Vessel-Guided Aneurysm Localization and Segmentation for TopAneu 2026

Farid Tasharofi[1,2], Mersad Shoaei Taklimi[1,2], Leonhard Rist[1,2], and Andreas Maier[1]

[1] Pattern Recognition Lab, Friedrich-Alexander-Universität Erlangen, Germany
[2] Siemens Healthineers AG, Forchheim, Germany
farid.tasharofi@fau.de

Abstract. We describe a vessel-guided submission to TopAneu 2026, covering vessel-wise aneurysm localization (Task 1) and location-labelled aneurysm segmentation (Task 2) in CTA and MRA. A shared vessel prior is selected from ten configurations derived from a native nnU-Net and released TopBrain 2025 models. On a 68-case internal evaluation subset, the selected, taxonomy-aligned UZH (re-labelled) model reaches 0.801 multiclass and 0.899 binary vessel Dice. A vessel-guided Task 1 classifier that corrects lateralized labels during mirror augmentation and uses rarity-aware sampling achieves a mean per-class precision/recall/MCC score of 0.195 on the primary split and 0.181 on an independent split; aneurysm copy-paste augmentation yields a 12% relative improvement over its matched baseline. Task 2 experiments show that the initial unweighted presence loss collapses under severe voxel imbalance, whereas the revised foreground-weighted Dice+BCE loss yields mean Dice scores of 0.516 on CT and 0.469 on MR validation. These experiments identify taxonomy alignment and severe class imbalance as the main shared constraints.

Keywords: Intracranial aneurysm · Vessel segmentation · Transfer learning

1 Introduction

Intracranial aneurysms are pathological dilations of intracranial arteries. Their rupture causes aneurysmal subarachnoid hemorrhage, while management of unruptured aneurysms requires balancing the estimated rupture risk against the risks of preventive treatment [1]. Automated detection and anatomical localization in CTA and MRA may support image review.

The TopAneu 2026 challenge [5] provides a multi-center cohort of CTA and MRA volumes for two complementary problems: vessel-wise aneurysm localization across 52 anatomical locations and location-labelled aneurysm segmentation. Organizer-provided silver-standard vessel segmentations with 36 anatomical classes additionally enable the use of explicit vascular anatomy as a prior for both problems.

However, transferring existing vessel-segmentation models to aneurysm analysis is not straightforward. Recent high-performing vessel models use different anatomical taxonomies, and mismatches between vessel definitions can directly invalidate downstream location labels. Moreover, localization methods that rely on lateralized vessel anatomy require augmentation strategies that preserve left–right label consistency. Since several recent models provide strong but structurally different vessel representations, a systematic comparison is also needed to determine which model transfers most reliably to heterogeneous CTA and MRA data.

Related benchmarks include ADAM for intracranial aneurysm detection and segmentation [4], TopCoW for Circle-of-Willis segmentation [7], and TopBrain for whole-brain vessel segmentation with anatomical labels [8]. Building on this literature and the nnU-Net framework [2], we systematically compare recent high-performing vessel-segmentation models, align their anatomical taxonomies, and use the selected vessel representation as a prior for aneurysm localization and location-labelled segmentation. For localization, we correct laterality under mirror augmentation and address rare anatomical classes through rarity-aware sampling; for segmentation, we factorize aneurysm presence from anatomical location and address severe voxel imbalance with a foreground-weighted Dice+BCE presence loss.

2 Methods

2.1 Data and task definitions

The TopAneu 2026 dataset contains 417 scans (109 CTA, 308 MRA) from four centers and 389 aneurysms in 306 cases. All 111 negative cases are MRA. Of 52 anatomical vessel-location classes (including lateralized classes), 43 occur at least once. We use a patient-grouped, multi-label-stratified split (333 training/84 validation), balanced by modality, center, aneurysm presence, and location [5].

2.2 Selection of the vessel-segmentation prior

To identify a suitable anatomical vessel prior for downstream aneurysm localization and segmentation, we compared ten vessel-segmentation configurations. The comparison included models trained from scratch on the TopAneu vessel annotations as well as transferred models from the TopBrain 2025 benchmark [8].

Among the transferred models, UZH, KDH, and ARG correspond to released high-performing TopBrain submissions with complementary design strategies. UZH employs anatomically grouped nnU-Net models initialized from its previous TopCoW models; KDH uses modality-specific vessel-segmentation pipelines with pretrained models; and ARG follows a two-stage 3D nnU-Net design in which a binary vessel probability map is provided as additional input to a subsequent multiclass segmentation stage [8]. We fine-tuned the released models on the TopAneu data while retaining their original architectures and training strategies where applicable.

