Online Resource 1

The Visual Computer

KORAL-Net: Ordered Volumetric Representation and Boundary-Preserving Reconstruction for Three-Phase 3D DCE-MRI Segmentation

Gaixing Wei, Dezhen Zeng, Da Zeng, Shenglin Qin

School of Big Data and Artificial Intelligence, Nanning College of Technology, Nanning, China
Corresponding author: Dezhen Zeng; 984374899@qq.com

This online resource provides the complete ablation ladder, seed-level estimates, paired effects, engineering records, case provenance, integrity audit, reproducibility checklist, and extended qualitative evidence.

Supplementary Table S1 | Cohorts, preprocessing, optimisation, and reproducibility

Panel A. Cohort partitions and evidence roles				
Cohort	Cases	Role	Input	Label
I-SPY2 training	786	Parameter learning	Pre/early/late 3D DCE-MRI	Binary tumour mask
I-SPY2 development validation	98	Model selection	Same	Same
I-SPY2 internal test	98	Internal evaluation and repeated-seed analysis	Same	Same
I-SPY1 cross-cohort transfer	172	Cross-cohort evaluation	Audited three-phase geometry	Binary tumour mask
Panel B. Preprocessing, optimization, and reproducibility				
Item	Recorded value			
Target spacing	1.0 x 1.0 x 1.0 mm			
Patch size	256 x 256 x 40 voxels			
Effective batch size	2			
Encoder channels	32, 64, 128, 256, 320, 320, 320			
Residual blocks per level	1, 3, 4, 6, 6, 6, 6			
Optimiser	SGD with Nesterov momentum 0.99			
Weight decay	3×10^{-5}			
Maximum learning rate	5×10^{-3} after five-epoch linear warm-up			
Decay	Polynomial, exponent 0.9			
Gradient clipping	Global norm 5			
Mixed precision	Enabled			
Iterations	250 training and 50 validation iterations per epoch			
Random seeds	2026, 2027 and 2028 for core models			
Checkpoint rule	Best development-validation checkpoint under the convergence-controlled budget			
Hardware	Two NVIDIA GeForce RTX 4090 GPUs; one model assigned per GPU during parallel runs			

Supplementary Table S2 | Ablation design, seed-level results, and paired effect estimates

Panel A. Prespecified ablation ladder							
ID	Spatial configuration	OKE	ESR	LPD	ARO	Falsifiable question	Status
A0	Plain nnU-Net v2				None	Framework baseline	Complete
C-RE	Residual-encoder control				None	How much arises from spatial capacity?	Complete
A1	Residual encoder	Yes	No	No	None	Does ordered input encoding add information?	Complete
A2	Residual encoder	Yes	Yes	No	None	Does reliability-conditioned skip routing help?	Complete
A3	Residual encoder	Yes	Yes	Yes	None	Does lesion-preserving decoding improve reconstruction?	Complete
A4	Residual encoder	Yes	Yes	Yes	Fixed λ=0.05	Does scheduling outperform a fixed recall weight?	Complete
A5	Residual encoder	Yes	Yes	Yes	Scheduled and capped	Is the original full candidate reproducible?	Complete; seed-sensitive
A6-N	Residual encoder	Yes	No	Yes	None	Does ARO change the final no-ESR topology?	Complete; three seeds and two tests
A6-F	Residual encoder	Yes	No	Yes	Fixed λ=0.05	Does scheduling outperform a fixed weight?	Complete; three seeds and two tests
A6 KORAL-Net	Residual encoder	Yes	No	Yes	Scheduled and capped	Does removing the unstable router improve reproducibility?	Complete; final stable configuration
Panel B. Seed-level Dice for repeated core evaluation							
Model	Seed	Internal I-SPY2	Cross-cohort I-SPY1				
nnU-Net v2	2026	0.8194	0.6785				
nnU-Net v2	2027	0.8204	0.6790				
nnU-Net v2	2028	0.8145	0.6772				
ESR-containing candidate	2026	0.8479	0.7006				
ESR-containing candidate	2027	0.7598	0.6426				
ESR-containing candidate	2028	0.8435	0.6955				
KORAL-Net	2026	0.8456	0.6997				
KORAL-Net	2027	0.8436	0.6980				
KORAL-Net	2028	0.8467	0.7017				
Panel C. KORAL-Net minus nnU-Net v2 paired effects after case-wise seed averaging							
Cohort	Metric	Mean difference (95% bootstrap CI)	Holm-adjusted P				
Internal I-SPY2	Dice	+0.0272 (+0.0200 to +0.0342)	3.47×10^{-10}				

