Vision-Language Model-Based Automated Image Quality Assessment of LGE-MRI for Cardiac Ablation Planning

Bipasha kundu, Cristian A. Linte

Rochester Institute of Technology
Rochester, NY, USA

Aims: This study aims to evaluate a two-stage vision-language model (VLM) framework for clinically grounded image quality assessment (IQA) of left atrial LGE-MRI to support ablation planning in atrial fibrillation patients. Specifically, it explores VLMs for generating structured radiology-style quality reports across five criteria, and a GPT-based reasoning module to map predicted reports to a binary clinical usability decision.

Methods: We curated 60 annotated MRI slices from 20 patients using the LAScarQS 2022 dataset with expert radiologist annotations across five criteria - Noise, Motion Artifact, LA Boundary Accuracy, Pulmonary Vein Region Accuracy, and Under-segmentation Severity - each scored on a 0–3 ordinal scale (unusable, poor, acceptable, good). In Stage 1, three state-of- the-art VLMs (InternVL2-2B, DeepSeek-VL-1.3B, and LLaVA-1.5-7B) were fine-tuned using LoRA (Rank 8) to generate radiology-style quality reports. In Stage 2, a GPT-based reasoning module parsed the free-text reports into structured scores and a binary clinical usability decision. Models were evaluated using accuracy (ACC), Pearson Linear Correlation Coefficient (PLCC), F1 score, and Cohen's kappa Coefficient (k).

Results: InternVL2 achieved the highest performance (ACC=0.65,PLCC=0.79) across all five criteria. DeepSeek demonstrated competitive criterion-level accuracy (ACC=0.60,PLCC=0.78) and achieved perfect clinical usability agreement (Acc=1.00,F1=1.00,k=1.00) on the binary ablation planning decision, demonstrating robust end-to-end reasoning even with imperfect criterion-level predictions. InternVL2 followed with strong usability performance (Acc=0.92,k=0.83). No model produced a two-level ordinal prediction error, confirming clinical safety of predictions across all architectures.

Conclusion: This work presents the first VLM-based framework for clinically grounded, multi-criteria Image Quality Assessment of LGE-MRI in the context of cardiac ablation planning. The two-stage design effectively separates visual perception from structured clinical reasoning, producing interpretable quality decisions that align well with expert radiologist judgment. These results establish proof-of-concept feasibility and motivate future validation on larger dataset for generalizability.