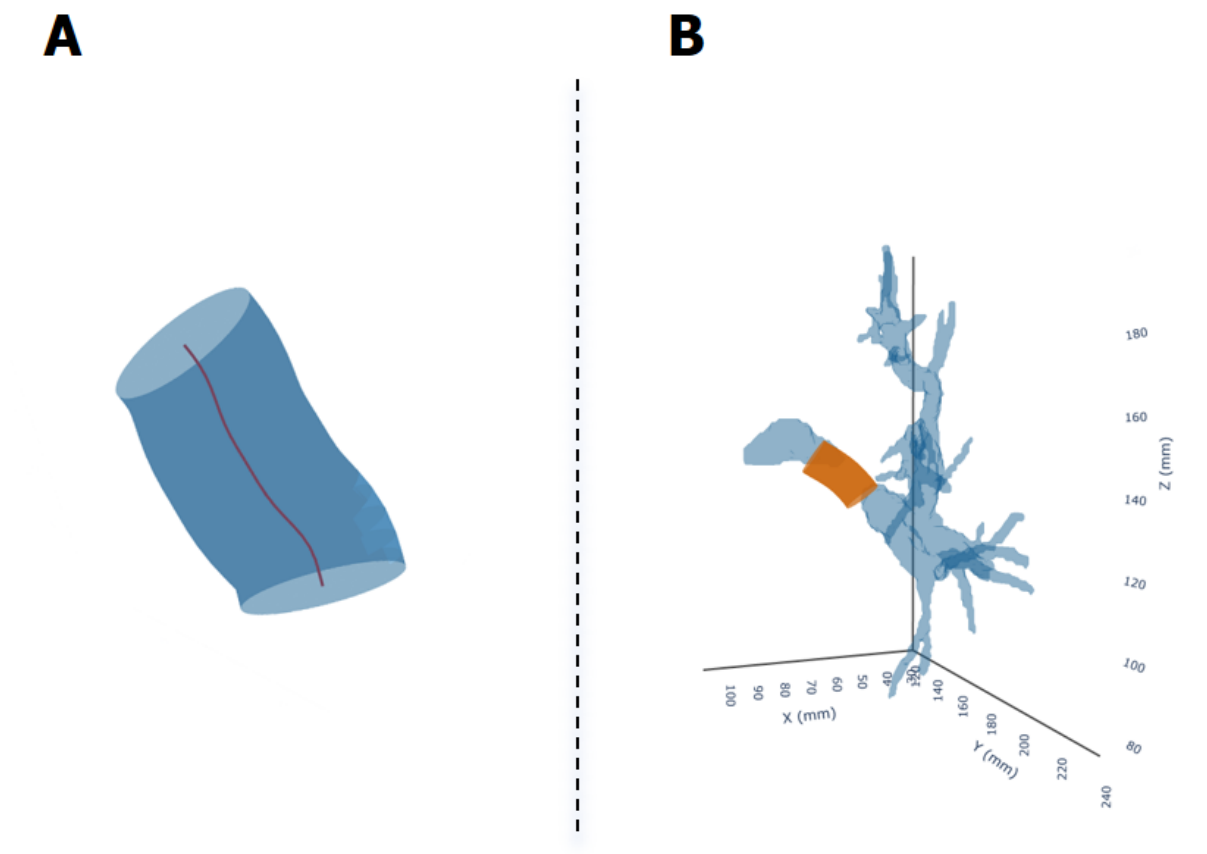
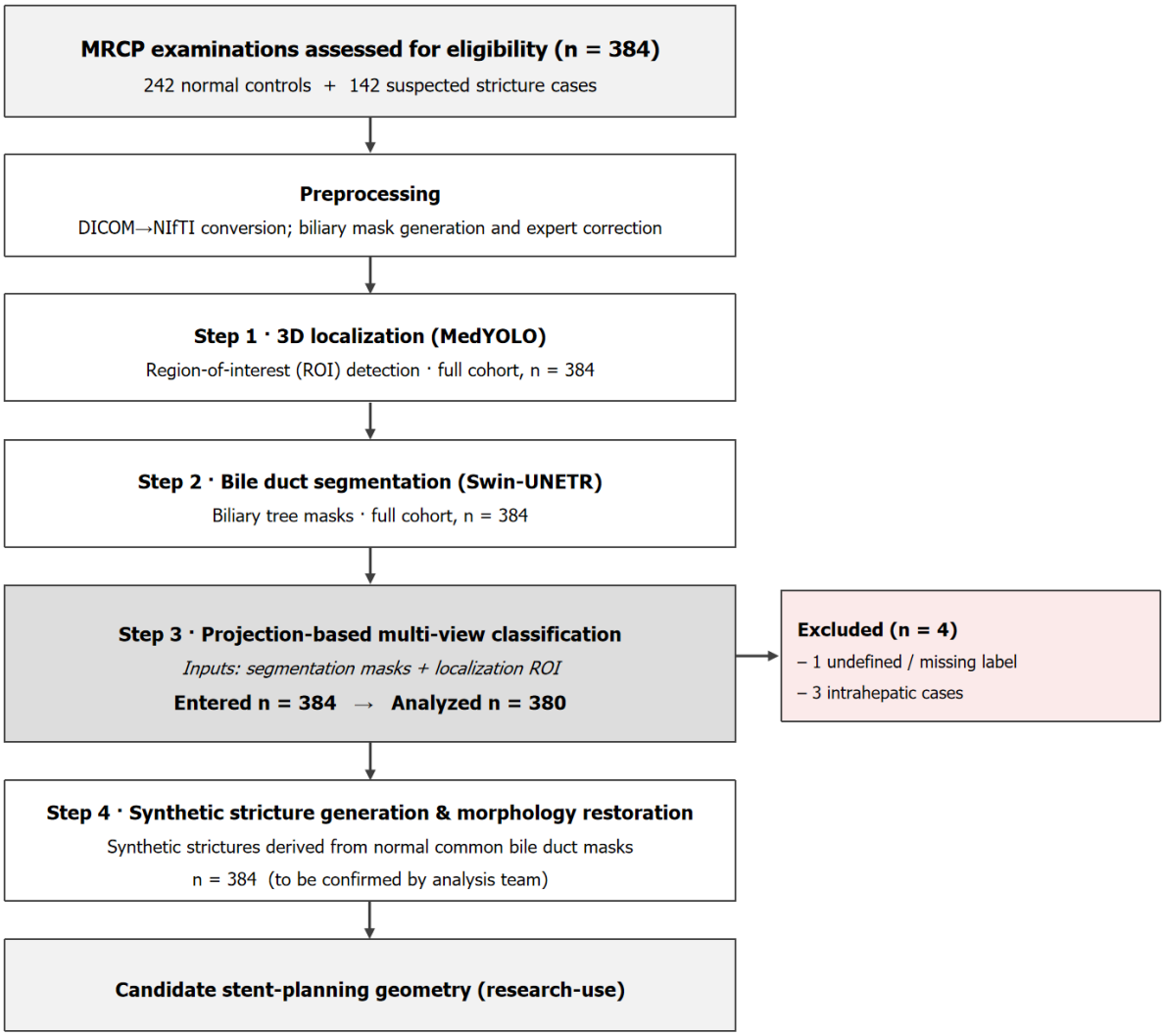
