多模型分步流程的收缩与舒张末期容积分析表

容积值(ml)	平均值	标准差	中位数	Q1	Q3	最小值	最大值
舒张末期(Simpson 法)	119.96	37.69	114	94	139.25	42	251
收缩末期(Simpson 法)	67.31	29.74	60	46	79.25	24	172
舒张末期(AL 法)	103.35	40.82	96.5	78	123	25	275
收缩末期(AL 法)	55.24	29.85	49.5	35	67	12	229

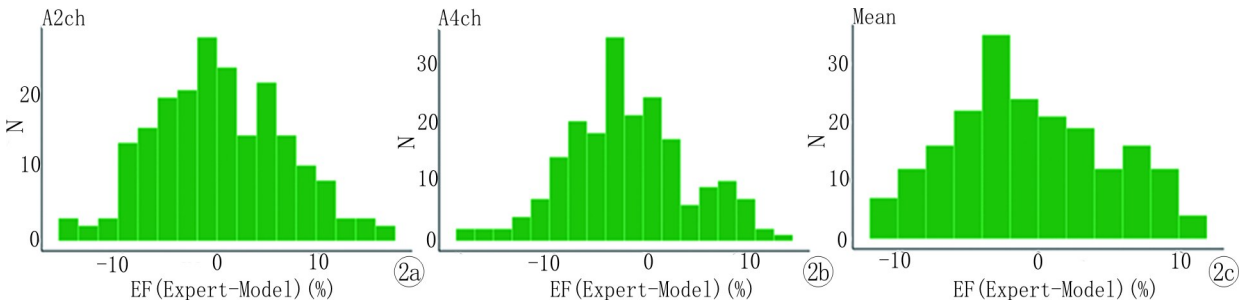


图 2 端到端模型人工测量值与各方法计算值差值的直方图 a:A2c;b:A4c;c:平均

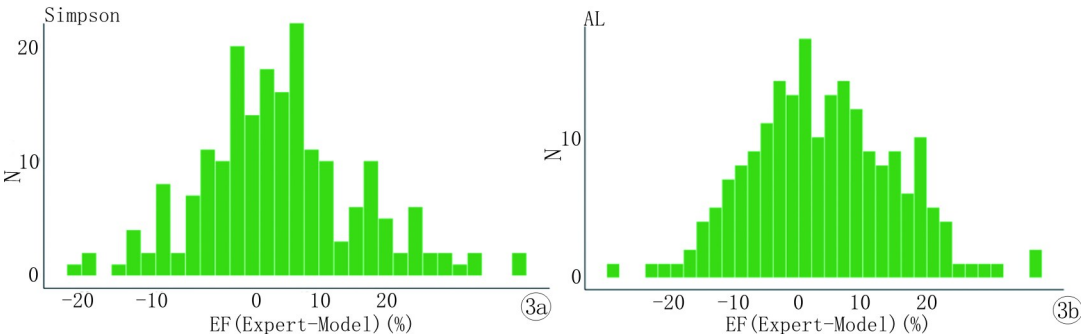


图 3 多模型分步流程人工测量值与各方法计算值差值的直方图 3a:Simpson 法;3b:AL 法

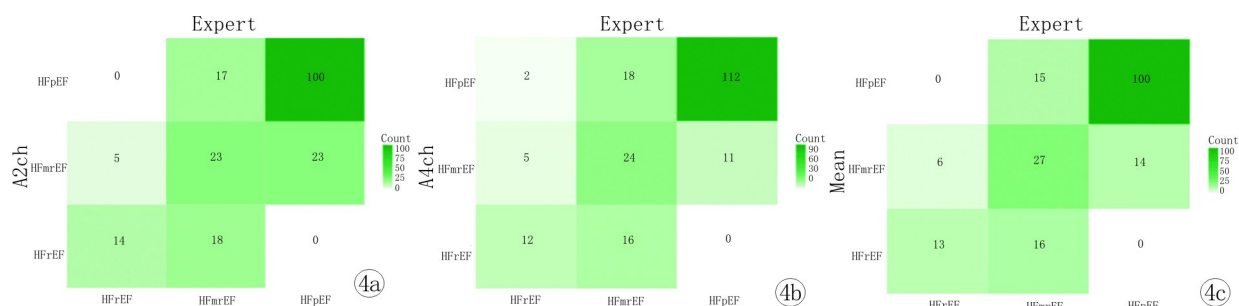


图4 端到端模型心衰分组热力图 a;A2c;b:A4c;c:平均

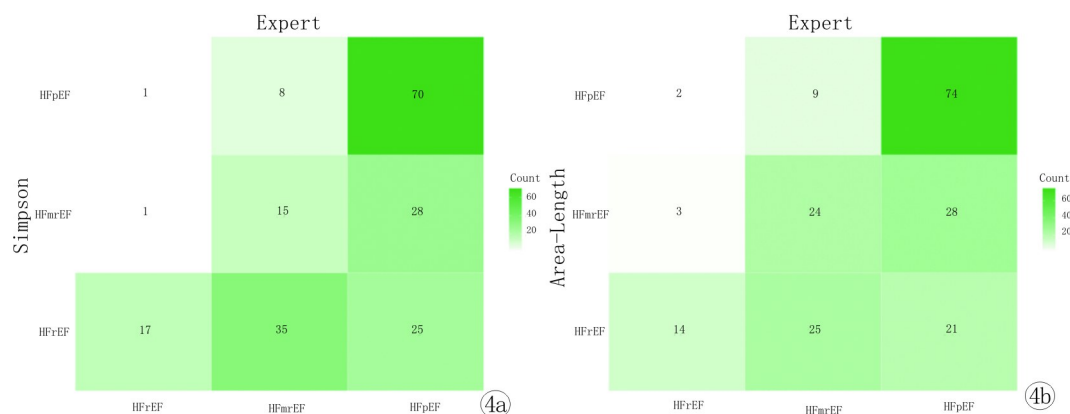


图5 多模型分步流程心衰分组热力图 a;Simpson 法;b:AL 法

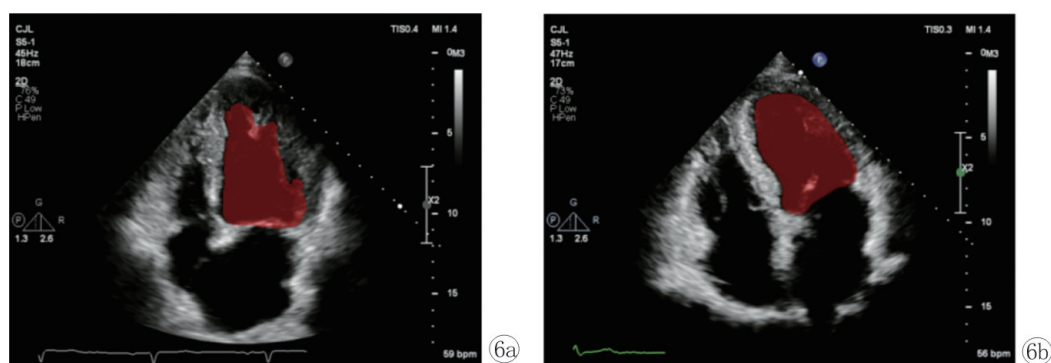


图6 AI掩码模型给出的分割结果 a;受噪声影响的不连续平滑分割;b:受噪声影响的连续平滑分割

3 讨论

尽管心脏磁共振被认为是测量 LVEF 的“金标准”,但由于成本高昂、耗时且不可移动,二维经胸超声心动图因其无创、实时、便携及低成本的优势,成为临床首选的成像模态。目前的临床标准操作流程通常采用 Simpson 双平面法,该方法要求超声医师在 A2C 和 A4C 上,分别识别舒张末期和收缩末期帧,并手动描绘心内膜边界。随着卷积神经网络的兴起,第一代深度学习采用“模仿人类”的策略,即训练模型自动分割左心室,然后应用 Simpson 公式计算 LVEF。虽然这种基于分割的方法提高了效率,但其核心痛点并未解决:它依然依赖于清晰的边界定义,且无法利用全周期的时空信息^[5-9]。因此尽管 AI 分割在其他影像学应用上取得了巨大进步,但超声心动图测量 LVEF 上仍然需评估各类

