

Echo-SAM: Zero-Shot Learning of Unseen Structures via Medical Knowledge Graph Grounding

Jiewen Yang · Xiaomeng Li*

The Hong Kong University of Science and Technology

01 Motivation & protocol

Many cardiac structures lack expert-drawn masks. Ultrasound appearance can be ambiguous; anatomical relationships provide spatial clues.

Where is the target relative to structures we can already see?

Seen in training

LA RA RV

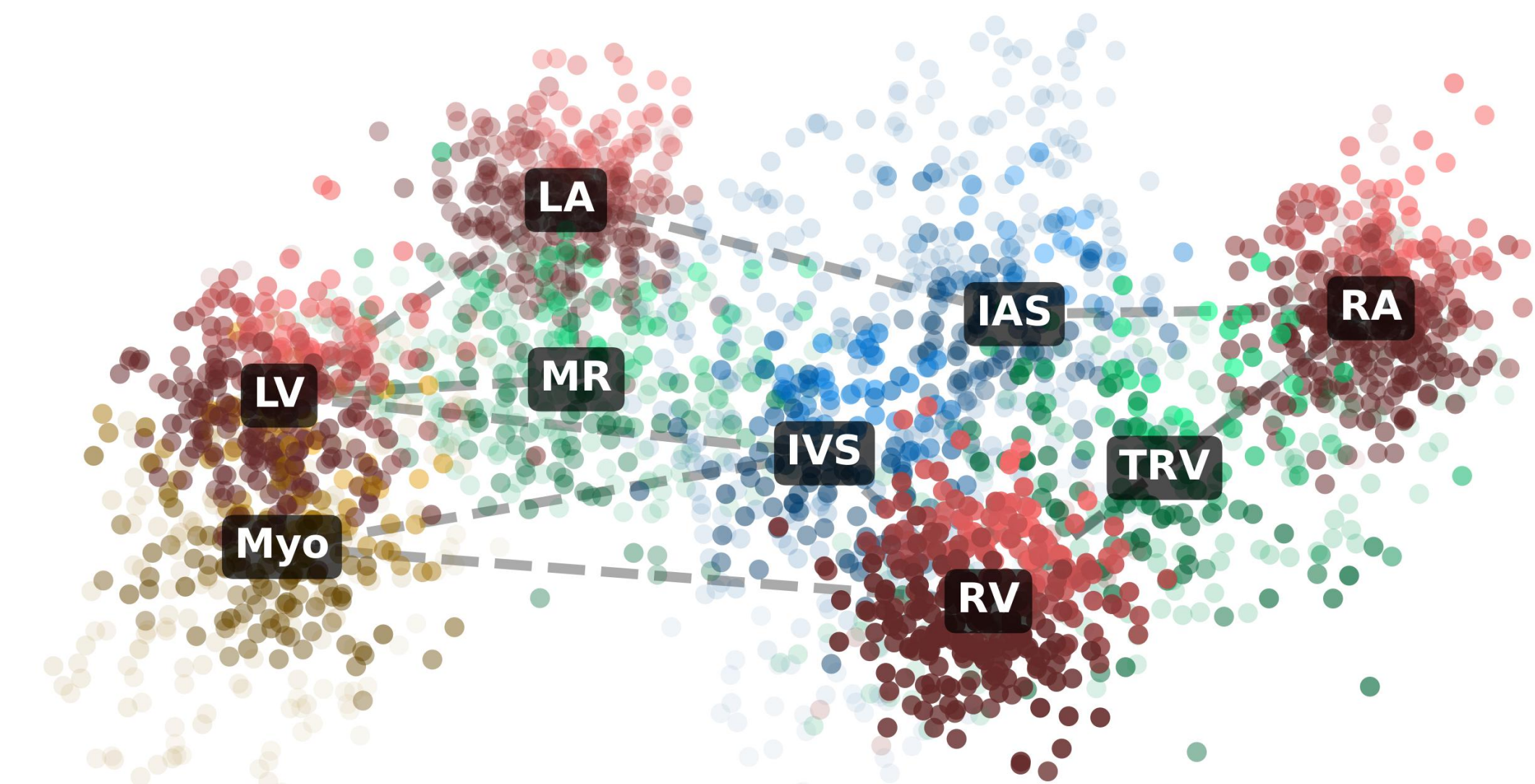
Unseen targets

LV Myocardium

Train on CardiacUDA using left/right atrial and right ventricular annotations. Unseen-target labels are excluded from training and model selection.

02 Anatomy as a graph

15 concepts · 42 anatomical relations



Learned cardiac representations in the four-chamber view.

03 Generality & evidence

One formulation for chambers, walls, septa, valves and great vessels.

In principle, any cardiac structure can be represented through valid, observable anatomical dependencies.