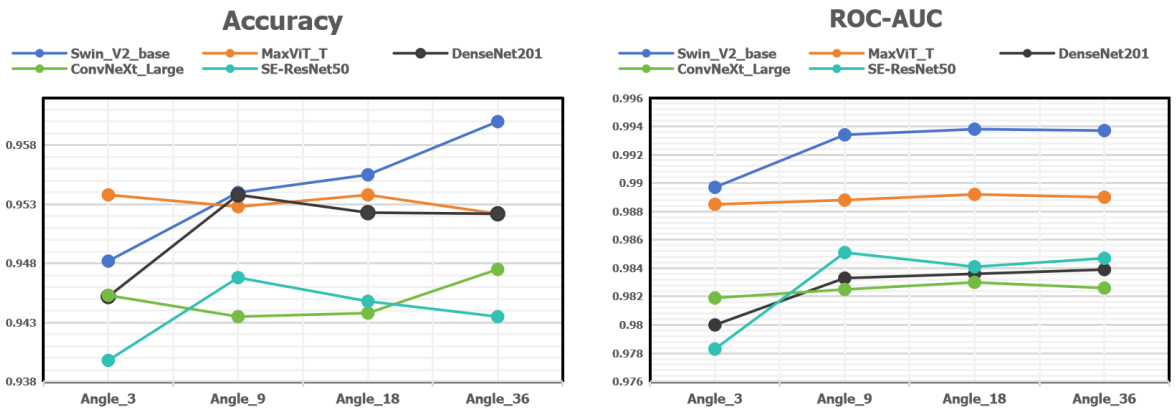