假设带来的误差。从 A2C 与 A4C 测量角度的推荐方法,一边是对边界细节敏感的 Simpson 法,另一边是更依赖于整体形态但比较稳健的 AL 法,把握实际场景下的适用性极其重要。AL 法由于容积计算仅依赖于两个截面面积和长轴长度,局部的分割噪声(如个别像素的凹陷或突起,通常可能是乳头肌在心腔中显示(图 6a),在积分成总面积的过程中被平均化了。这意味着,即使 AI 分割掩码不够完美,只要整体轮廓覆盖正确,AL 法计算出的容积往往比较稳定。

本研究中,在 CAMUS 数据集上利用分割模型计算 LVEF 的实验显示面积长度法在 AI 分割的结果上计算 LVEF 较双平面 Simpson 法要接近于原文^[5]的计算结果。这类差异来自于采用的模型,原文中采用了经典的 U-Net 架构,而本研究采用的则

是自监督预训练+弱监督分割学习^[10],因此掩码结果略有差异,使采用 Simpson 法平均测量值降低了 6.73%。根据已有文献的报道,心脏磁共振检查采用 AL 法平均会比 Simpson 法高 5%^[11],因而采用超声心动图检查误差推测大概在 5%~10%以内,恰好补齐了低估的部分误差。表 5 中对于两种测量方法在舒张和收缩末期容积的计算值可以发现,AL 法普遍会比 Simpson 法测算低约 10~20ml,因此如果容积变化量在两者是近似的时候,往往 Simpson 法会低估 LVEF,与既往在心脏磁共振领域研究的文献的结论是相一致的^[12, 13]。在外部验证结果中, Simpson 法的结果与人工测量的趋势比 AL 法更近似,且依然整体比 AL 法更加低估 LVEF (Bias: -6.64 vs -3.85)。进一步的心衰类型分组显示(见图 4 和图 5),对于射血分数保留型和降低型心力衰竭,所有的模型结果都能比较准确地分类出来,而轻度降低型则比较难以区分,端到端模型两端分布比较均匀,而多模型分步流程的 Simpson 法和 AL 法均有明显低估趋势。在 LVEF 差值直方图中,也能看出多模型分步流程结果(Simpson 法与 AL 法)的整体右偏(预测值低估)。需要指出的是,在实际临床工作中,现有的指南更推荐使用双平面 Simpson 法,这是在能够十分准确地获取心内膜边界的前提下进行的,而利用 AL 法更稳健的前提则是整体的内膜描绘因边缘模糊或 AI 泛化性不足而不够准确时。临床上已经应用的软件如 AutoEF 是基于心肌在收缩或舒张期的位移、应变、应变率等参数完成自动 LVEF 测量,因而具备极好的准确度^[14]。