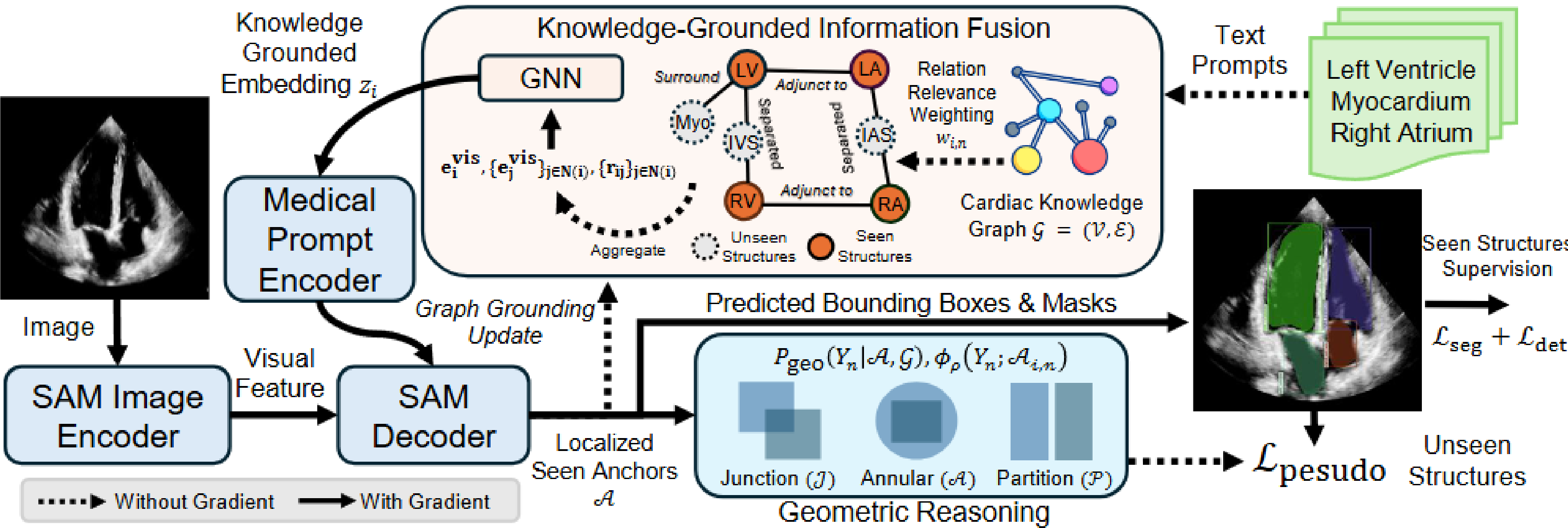
CURRENT QUANTITATIVE EVIDENCE

LV and myocardium on five public datasets, primarily in the four-chamber view. Evaluation coverage follows available images and reference masks.

Next: temporal reasoning and pathology-conditioned relations.

04 Anatomy → support → segmentation

Image-specific anchors connect a cardiac graph to SAM3 (ViT-L/14).



RELATION-CONDITIONED FUSION

A GNN combines image evidence with node and edge embeddings to enhance target prompts and weight anatomical relations.