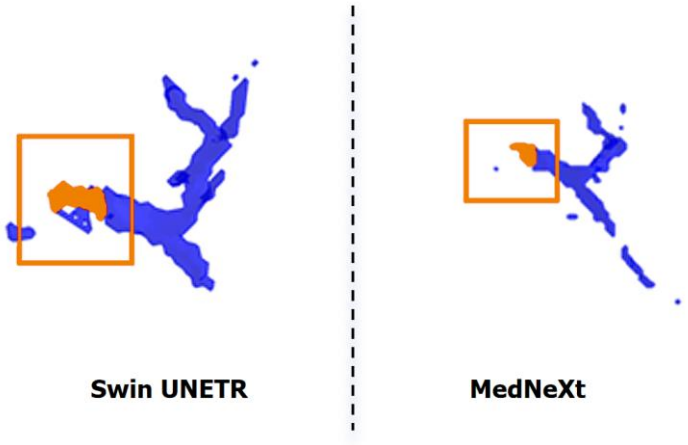
Supplementary Figure 1 Centerline-guided candidate stent-planning geometry.
A: Tubular mesh generated along the restored centerline; B: Candidate mesh overlaid on patient-specific biliary anatomy.



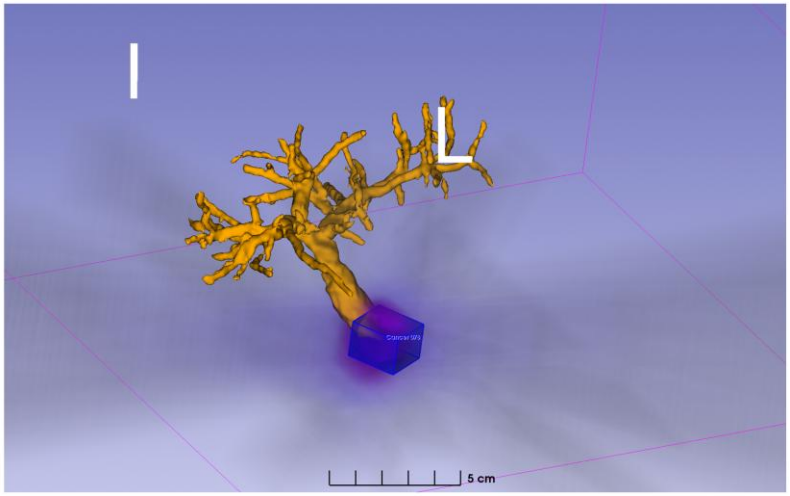
Supplementary Figure 2 Flow of magnetic resonance cholangiopancreatography examinations through the sequential analytical pipeline (STARD/STROBE-style inclusion and exclusion). The four analytical modules form a connected pipeline in which projection-based classification uses the bile-duct segmentation masks and the localization region-of-interest as inputs. For each module — three-dimensional localization (MedYOLO), bile duct segmentation (Swin-UNETR), projection-based classification, and morphology restoration — the number of examinations entered, the number and reason for any exclusions, and the number analyzed are shown. Of the 384 examinations (242 controls and 142 stricture cases), the projection-based classification step excluded one undefined/missing-label case and three intrahepatic cases, yielding 380 analyzed examinations; the remaining four intrahepatic and all 29 perihilar strictures were retained.



Supplementary Figure 3 Classification performance by projection view count and image-classification backbone. Accuracy and ROC AUC are shown for increasing projection angles across model architectures.

