This MICCAI paper is the Open Access version, provided by the MICCAI Society. It is identical to the accepted version, except for the format and this watermark; the final published version is available on SpringerLink.

Topology-Aware Staged Diffusion for Multi-Vendor Retinal OCT Fluid Segmentation

Wa'ad Awamleh[1], Hrvoje Bogunović[2], Botond Fazekas[2]

[1] Institute for Machine Learning,
Johannes Kepler University Linz, Austria
[2] Christian Doppler Laboratory for Artificial Intelligence in Retina, Institute of Artificial Intelligence, Center for Medical Data Science, Medical University of Vienna, Austria

Abstract. Accurate segmentation of retinal fluids, i.e., intraretinal fluid (IRF), subretinal fluid (SRF), and pigment epithelial detachment (PED), from optical coherence tomography (OCT) is essential for monitoring exudative retinal diseases and guiding anti-VEGF therapy. Conventional deep learning methods often yield topologically fragmented predictions that are clinically implausible. We benchmark UniSegDiff, a published staged conditional diffusion model, on multi-vendor retinal OCT fluid segmentation and integrate a connectivity prior loss inspired by LTSeg to enforce topological coherence during training. to enforce topological coherence during training. On the RETOUCH benchmark (112 volumes: 70 training, 42 test; three vendors), our model achieves a mean Dice of 0.780, outperforming CFENet (0.718), nnUNet (0.716), and base UniSegDiff without the prior (0.702). The connectivity prior yields the largest improvement for SRF (+0.126 DSC), the fluid type most prone to fragmentation, and reduces mean absolute volume difference (AVD) by 57% over CFENet. Zero-shot cross-dataset evaluation on the AROI dataset demonstrates moderate generalization (overall Generalized Dice 0.704), with a clear fluid-type-dependent transferability hierarchy. Our results show that combining diffusion-based iterative refinement with explicit topological priors produces segmentations that are both quantitatively accurate and structurally plausible.

Keywords: Retina · Optical Coherence Tomography · Image segmentation · Diffusion models · Topological priors

1 Introduction

Retinal diseases such as age-related macular degeneration (AMD) are among the leading causes of irreversible vision loss worldwide [18]. Optical coherence tomography (OCT) enables clinicians to visualize and quantify pathological fluid accumulations, including intraretinal fluid (IRF), subretinal fluid (SRF), and pigment epithelial detachment (PED), as key biomarkers for anti-VEGF treatment

decisions [8,13]. Deep learning methods [12,9] have achieved strong segmentation results, but frequently yield topologically fragmented predictions, in the form of disconnected components and spurious regions that are clinically implausible and corrupt volume estimates. These artifacts are especially problematic for thin structures such as SRF, where pixel-wise losses fail to enforce spatial coherence.

A second practical challenge is multi-vendor deployment. OCT devices differ in resolution, noise characteristics, and contrast, so a clinically useful method must generalize across vendors without per-device retraining. The RETOUCH challenge [2] provides a controlled multi-vendor benchmark for this setting.

We choose UniSegDiff [7] as the starting architecture because its staged conditional diffusion framework offers a natural insertion point for topological constraints: by Stage 3 the predicted mask is already refined, so a connectivity prior can correct fragmentation without destabilizing early training. UniSegDiff was originally proposed for unified lesion segmentation across multiple organs and modalities, but it has not been evaluated on multi-vendor retinal OCT and does not enforce topological coherence. We emphasize that UniSegDiff is an *existing* architecture [7]; our contribution lies in the domain adaptation and the topological integration described below.

In this work, we make the following contributions:

1. We **adapt and benchmark** UniSegDiff on the RETOUCH multi-vendor retinal OCT fluid segmentation benchmark.
2. We **integrate a connectivity prior loss**, inspired by LTSeg [20], into the final stage of diffusion training to enforce topological coherence.
3. We conduct **systematic ablations** isolating the effects of the connectivity prior and post-processing.
4. We perform **zero-shot cross-dataset evaluation** on the AROI database [11], including a strict cross-vendor setting with the target device excluded from training.