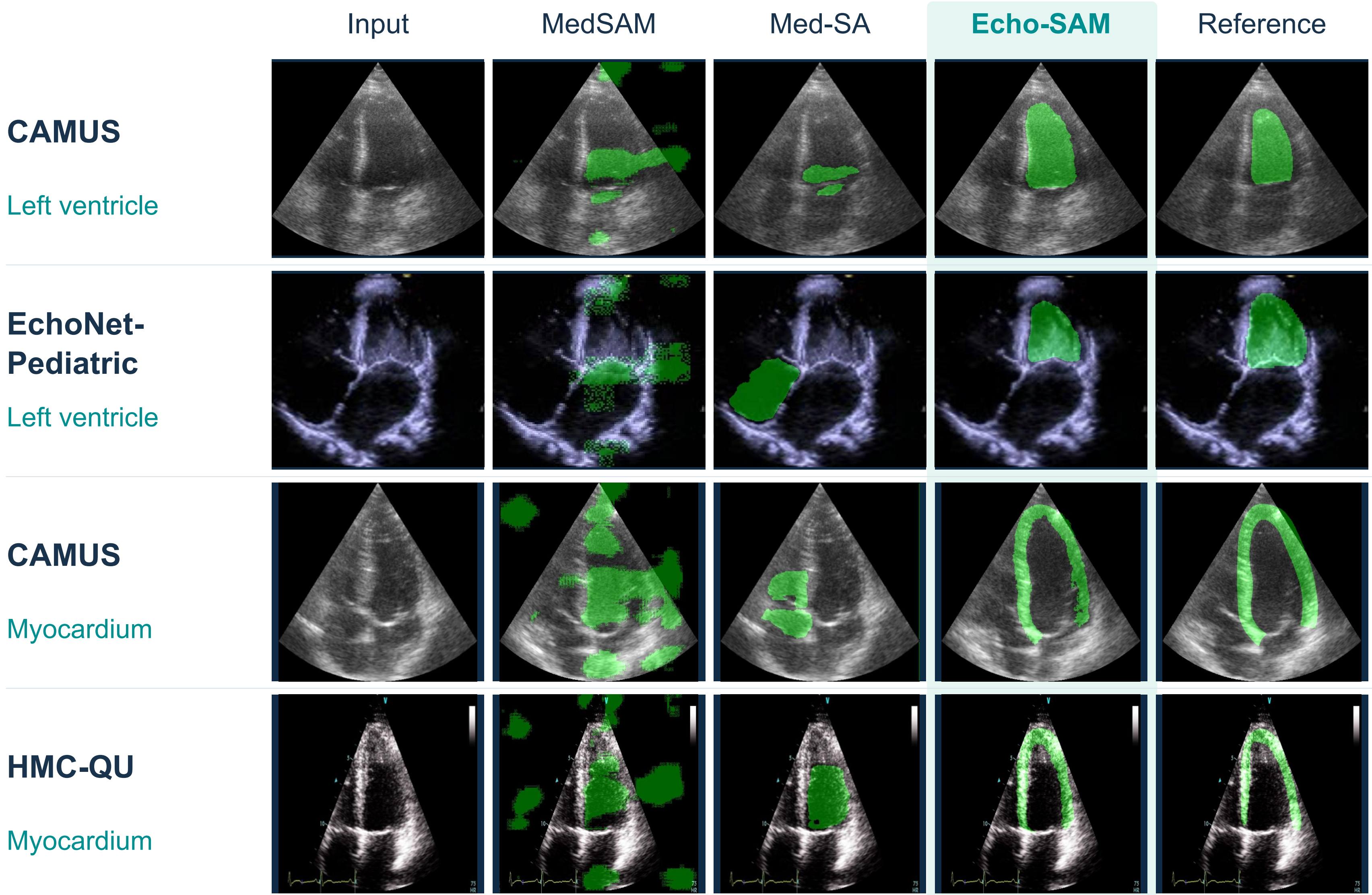
LATENT GEOMETRIC INFERENCE

Candidate supports adapt to the anchors through latent scale, thickness and offset. Posterior-guided scores select plausible target support.

Training: seen-label supervision + graph-derived pseudo targets

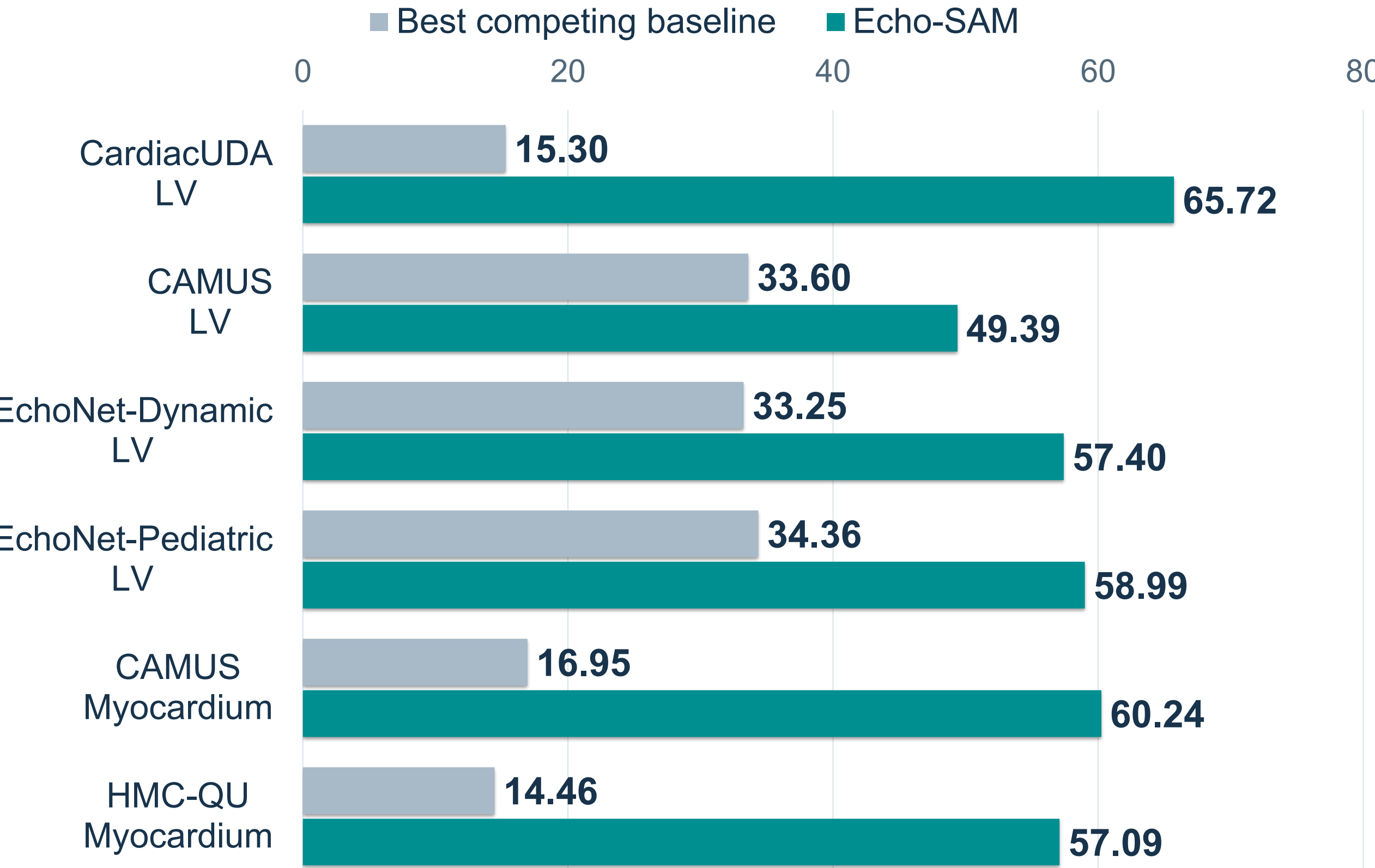
05 Qualitative segmentation across datasets