As internal baselines, we additionally trained nnU-Net ResEncUNet-M models at 3D full resolution [2]. Generic channel naming was used to enforce per-channel Z-score normalization for the mixed CTA/MRA data. The native model was trained directly on TopAneu's 36-class vessel taxonomy. We additionally evaluated a UZH+ARG hybrid, which supplies ARG's binary probability map to the regional UZH models, and Skeleton Recall Loss [3] with $\lambda_{\mathrm{SRL}} \in \{1.0, 0.1\}$. The four configurations selected for downstream analysis are summarized in Table 1.

Taxonomy alignment. TopBrain uses one ICA class, whereas TopAneu separates R/L-ICA-C6-C7 from R/L-ICA-C1-C5. Naive transfer therefore produced zero exact-label Dice for C1-C5 in three pretrained configurations. We rebuilt the output head with dedicated C1-C5 channels and continued fine-tuning from the shared network weights; incompatible segmentation-head weights were reinitialized.

Evaluation. All configurations were evaluated using binary and 36-class Dice on the 68 validation cases not used to train any compared checkpoint. These scores were used only to select the vessel prior and are separate from the downstream aneurysm-localization and segmentation evaluations.

2.3 Vessel-guided aneurysm localization

For vessel-wise aneurysm localization, we adapt the vessel-guided ROI classification approach of the RSNA 2025 first-place solution [6]. The localization model reuses the nnU-Net backbone trained for multiclass vessel segmentation. Its weights are initialized from the corresponding vessel-segmentation model, and a classification head is used for aneurysm localization. The resulting network is then fine-tuned for the localization objective, thereby transferring the anatomical representation learned during vessel segmentation to the downstream classification problem.

The selected vessel segmentation additionally defines the anatomical regions of interest used by the classifier. For each of the 52 aneurysm-location classes, the corresponding region is derived from the relevant classes of the 36-class vessel segmentation. A predefined binary mapping associates 33 aneurysm locations with a single vessel class, while the remaining 19 junction locations combine two vessel classes, or three for the VA–BA junction. All mappings were manually checked against the released anatomical taxonomies.

Because many location classes are strongly imbalanced, training cases are sampled according to class rarity. For class c with $n_c > 0$ positive cases among N training cases, we define $w_c = \min\{20, \max[1, (N - n_c)/n_c]\}$, and sample each case in proportion to the largest weight among its positive labels. Mirror augmentation is made anatomically consistent by jointly mirroring the image and vessel masks and swapping all left–right location labels.

As an additional augmentation for rare classes, we generate synthetic training cases by copying real aneurysm patches to anatomically compatible vessel

locations in other scans. For five classes absent from training, contralateral donor aneurysms are mirrored; for rare classes with 1–5 positives, same-class donors are reused. The patches undergo mild geometric perturbation and intensity-matched blending, yielding 99 additional training cases while leaving the validation set unchanged.

The localization network is trained with AdamW (learning rate 10^{-4}, weight decay 10^{-2}), batch size 1 with eight-step gradient accumulation, and mixed precision. The checkpoint with the best validation score within 200 epochs is retained.

2.4 Location-labelled aneurysm segmentation

The segmentation model uses an nnU-Net v2 ResEnc-M encoder–decoder backbone [2], trained jointly on CTA and MRA. The image is the only network input; the vessel prior restricts the training loss to within 5 mm of the predicted vasculature. A single 53-channel output head produces one aneurysm-presence logit and 52 anatomical-location logits.

We factorize voxel-wise aneurysm presence and anatomical localization as

$$P(y{=}k \mid x) = P(A{=}1 \mid x)\, P(y{=}k \mid A{=}1, x),$$

where $\hat{A} = \sigma(z_A)$ estimates aneurysm presence. The objective is

$$\begin{aligned} \mathcal{L} &= \mathcal{L}_{\text{presence}} + \mathcal{L}_{\text{loc}}, \\ \mathcal{L}_{\text{loc}} &= \frac{1}{N} \sum_{i=1}^{N} \frac{1}{|V_i|} \sum_{v \in V_i} \text{CE}(\hat{y}_v, y_v), \end{aligned} \tag{1}$$

where V_i denotes aneurysm i; averaging within each aneurysm reduces domination by large aneurysms.

Voxel-wise BCE for presence collapsed to an all-background solution, and adding an unweighted Dice term alone did not resolve this failure. We therefore use $\mathcal{L}_{\text{presence}} = \mathcal{L}^{w}_{\text{BCE}} + \mathcal{L}_{\text{Dice}}$, with positive-class weight $w = 568$, corresponding to the measured background-to-foreground voxel ratio in the training data. The location loss remains unchanged.