2 Related Work

Retinal OCT fluid segmentation. The RETOUCH challenge [2] established a multi-vendor benchmark for IRF/SRF/PED segmentation. Subsequent methods include RetFluidSeg [10], ReLayNet [14]. nnUNet [9] serves as a strong self-configuring baseline but does not incorporate structural priors.

Diffusion models for segmentation. Diffusion models [5] have been applied to medical segmentation in single-stage [19] and iterative [1] settings. UniSegDiff [7] addresses uneven attention across diffusion timesteps by dynamically adjusting prediction targets at different training stages. Its architecture consists of a frozen Conditional Feature Extraction Network (CFENet), a standard U-Net pre-trained on segmentation, and a trainable Denoising Network (DNet) with dual decoders that predict the clean mask and noise respectively. CFENet injects multi-scale features into DNet via Dual Cross-Attention, bridging the modality gap between raw images and noisy masks; we augment this framework with a topological prior.

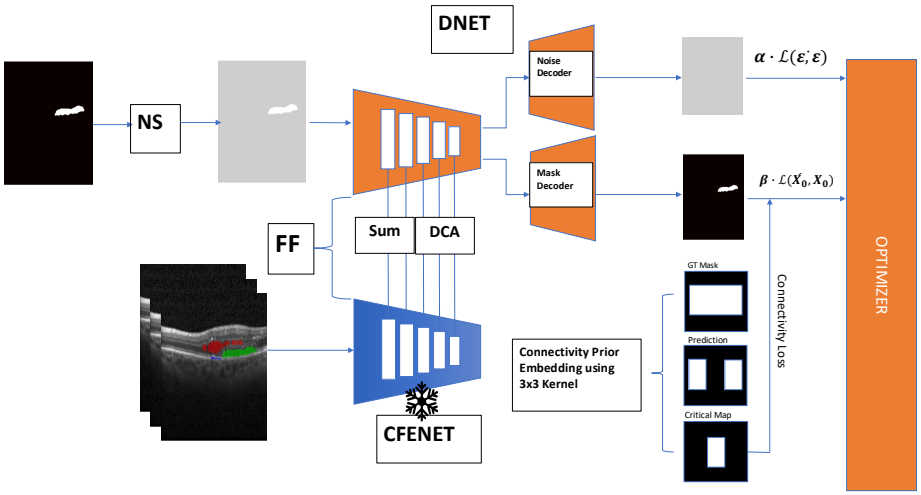

Fig. 1. Overview of our adapted UniSegDiff pipeline. The frozen CFENet provides multi-scale features to the trainable DNet via Feature Summation (levels 1–2) and Dual Cross-Attention (levels 3–5). The connectivity prior loss is applied only at Stage 3 (t=0), where predictions are refined enough for meaningful topological constraints. NS: Noise Schedule, FF: Features Fusion, DCA: Dual Cross-Attention, CFENET: Conditional Feature Extraction Net.

Topology-aware segmentation. Persistent homology losses [6,3] enforce topological constraints but are computationally expensive. clDice [15] targets tubular structures via skeleton-based metrics. LTSeg [20] proposed a lightweight connectivity prior for rib segmentation: it dilates the predicted binary mask, identifies ground-truth pixels that fall in the dilated region but not the prediction itself (i.e., gaps), and penalizes these critical pixels to encourage connected structures. TopoDiffusionNet [4] integrated persistent homology into the diffusion denoising process to enforce topological constraints during generation. We adapt the LTSeg connectivity prior to retinal fluids within a staged diffusion segmentation framework.

3 Methods

3.1 Architecture

As illustrated in Fig. 1, our pipeline follows UniSegDiff [7]: a frozen CFENet (5-level U-Net, 56.6M params) provides multi-scale conditional features CF_i (i=1∼5) to a trainable DNet (58.5M params) with dual decoders for mask and noise prediction. DNet fuses these features via Feature Summation at levels 1–2 and Dual Cross-Attention (4 heads) at levels 3–5, with cross-attention dimensions reduced to fit within a single A100 GPU. Total: 115.1M parameters (56.6M frozen, 58.5M trainable).