Representative predictions from the paper; green denotes the target.



Unseen chambers and thin myocardial walls are localized from anatomical context.

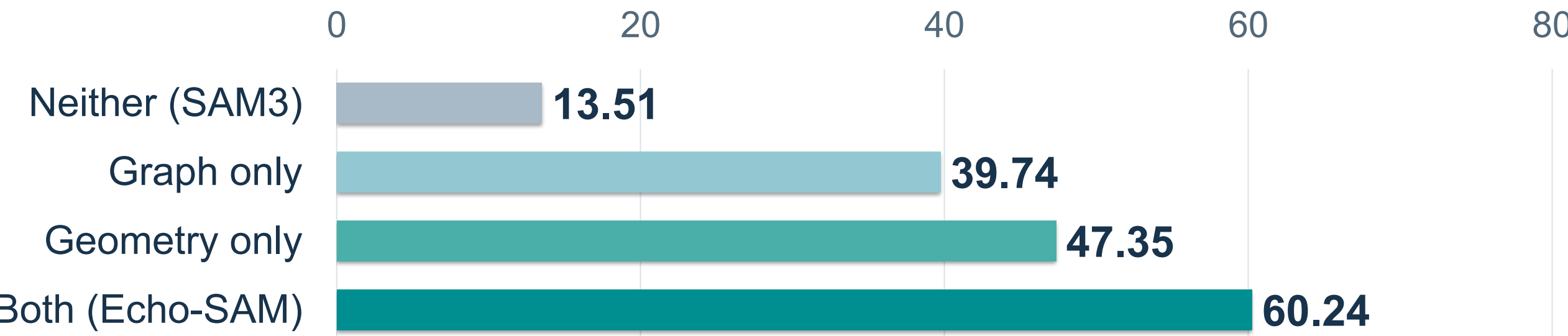
06 Zero-shot performance



Best baseline, top to bottom: MedSAM, MedSAM, Med-SA, Med-SA, MedSAM, SAM3. Main-paper zero-shot protocol.

07 Both modules matter

Module ablation · CAMUS myocardium · Dice (%)



Graph grounding adds +12.89 Dice points to geometry-only inference.

08 Anchor robustness

Retained anchors at inference

Available	CAMUS LV	CAMUS Myo	HMC-QU Myo
3 anchors	49.39 ± 1.10	60.24 ± 2.87	57.09 ± 2.44
2 anchors	48.48 ± 1.37	56.52 ± 2.55	55.76 ± 2.81
1 anchor	47.69 ± 1.49	57.77 ± 2.04	54.19 ± 2.22
0 anchors	32.93 ± 5.66	39.74 ± 2.27	35.98 ± 2.70

Dice (%), mean ± SD over 3 independent runs.

Partial anchor loss preserves useful spatial evidence. Complete loss removes geometric support.

Further challenges: weak boundaries, truncated images and abnormal topology.