Internal I-SPY2	IoU	+0.0416 (+0.0327 to +0.0502)	1.87×10^{-11}				
Internal I-SPY2	Precision	+0.0443 (+0.0342 to +0.0548)	1.61×10^{-11}				
Internal I-SPY2	Recall	+0.0029 (-0.0141 to +0.0168)	0.4634				
Internal I-SPY2	HD95 (mm)	-1.199 (-2.641 to +0.198)	4.37×10^{-5}				
Internal I-SPY2	ASD (mm)	-0.231 (-0.471 to -0.009)	9.61×10^{-7}				
Cross-cohort I-SPY1	Dice	+0.0216 (+0.0130 to +0.0304)	7.88×10^{-6}				
Cross-cohort I-SPY1	IoU	+0.0256 (+0.0166 to +0.0349)	1.60×10^{-6}				
Cross-cohort I-SPY1	Precision	+0.0455 (+0.0350 to +0.0564)	7.91×10^{-15}				
Cross-cohort I-SPY1	Recall	-0.0316 (-0.0421 to -0.0213)	5.41×10^{-9}				
Cross-cohort I-SPY1	HD95 (mm)	-0.936 (-2.923 to +1.297)	3.40×10^{-5}				
Cross-cohort I-SPY1	ASD (mm)	-0.669 (-1.522 to +0.075)	2.56×10^{-6}				

Supplementary Table S3 | Efficiency, case provenance, data integrity, and reproducibility

Panel A. Archived operational measurements under the recorded workstation configuration					
Model	Unique trainable parameters (M)	Median logged epoch time (s)	Peak training allocation (GiB)	Archived prediction + export (s/case)	Compact checkpoint (MiB)
nnU-Net v2	44.70	156.70	Not recorded comparably	0.594	Not recorded
Residual spatial control	140.74	134.35	Not recorded comparably	Not timed	537.19
ESR-containing candidate	Not uniquely profiled	185.18	Not recorded comparably	1.483	537.15
KORAL-Net	141.68	168.37	18.60	1.433	540.65
Panel B. Qualitative-case provenance and display strata					
Cohort	nnU-Net identifier	Source case identifier	Target voxels	KORAL Dice	Baseline Dice
Internal	IT_0076	ACRIN-6698-530095_spy2_vis1	6,462	0.9346	0.8847
Internal	IT_0001	ISPY2-209397_spy2_vis1	38,314	0.9305	0.8931
Internal	IT_0010	ISPY2-366120_spy2_vis1	52,213	0.9599	0.9349
Cross-cohort	EX1_0065	ISPY1_1204_spy1_vis1	4,212	0.9495	0.9014
Cross-cohort	EX1_0117	ISPY1_1075_spy1_vis1	10,159	0.8901	0.8337
Cross-cohort	EX1_0115	ISPY1_1099_spy1_vis1	20,315	0.9036	0.8097
Panel C. Dataset and split integrity audit					
Audit item	Result				
Total examinations audited	1,154				
Training/development/internal/cross-cohort counts	786/98/98/172				
Cross-split identifier overlaps	0				
Phase-to-label shape mismatches	0				

Axis-code mismatches	0				
Affine discrepancies greater than 10^{-3}	0				
Maximum recorded affine error	0.0				
AMBL cases audited	52				
AMBL empty masks	16				
AMBL non-empty mask volume, median (range), mL	2.950 (0.040-36.588)				
Panel D. Reproducibility and reporting checklist					
Requirement	Location/status				
Fixed cohort manifest and patient-level overlap audit	research/audits/koral_dataset_metadata_*				
Architecture and module definitions	Figures 1 and 2a-d; editable draw.io files				
Random seeds and checkpoint rule	Section 2.7 and Supplementary Table S1				
Per-case metrics and seed summaries	research/statistics/kep_multiseed_*				
Planned paired statistics and correction family	Section 2.7 and the analysis manifest				
Foreground-error burden	Table 4; research/statistics/fp_burden/				
Efficiency scope and timing records	Table 6 and Supplementary Table S4; research/statistics/efficiency/				
Qualitative source data and crop/window settings	Supplementary Table S3, panel B and figure manifests				
Failed/invalid comparison runs	Reported explicitly; excluded from quantitative ranking				
Unavailable NME, motion and image-quality labels	Reported as not estimable				
Author, ethics, funding and repository metadata	Author confirmation form supplied separately				

Supplementary Table S4 | Standardized common-patch inference complexity

Model	Technical family	Parameters (M)	GFLOPs	Peak VRAM (GiB)	Latency (ms)	Input
SegResNet	Residual convolution	4.70	147.07	0.30	13.77	1 × 3 × 128 × 128 × 64
SwinUNETR	Windowed transformer	15.70	186.40	1.74	37.46	1 × 3 × 128 × 128 × 64
MedNeXt-S	Large-kernel ConvNeXt	5.55	126.62	0.89	26.98	1 × 3 × 128 × 128 × 64
3D UX-Net	Large-kernel convolution	53.04	1382.63	1.09	52.28	1 × 3 × 128 × 128 × 64
nnU-Net v2	Self-configuring 3D U-Net	44.70	491.27	1.50	11.02	1 × 3 × 128 × 128 × 64
KORAL-Net	Ordered representation + bounded reconstruction	141.68	772.72	1.80	30.71	1 × 3 × 128 × 128 × 64