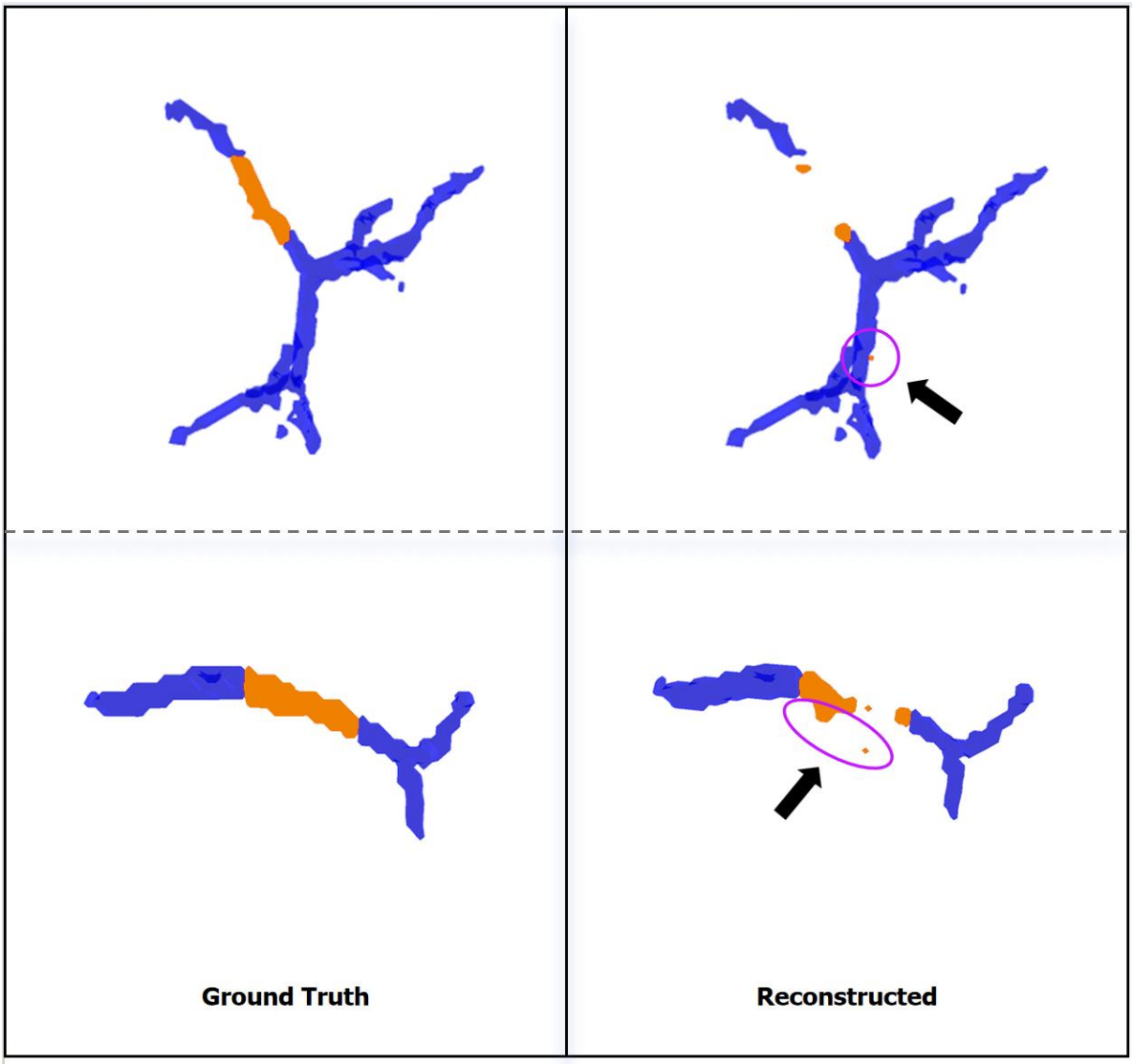


Supplementary Figure 4 Model comparison for synthetic stricture restoration. In the representative case, Swin-UNETR preserved the elongated ductal target better than MedNeXt.



Supplementary Figure 5 Multi-view projection-support visualization for

morphology restoration. The magenta-blue support map visualizes the spatial accumulation of back-projected evidence across projection views around the manually annotated bounding-box region. The map reflects cross-view geometric consistency and should not be interpreted as calibrated generative confidence.