3.2 Staged Training

The diffusion process uses 1,000 timesteps with a linear β schedule from 10^{-4} to 0.02. Stage boundaries are set at t=599 and t=299, dividing the timestep range into three parts. The mask loss combines BCE and Dice: $\mathcal{L}_{\text{mask}} = \mathcal{L}_{\text{BCE}} + \mathcal{L}_{\text{Dice}}$. The noise loss is $\mathcal{L}_{\text{noise}} = \text{MSE}(\hat{\epsilon}, \epsilon)$. Training has two phases: CFENet is first pre-trained standalone with Dice-Focal loss (30 epochs), then frozen while DNet is trained with the staged loss (100 epochs, $\sim$26 min/epoch on an A100 GPU).

3.3 Connectivity Prior Loss

A key contribution of this work is integrating a topological constraint into UniSegDiff, inspired by LTSeg [20]. For each fluid class, we compute a critical map C identifying ground-truth pixels adjacent to, but not covered by, the prediction:

$$C = \text{dilate}(\hat{m}) \odot (1 - \hat{m}) \odot m_0, \tag{1}$$

where $\odot$ denotes element-wise multiplication, $\hat{m}$ is the prediction binarized using a straight-through estimator (STE), m_0 the ground truth, and dilate$(\cdot)$ applies morphological dilation with a 3×3 all-ones structuring element, expanding the mask by one pixel in each direction to capture immediately adjacent regions. Intuitively, the loss focuses learning on gap pixels where the prediction breaks apart, forcing the network to "fill in" disconnected regions while the sparsity term prevents it from over-predicting to trivially satisfy connectivity:

$$\mathcal{L}_{\text{conn}} = \lambda_c \frac{\sum C \odot \text{BCE}(\hat{x}_0, m_0)}{\sum C + \epsilon} + \lambda_s \frac{\text{ReLU}(\text{vol}(\hat{m}) - \text{vol}(m_0))}{HW}, \tag{2}$$

with λ_c=0.1 and λ_s=0.5. The connectivity loss is applied only during training at the final timestep (t=0) of Stage 3, where predictions are sufficiently refined. The total Stage 3 loss is:

$$\mathcal{L}_{\text{S3}} = \alpha \mathcal{L}_{\text{noise}} + \beta \mathcal{L}_{\text{mask}} + \mathcal{L}_{\text{conn}}. \tag{3}$$

3.4 Inference

Inference follows a three-stage pipeline: (1) from pure noise x_T, the model produces an initial coarse segmentation; (2) this is re-noised to t=599 and iteratively refined via DDIM sampling to t=299; (3) each endpoint undergoes 10 MC dropout forward passes during t=299 to t=0, producing 20 candidate masks fused via STAPLE [17]. Optional morphological post-processing includes hole filling, small component removal (2D: <10 px, 3D: <30 vx), and argmax-based overlap resolution.

4 Experiments

4.1 Datasets

RETOUCH. The RETOUCH benchmark [2] contains 112 OCT volumes (70 train, 42 test) from three vendors: Cirrus (24/14), Spectralis (24/14), and Topcon (22/14). Voxel-level annotations for IRF, SRF, and PED were performed by expert consensus. The test set is evaluated via the official challenge server.

AROI. The AROI database [11] comprises 1,136 annotated B-scans from 24 AMD patients (Cirrus, Zagreb). We extract three fluid classes and use AROI exclusively for zero-shot evaluation without fine-tuning.

Preprocessing and augmentation. 3D MHD volumes are sliced into 2D B-scans, normalized to $[0, 1]$, and resized to 256×256 with aspect-ratio-preserving padding; masks use nearest-neighbor interpolation and 3-channel one-hot encoding. Augmentation follows Unterdechler et al. [16]: horizontal flipping, small rotations ($\pm10°$), brightness/contrast jitter, and Gaussian noise.

Baselines. We compare four configurations: (1) **CFENet**, the frozen feature extractor trained standalone (56.6M params); (2) **nnUNet**, a compact nnUNet-inspired baseline (11.2M params) [9]; (3) **UniSegDiff**, our adapted implementation without connectivity prior; (4) **UniSegDiff + CP**, with the connectivity prior at Stage 3. All use identical preprocessing, augmentation, and splits.

Metrics. On RETOUCH, we report per-vendor, per-class Dice Similarity Coefficient (DSC) and Absolute Volume Difference (AVD), computed via the official challenge server. On AROI, we report 2D per-slice Dice and 3D per-patient Generalized Dice (GD).