Supplementary Table S4 note. Batch one, fp16, one RTX 4090, three warm-ups and 20 timed runs. FLOPs count Conv3d, ConvTranspose3d, and Linear operators at two FLOPs per multiply-add. The archived runtime was PyTorch 2.3 for the four comparator models and PyTorch 2.5 for nnU-Net v2 and KORAL-Net.

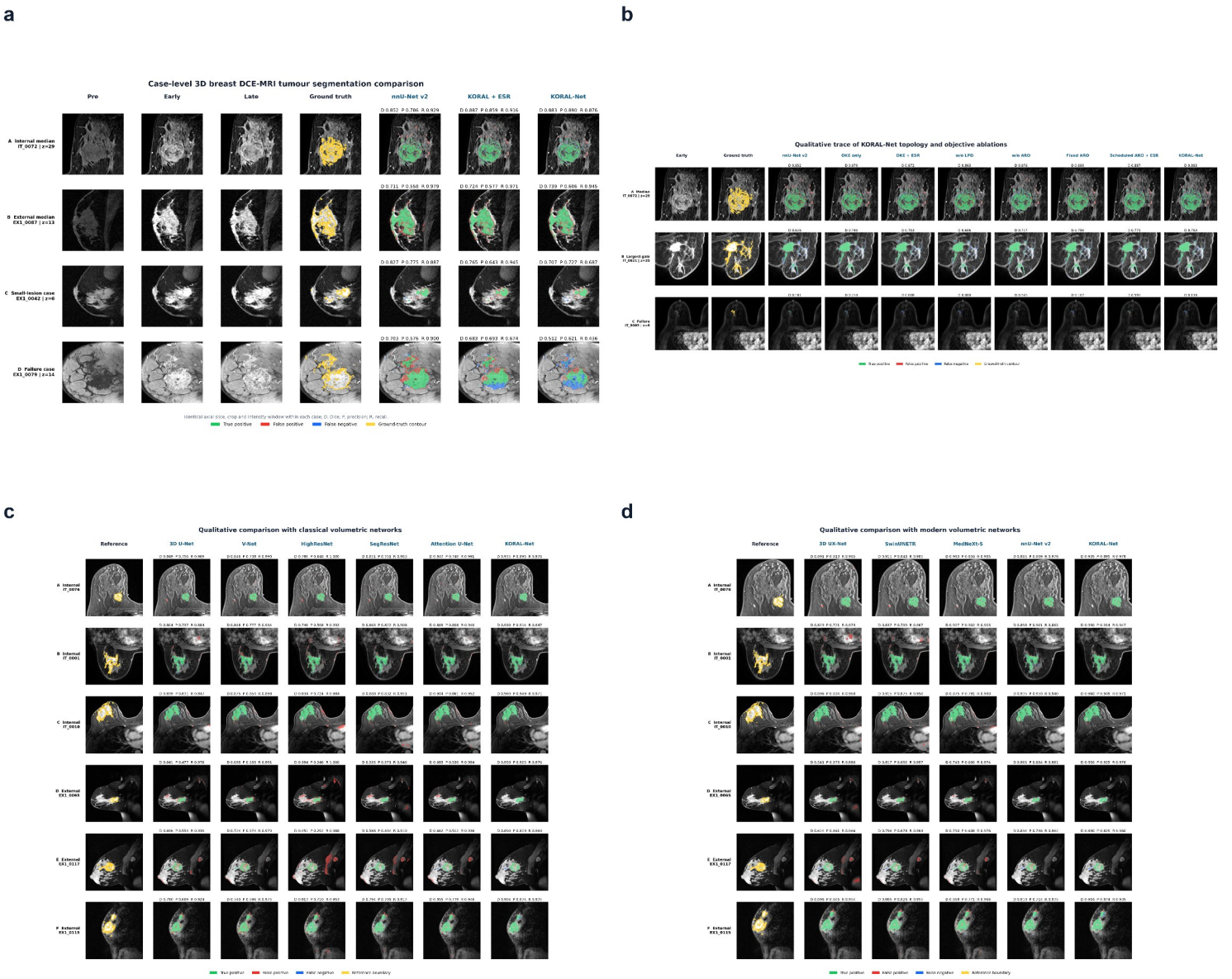

Fig. S1 Extended qualitative evidence. a, Balanced and failure audit. b, Case-level ablation matrix. c, Classical 3D segmenters. d, Modern 3D segmenters