Supplementary Figure 6 Representative failure cases in morphology restoration. The upper and lower rows show two representative cases, with ground truth on the left and reconstructed output on the right. Circles and arrows identify discontinuous restoration and isolated false-positive stenotic fragments. Orange regions denote target or generated stenotic segments, and reconstructed regions are model-generated hypotheses rather than observed anatomy.

Supplementary Table 1 Restoration Dice score according to lesion weight

Lesion weights	Dice score
3	0.5691
5	0.6597
7	0.6804
9	0.7245

Supplementary Table 2 Restoration Dice score according to training iteration

Training iteration	Dice score
30000	0.6957
150000	0.7282
300000	0.7508

Supplementary Table 3 Restoration Dice score according to loss function

Loss function	Dice score
Weighted Dice CE	0.6532
DiceFocalLoss	0.7531
TverskyLoss	0.6880
Generalized Dice Loss	0.4537
Top-K% Loss	0.5796

Supplementary Table 4 Distance-based and centerline-based tubular morphology performance

Model/loss setting	Dice (95%CI)	HD95 (95%CI)	ASSD (95%CI)	clDice (95%CI)	Mean centerline distance (95%CI)
Center loss, lesion weight 9	0.6495 (0.5979-0.7011)	9.6683 (4.6829-14.6537)	1.5282 (0.8770-2.1795)	0.6897 (0.6150-0.7644)	2.4994 (1.6672-3.3316)
Center loss, lesion weight 9 with bootstrap	0.6757 (0.6125-0.7388)	15.2185 (3.4069-27.0300)	2.0189 (0.9719-3.0659)	0.6851 (0.6116-0.7586)	2.9600 (1.9469-3.9731)

sampling						
DiceFocal loss	0.5937 (0.5192-0.6683)	12.5000 (8.0451-16.9549)	2.3131 (1.2280-3.3983)	0.6411 (0.5576-0.7245)	2.9487 (2.0717-3.8257)	
Generalized Dice loss	0.6213 (0.5712-0.6713)	18.9286 (6.0245-31.8328)	2.7559 (1.1598-4.3520)	0.6496 (0.5727-0.7265)	3.7415 (1.8142-5.6688)	
Tversky loss	0.6105 (0.5589-0.6621)	11.2617 (6.1321-16.3912)	1.8668 (1.2527-2.4809)	0.6617 (0.5989-0.7244)	2.7728 (1.9347-3.6109)	
Weighted Dice cross-entropy	0.6741 (0.6062-0.7419)	23.0182 (2.7400-43.2963)	2.8408 (0.6422-5.0395)	0.6954 (0.6200-0.7708)	3.6783 (1.7648-5.5919)	

CI: Confidence interval; HD95: 95% Hausdorff distance; ASSD: Average symmetric surface distance; clDice: Centerline Dice.

Supplementary Table 5 Swin-UNETR segmentation performance by input resizing and window size

Resize	Window size	Dice score
(100, 200, 200)	(32, 96, 96)	0.7974
(134, 446, 446)	(32, 96, 96)	0.7356
(128, 384, 384)	(96, 96, 96)	0.7413
(100, 100, 100)	(96, 96, 96)	0.8030
(100, 200, 200)	(96, 96, 96)	0.8113

Dice score is the volumetric overlap score; bold indicates the best result.

Supplementary Table 6 Classification performance by projection weighting method using the SwinV2-Base backbone

Projection Method	Accuracy	Balanced Acc.	Macro F1	ROC AUC

Constant	0.97368	0.96757	0.97127	0.99423
Linear	0.97039	0.96499	0.96775	0.99404
Gaussian	0.97697	0.97211	0.97491	0.99526

Supplementary Table 7 Classification accuracy according to projection-plane combination and number of projection views

X/Y/Z	**3 view**	**9 view**	**18 view**	**36 view**
X	0.951	0.951	0.949	0.949
Y	0.950	0.953	0.958	0.955
Z	0.933	0.938	0.942	0.951
XY	0.958	0.955	0.955	0.960
YZ	0.942	0.963	0.963	0.972
ZX	0.953	0.955	0.960	0.906
All planes	0.953	0.962	0.962	0.976