Experiments. We conduct five experiments: (1) RETOUCH benchmark comparison; (2) ablation of the connectivity prior; (3) ablation of post-processing; (4) zero-shot cross-dataset evaluation on AROI (all RETOUCH vendors in training); (5) strict cross-vendor evaluation on AROI (Cirrus excluded from training).

5 Results

5.1 RETOUCH Benchmark (Exp. 1)

Table 1 presents per-vendor per-class DSC and mean AVD. UniSegDiff + CP achieves the highest mean DSC (0.780), outperforming CFENet (0.718), nnUNet (0.716), and base UniSegDiff (0.702) across all vendors. The mean AVD of UniSegDiff + CP (0.046) represents a 57% reduction over CFENet (0.108) and 70% over nnUNet (0.156). Even base UniSegDiff achieves lower AVD (0.083) than both baselines despite lower DSC, suggesting iterative denoising inherently improves volumetric calibration.

Fig. 2 shows a qualitative comparison across all four models. CFENet, nnUNet, and UniSegDiff produce a spurious IRF prediction outside the retinal layers, which UniSegDiff + CP avoids.

Table 1. RETOUCH test set (42 volumes) DSC (↑) per vendor and fluid type, and mean AVD (↓) averaged across vendors, evaluated via the official challenge server. Best in **bold**.

	Cirrus DSC			Spectralis DSC			Topcon DSC			Mean	Mean
Model	IRF	SRF	PED	IRF	SRF	PED	IRF	SRF	PED	**DSC↑**	**AVD↓**
CFENet	.758	.751	.707	.731	.563	.761	.672	.817	.704	.718	.108
nnUNet	.745	.733	.762	.730	.523	.757	.672	.814	.703	.716	.156
UniSegDiff	.771	.624	.803	.717	.521	.741	.649	.799	.691	.702	.083
UniSegDiff+CP	**.828**	**.763**	**.872**	**.779**	**.661**	**.761**	**.723**	**.897**	**.741**	**.780**	**.046**

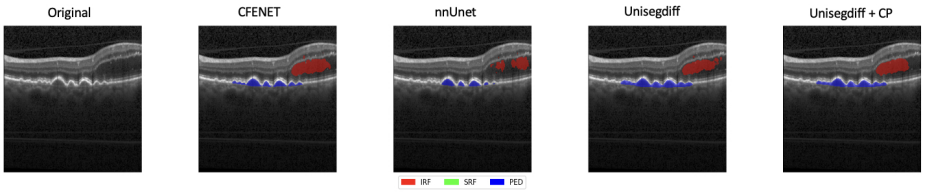

Fig. 2. Side-by-side model comparison on a RETOUCH B-scan. Left to right: OCT input, CFENet, nnUNet, UniSegDiff, UniSegDiff + CP. IRF (red), SRF (green), PED (blue). A spurious IRF prediction is visible in all models except UniSegDiff + CP. The ground-truth panel is omitted because the depicted OCT volume originates from the RETOUCH test set, for which reference annotations are not publicly available.

5.2 Connectivity Prior Ablation (Exp. 2)

Adding the connectivity prior improves mean DSC from 0.702 to 0.780 (+0.078). The largest per-class gain is SRF (+0.126), the fluid most prone to fragmentation due to its thin morphology.

Table 2 presents an error categorization on a held-out RETOUCH split (10 patients). The connectivity prior dramatically reduces topological fragmentation (SRF: 31.8%→12.6%; IRF: 21.4%→9.3%) while leaving under-segmentation rates largely unchanged, confirming it acts as a targeted topological intervention.

5.3 Post-Processing Ablation (Exp. 3)

Morphological post-processing (hole filling, small component removal) improves mean DSC from 0.712 to 0.780 (+0.068), with the largest gain for PED (+0.141). The connectivity prior and post-processing are complementary: the prior regularizes learned representations during training, while post-processing addresses residual artifacts at inference. Notably, raw UniSegDiff + CP predictions without post-processing (0.712) already match CFENet with post-processing (0.718).