3 Results

3.1 Vessel-segmentation model comparison

Table 1 summarizes the four configurations selected for downstream analysis from the ten evaluated models. Re-labelled UZH achieved the highest multiclass Dice of 0.801, closely followed by UZH+ARG with 0.800. Although the Native baseline achieved the highest binary Dice of 0.922, its multiclass Dice was lower at 0.634; KDH achieved 0.713. The best standalone ARG variant reached a multiclass Dice of 0.347 and was not selected for downstream analysis. Taxonomy alignment introduced dedicated ICA-C1-C5 channels, changing their exact-label Dice from zero to nonzero values without degrading whole-ICA performance (Fig. 1).

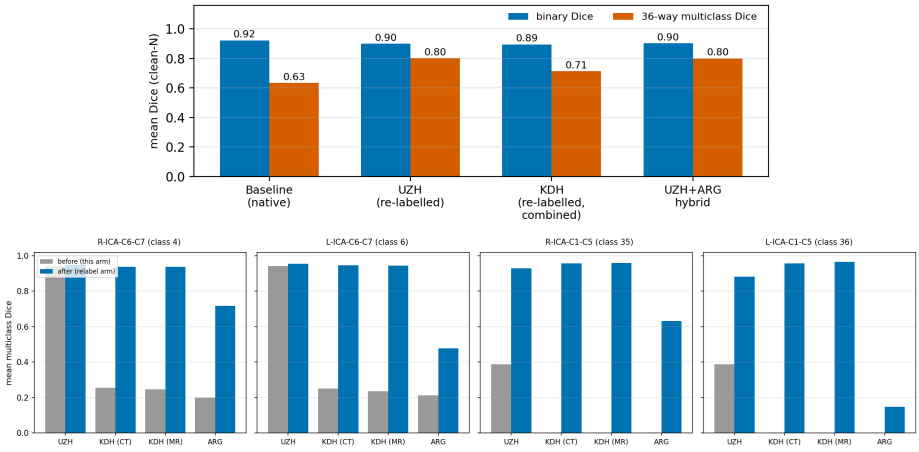

Fig. 1. Top: binary and multiclass Dice for the four vessel-segmentation configurations selected for downstream analysis on the 68-case evaluation subset. **Bottom**: effect of R/L-ICA-C1-C5 taxonomy alignment on per-class Dice before and after introducing dedicated output channels for these vessel segments.

3.2 Aneurysm localization results

Table 2 reports the Task 1 competition metric (mean per-class precision, recall, and MCC). Capped positive weighting and rarity sampling raise the plain baseline from 0.089 to 0.164; correcting mirrored labels further raises it to 0.195 (best checkpoint: epoch 12/200). An independent fold scores 0.181 (−7.2%), supporting an approximate 0.18–0.20 range (per-class Pearson $r = 0.39$). Aneurysm copy-paste augmentation improves its matched baseline from 0.164 to 0.184 (+12.0% relative). None of the five zero-example target classes has a positive validation example, so their recovery cannot be directly assessed. Detect-then-route collapses on 4 out of 5 folds; an independently unevaluated hierarchical variant is omitted.

Table 1. Vessel-segmentation performance on the 68-case evaluation subset for the four configurations selected for downstream analysis from the ten evaluated models. KDH combines separate CT and MR sub-models. The selected vessel prior is shown in bold.

Model	Binary Dice	36-class Dice
Native (baseline)	0.922	0.634
UZH (re-labelled)	**0.899**	**0.801**
KDH (re-labelled, combined)	0.894	0.713
UZH+ARG hybrid	0.903	0.800

Table 2. Task 1 competition score (mean per-class precision, recall, and MCC). Rows 1–3 show the baseline and two sequentially added components; row 4 evaluates the same configuration as row 3 on an independent fold; row 5 evaluates aneurysm copy-paste augmentation against the row-2 baseline in isolation.

Configuration	Competition score
Full-pool baseline (plain BCE, plain shuffle, 800 ep.)	0.089
+ capped pos-weight + rarity-weighted sampler	0.164
+ mirror-augmentation fix (fold 0)	0.195
+ mirror-augmentation fix (fold 1, indep. split)	0.181
+ aneurysm copy-paste augmentation (isolated ablation)	0.184

3.3 Location-labelled aneurysm segmentation results

The initial factorized loss for presence and location (Sect. 2.4) collapsed to an all-background presence prediction. Predicted foreground probability at ground-truth-positive voxels never exceeded 0.006; consequently, no positive voxels crossed the 0.5 threshold on either held-out or training-set voxels, indicating a training-dynamics failure rather than a generalization gap. Bypassing this gate, the conditional 52-way location softmax reached 41.4% top-1 accuracy at positive voxels (vs. approximately 1.9% chance), showing that the location head had nevertheless learned anatomical information.