近年来的研究表明,直接回归 EF 的 AI 模型能达到与人工测量相当的误差水平^[15]。其理论优势在于,全周期时空建模能够捕捉心脏运动的连续性特征,从而通过纹理和运动模式推断功能^[16, 17]。由于整个模型是联通的,训练过程中的所有参数都可以通过统一的目标函数进行优化。这使得模型能在训练时同时优化所有子任务的参数,从而提高整体性能。另外,其不再受限于 Simpson 法的几何假设,理论上能够学习到更复杂的非线性映射关系,可适用于心脏形态异常的患者。例如,Asch 等在 99 例成人中验证了万例超声心动图研究训练的模型,结果显示自动 EF 与专家手工 Simpson 法测量的参考 EF 相比平均 Bias 仅约 1.0%,95% LoA 为 $\pm 11.8\%$,与人工之间测量的 Bias (1.4%),95% LoA ($\pm 13.4\%$)相当,自动 EF 与人工测量平均 EF 的相关系数达到 $r=0.95$ ^[18]。

此外,Szjártó 等提出的 QUEST-EF 为端到端、分割无关的深度学习流水线,仅基于单一 A4C 视频

即可同时预测 3DE 标注的左心和右心 EF;在外部多中心测试中,其 LVEF 预测的平均绝对误差约为 4.60 个百分点(95% CI 4.41~4.80)^[19]。端到端的结构可以在同一个模型中实现多任务学习或多个视图的学习。通过共享信息和联合优化,模型可能比独立训练多个模块的传统方法获得更强的泛化能力和性能。目前的研究已经实现了多视图同时评估多项心脏参数指标^[20]。然而,端到端模型劣势在于中间处理步骤宛如一个黑匣子,并没办法了解到模型给出结果时是否基于准确的判断。这种“不可解释性”使得临床医生在依赖模型做决策时缺乏足够的信任度,尤其是在关键疾病诊断中。采用特征重要性分析、注意力机制或可视化方法,能够揭示模型的关键决策区域,从而部分弥补模型可解释性的不足。

本研究中使用的分步方法是一种较为符合临床测量 LVEF 流程的 AI 自动化结构,但也正是因为不同的模型之间会有各自的参数误差,使得最终的结果要劣于端到端回归的模型。因此,在临床应用中,如何平衡精度与可解释性,结合不同模型的优势,将是未来研究的关键方向。

【参考文献】

- [1] Ouyang D, He B, Ghorbani A, et al. Video-based AI for beat-to-beat assessment of cardiac function[J]. *Nature*, 2020, 580(7802): 252-256.
- [2] Alvé J, Hagberg E, Hagerman D, et al. A deep multi-stream model for robust prediction of left ventricular ejection fraction in 2D echocardiography[J]. *Sci Rep*, 2024, 14(1): 2104.
- [3] Lamoureux E, Ayromlou S, Ahmadi Amiri SN, et al. Segmenting cardiac ultrasound videos using self-supervised learning[C]//2023 45th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC). 2023, Sydney, Australia. IEEE, 2023: 1-7.
- [4] 中华医学会心血管病学分会,中国医师协会心血管内科医师分会,中国医师协会心力衰竭专业委员会,等.中国心力衰竭诊断和治疗指南 2024[J]. *中华心血管病杂志*, 2024, 52(3): 235-275.
- [5] Leclerc S, Smistad E, Pedrosa J, et al. Deep learning for segmentation using an open large-scale dataset in 2D echocardiography[J]. *IEEE Trans Med Imaging*, 2019, 38(9): 2198-2210.
- [6] Zhang J, Gajjala S, Agrawal P, et al. Fully automated echocardiogram interpretation in clinical practice: feasibility and diagnostic accuracy[J]. *Circulation*, 2018, 138(16): 1623-1635.
- [7] Liu F, Wang K, Liu D, et al. Deep pyramid local attention neural network for cardiac structure segmentation in two-dimensional echocardiography[J]. *Med Image Anal*, 2021, 67: 101873.
- [8] Jafari MH, Girgis H, van Woudenberg N, et al. Automatic biplane left ventricular ejection fraction estimation with mobile point-of-care ultrasound using multi-task learning and adversarial training[J]. *Int J Comput Assist Radiol Surg*, 2019, 14(6): 1027-1037.