5.4 Zero-Shot Cross-Dataset Evaluation (Exp. 4–5)

Table 3 presents cross-dataset results on AROI. When trained on all three RETOUCH vendors, UniSegDiff + CP achieves an overall GD of 0.704, substan-

Table 2. Error categorization on held-out RETOUCH split (10 patients, from the validation set). Under-seg. = FN/GT pixels; Topol. frag. = non-largest-component pixels / predicted pixels.

Model	Fluid	Under.	Over.	Frag.	Miss
CFENet	IRF	34.2	18.6	21.4	8.3
	SRF	49.7	12.1	31.8	14.2
	PED	38.5	15.3	11.2	17.6
Ours	IRF	32.8	16.1	9.3	5.1
	SRF	47.3	11.4	12.6	7.8
	PED	37.1	13.9	4.7	9.4

Table 3. Zero-shot evaluation on AROI (3D Generalized Dice). Exp. 4: all RETOUCH vendors in training. Exp. 5: Cirrus excluded. Best per experiment in **bold**.

Exp.	Model	IRF	SRF	PED	Overall
4	nnUNet	.001	**.461**	.003	.447
	Ours	**.703**	.454	**.326**	**.704**
5	nnUNet	.022	**.371**	.000	.468
	Ours	**.550**	.320	**.220**	**.450**

tially outperforming nnUNet (0.447). A clear transferability hierarchy emerges: IRF (0.703) > SRF (0.454) > PED (0.326), correlating with the morphological distinctiveness of each fluid type.

In the stricter cross-vendor setting (Exp. 5, Cirrus excluded from training), overall GD drops from 0.704 to 0.450 (−36% relative), confirming that vendor-matched training data is critical. nnUNet achieves marginally higher overall GD (0.468) but its performance is carried entirely by SRF, as IRF collapses to 0.022 and PED to 0.000, whereas UniSegDiff + CP retains non-trivial performance across all classes.

Fig. 3 shows a representative prediction with per-class probability maps and MC dropout uncertainty.

6 Discussion

The connectivity prior is the critical ingredient. Without it, UniSegDiff (0.702) underperforms discriminative baselines; the diffusion process alone does not produce topologically coherent predictions. Table 2 reveals the mechanism: fragmentation drops selectively (SRF: 31.8%→12.6%; IRF: 21.4%→9.3%) while under-segmentation rates remain largely unchanged, confirming a targeted topological intervention. Further, base UniSegDiff already achieves lower AVD (0.083) than CFENet (0.108) and nnUNet (0.156) despite lower DSC, showing that iterative denoising inherently calibrates volumes; with the prior, AVD drops to 0.046, within typical inter-grader variability [2]. The consistent multi-vendor performance

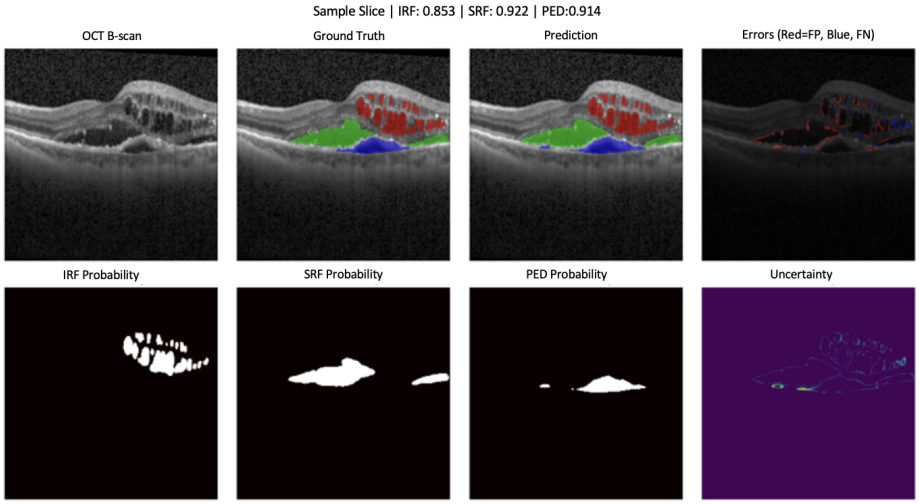

Fig. 3. UniSegDiff + CP prediction on a RETOUCH Spectralis B-scan. Top: input, ground truth, prediction, error map (red=FP, blue=FN). Bottom: per-class probabilities (IRF, SRF, PED) and MC dropout uncertainty. Per-class Dice: IRF = 0.853, SRF = 0.922, PED = 0.914.