With the revised foreground-weighted Dice+BCE formulation, the segmentation model achieved a Dice of 0.516 ± 0.365 on CT ($n = 22$) and 0.469 ± 0.333 on MR ($n = 62$). Corresponding VS scores were 0.672 ± 0.307 and 0.504 ± 0.396, with HD95 values of 29.3 ± 25.4 mm and 39.1 ± 22.6 mm, respectively (Table 3).

4 Discussion and Conclusion

Our internal experiments suggest two main lessons. First, the selected pretrained configuration outperformed the native baseline (0.801 vs. 0.634 multiclass Dice), although architectural and training differences prevent attributing the difference solely to pretraining. Second, class imbalance is a shared bottleneck: 9/52 location classes are absent, while small vessel classes are sparsely represented. Rarity-aware sampling and aneurysm copy-paste augmentation improve aneurysm localization, whereas unweighted voxel-wise BCE fails for segmentation and the

Table 3. Location-labelled aneurysm segmentation on the internal validation set (mean±SD).

Split	Dice	Volumetric Similarity	HD95 [mm]
CT ($n = 22$)	0.516±0.365	0.672±0.307	29.3±25.4
MR ($n = 62$)	0.469±0.333	0.504±0.396	39.1±22.6

foreground-weighted Dice+BCE formulation enables non-trivial segmentation performance.

The study is limited by the absence of challenge test-set results and use of a single primary Task 1 split. Consequently, all values are internal-validation results and should not be interpreted as challenge test-set performance.

In summary, we select a taxonomy-aligned vessel prior, report fold-checked aneurysm-localization results, and show that foreground-weighted Dice+BCE enables location-labelled aneurysm segmentation on both CT and MR.

Acknowledgments. The authors gratefully acknowledge the scientific support and HPC resources provided by the Erlangen National High Performance Computing Center (NHR@FAU) of the Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU). The hardware is partially funded by the German Research Foundation (DFG). This research has received funding from the scholarship program "Dhip campus-bavarian aim".

AI-assisted writing disclosure. During manuscript preparation, the authors used ChatGPT (OpenAI) and Claude (Anthropic) to assist with drafting and language refinement of selected passages. All AI-assisted content was reviewed, verified, and edited by the authors, who take full responsibility for the final manuscript.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Etminan, N., Rinkel, G.J.: Unruptured intracranial aneurysms: development, rupture and preventive management. Nature Reviews Neurology **12**(12), 699–713 (2016). https://doi.org/10.1038/nrneurol.2016.150
2. 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
3. Kirchhoff, Y., Rokuss, M.R., Roy, S., et al.: Skeleton Recall Loss for Connectivity Conserving and Resource Efficient Segmentation of Thin Tubular Structures. In: Leonardis, A., Ricci, E., Roth, S., Russakovsky, O., Sattler, T., Varol, G. (eds.) Computer Vision – ECCV 2024, LNCS, vol. 15135, pp. 218–234. Springer, Cham (2024). https://doi.org/10.1007/978-3-031-72980-5˙13
4. Timmins, K.M., van der Schaaf, I.C., Bennink, E., et al.: Comparing methods of detecting and segmenting unruptured intracranial aneurysms on TOF-MRAs: The ADAM challenge. NeuroImage **238**, 118216 (2021). https://doi.org/10.1016/j.neuroimage.2021.118216
5. TopAneu 2026 Organizing Committee: TopAneu 2026 Challenge: Intracranial Aneurysm Detection and Segmentation. `https://topaneu-26.grand-challenge.org/`. Last accessed 18 Aug 2026
6. uchiyama33: RSNA Intracranial Aneurysm Detection: 1st Place. GitHub repository, `https://github.com/uchiyama33/rsna2025_1st_place`. Last accessed 18 Aug 2026

7. Yang, K., Musio, F., Ma, Y., et al.: Benchmarking the CoW with the TopCoW Challenge: topology-aware anatomical segmentation of the Circle of Willis for CTA and MRA. arXiv:2312.17670 (2023). https://doi.org/10.48550/arXiv.2312.17670
8. Yang, K., Shi, P., Huang, H., et al.: TopBrain segmentation challenge for whole brain vessel anatomy. medRxiv (2026). https://doi.org/10.64898/2026.05.28.26354312