(Cirrus: 0.821, Topcon: 0.787, Spectralis: 0.734) demonstrates vendor-agnostic representations suitable for multi-site deployment.

Cross-dataset generalization. The zero-shot hierarchy (IRF>SRF>PED) reflects morphological distinctiveness: IRF cysts are device-consistent, while SRF and PED boundaries are sensitive to acquisition protocol. Excluding Cirrus from training causes a 36% GD decline, confirming cross-vendor transfer is a compounding challenge. nnUNet's IRF collapses to near-zero GD in both settings, whereas UniSegDiff + CP retains 0.703 and 0.550, indicating more generalizable representations.

Limitations. UniSegDiff + CP is ∼200× slower than nnUNet at inference (∼2.4 s vs. ∼4 ms per slice) due to multi-stage DDIM sampling and MC dropout. This processing time precludes deployment in high-throughput routine clinical settings. However, in offline analysis environments such as clinical trials, real-time processing is not a prerequisite. In these contexts, the requirement for precise volumetric quantification and topological coherence supersedes processing speed, and the substantial reduction in AVD justifies the extended computational duration. To facilitate future routine clinical application, concrete acceleration strategies, including knowledge distillation, sampling step reduction, and early exiting, will be evaluated. All models operate on 2D B-scans, ignoring inter-slice context. Single-dataset training limits cross-institutional transfer, as evidenced by the moderate zero-shot AROI results, which may also reflect differences in annotation protocols between datasets.

7 Conclusion

We benchmarked UniSegDiff on multi-vendor retinal OCT fluid segmentation and showed that integrating a connectivity prior loss yields consistent improvements over both discriminative baselines and the base diffusion model across all vendors and fluid types. Diffusion models are not inherently superior without structural priors; their value lies in providing natural insertion points for topology-aware losses. Future work should address 3D modeling, efficient inference distillation, and domain adaptation for cross-institutional transfer.

References

1. Amit, T., Shaharbany, T., Nachmani, E., Wolf, L.: SegDiff: Image Segmentation with Diffusion Probabilistic Models. arXiv preprint arXiv:2112.00390 (2021)
2. Bogunović, H., Venhuizen, F., Klimscha, S., Apostolopoulos, S., Bab-Hadiashar, A., Bagci, U., Beg, M.F., Bekalo, L., Chen, Q., Ciller, C., et al.: RETOUCH – The Retinal OCT Fluid Detection and Segmentation Benchmark and Challenge. IEEE Transactions on Medical Imaging (2019). https://doi.org/10.1109/TMI.2019.2901398
3. Clough, J.R., Byrne, N., Oksuz, I., Zimmer, V.A., Schnabel, J.A., King, A.P.: A Topological Loss Function for Deep-Learning based Image Segmentation using Persistent Homology. IEEE Transactions on Pattern Analysis and Machine Intelligence **44**(12), 8766–8778 (2022). https://doi.org/10.1109/TPAMI.2020.3013679
4. Gupta, S., Samaras, D., Chen, C.: TopoDiffusionNet: A Topology-aware Diffusion Model. In: International Conference on Learning Representations (ICLR) (2025)
5. Ho, J., Jain, A., Abbeel, P.: Denoising Diffusion Probabilistic Models. In: Advances in Neural Information Processing Systems (NeurIPS). vol. 33, pp. 6840–6851 (2020)
6. Hu, X., Li, F., Samaras, D., Chen, C.: Topology-Preserving Deep Image Segmentation. In: Advances in Neural Information Processing Systems (NeurIPS). vol. 32 (2019)
7. Hu, Y., Chang, S., Zhang, L., Tian, F., Sun, W., Lu, H.: Unisegdiff: Boosting unified lesion segmentation via a staged diffusion model. In: Medical Image Computing and Computer Assisted Intervention – MICCAI 2025. Lecture Notes in Computer Science, vol. 15961, pp. 663–673. Springer (2025)
8. Huang, D., Swanson, E.A., Lin, C.P., Schuman, J.S., Stinson, W.G., Chang, W., Hee, M.R., Flotte, T., Gregory, K., Puliafito, C.A., et al.: Optical Coherence Tomography. Science **254**(5035), 1178–1181 (1991). https://doi.org/10.1126/science.1957169
9. Isensee, F., Jaeger, P.F., Kohl, S.A.A., Petersen, J., Maier-Hein, K.H.: nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. Nature Methods **18**(2), 203–211 (2021). https://doi.org/10.1038/s41592-020-01008-z
10. Lu, D., Heisler, M., Lee, S., Bhalla, M., Kirat, G., Beg, M.F., Sarunic, M.V.: Deep-Learning Based Multiclass Retinal Fluid Segmentation and Detection in Optical Coherence Tomography Images Using a Fully Convolutional Neural Network. Medical Image Analysis **54**, 100–110 (2019)
11. Melinščak, M., Radmilović, M., Vatavuk, Z., Lončarić, S.: AROI: Annotated Retinal OCT Images Database. In: Proceedings of the International Symposium on Image and Signal Processing and Analysis (ISPA) (2021)

12. Ronneberger, O., Fischer, P., Brox, T.: U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Medical Image Computing and Computer Assisted Intervention – MICCAI 2015. pp. 234–241. Springer (2015)
13. Rosenfeld, P.J., Brown, D.M., Heier, J.S., Boyer, D.S., Kaiser, P.K., Chung, C.Y., Kim, R.Y.: Ranibizumab for Neovascular Age-Related Macular Degeneration. New England Journal of Medicine **355**(14), 1419–1431 (2006). https://doi.org/10.1056/NEJMoa054481
14. Roy, A.G., Conjeti, S., Karri, S.P.K., Sheet, D., Katouzian, A., Wachinger, C., Navab, N.: ReLayNet: Retinal Layer and Fluid Segmentation of Macular Optical Coherence Tomography using Fully Convolutional Networks. Biomedical Optics Express **8**(8), 3627–3642 (2017). https://doi.org/10.1364/BOE.8.003627
15. Shit, S., Paetzold, J.C., Sekuboyina, A., Ezhov, I., Unber, A., Zhylka, A., Pluim, J.P.W., Bauer, U., Menze, B.H.: clDice – a Novel Topology-Preserving Loss Function for Tubular Structure Segmentation. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) pp. 16560–16569 (2021)
16. Unterdechler, M., Fazekas, B., Aresta, G., Bogunović, H.: Comparative analysis of data augmentation for retinal oct biomarker segmentation. In: Ophthalmic Medical Image Analysis. pp. 94–103. Springer Nature Switzerland, Cham (2025)
17. Warfield, S.K., Zou, K.H., Wells, W.M.: Simultaneous Truth and Performance Level Estimation (STAPLE): An Algorithm for the Validation of Image Segmentation. IEEE Transactions on Medical Imaging **23**(7), 903–921 (2004). https://doi.org/10.1109/TMI.2004.828354
18. Wong, W.L., Su, X., Li, X., Cheung, C.M.G., Klein, R., Cheng, C.Y., Wong, T.Y.: Global prevalence of age-related macular degeneration and disease burden projection for 2020 and 2040: a systematic review and meta-analysis. The Lancet Global Health **2**(2), e106–e116 (2014). https://doi.org/10.1016/S2214-109X(13)70145-1
19. Wu, J., FU, R., Fang, H., Zhang, Y., Yang, Y., Xiong, H., Liu, H., Xu, Y.: Medsegdiff: Medical image segmentation with diffusion probabilistic model. In: Oguz, I., Noble, J., Li, X., Styner, M., Baumgartner, C., Rusu, M., Heinmann, T., Kontos, D., Landman, B., Dawant, B. (eds.) Medical Imaging with Deep Learning. Proceedings of Machine Learning Research, vol. 227, pp. 1623–1639. PMLR (10–12 Jul 2024), https://proceedings.mlr.press/v227/wu24a.html
20. Zhao, X., Li, C., Tang, J., Li, C.: Learning with Explicit Topological Priors for Chest X-ray Rib Segmentation. In: Medical Image Computing and Computer Assisted Intervention – MICCAI 2025. Lecture Notes in Computer Science, Springer (2025)