DINO-3DRA: Leveraging 2D Foundation Model Semantics for 3D Cerebral Aneurysm Segmentation

Jiayang Lu[1], Fengming Lin[1,2,3*], Alejandro F. Frangi[1,2,3,4,5*], and Ali Sarrami-Foroushani[1,2,3*†]

[1] Centre for Computational Imaging and Modelling in Medicine (CIMIM), University of Manchester, Manchester, UK
[2] Christabel Pankhurst Institute, University of Manchester, Manchester, UK
[3] Department of Computer Science, University of Manchester, Manchester, UK
[4] Division of Informatics, Imaging & Data Sciences, University of Manchester, Manchester, UK
[5] NIHR Manchester Biomedical Research Centre, Manchester Academic Health Sciences Centre, University of Manchester, Manchester, UK
a.sarrami@manchester.ac.uk

Abstract. Accurate aneurysm segmentation in 3D rotational angiography (3DRA) is hindered by extreme class imbalance, morphological similarity to vessels, and absent large-scale 3D pretraining. 2D vision foundation models encode dense structural priors from ∼1.7 billion images, yet naïve slice-wise transfer fragments anatomical continuity and destabilises optimisation. We propose DINO-3DRA, a dual-path framework achieving effective cross-dimensional semantic transfer by injecting frozen DINOv3 features into a 3D U-Net backbone via Room-Lite spatial mixing and calibrated residual fusion. On multi-centre 3DRA data, DINO-3DRA achieves state-of-the-art aneurysm segmentation (Dice: 0.758; HD95: 2.75 mm; +13% over nnU-Net) with only 5.72M trainable parameters. Ablation studies confirm that gains arise from structured cross-dimensional transfer rather than loss design alone, with bridged foundation features improving anatomical continuity between aneurysms and parent vessels. Without fine-tuning on CADA and SHINY-ICARUS, DINO-3DRA removes all failures (Dice < 0.5) observed in baseline architectures, indicating improved stability across heterogeneous protocols. Code and weights are available at https://github.com/JiayangDS/Dino3DRA.

Keywords: Aneurysm Segmentation · 3D Rotational Angiography · Vision Foundation Models · Cross-Dimensional Feature Transfer

1 Introduction

Intracranial aneurysms affect 3–5% of the population with significant rupture mortality [22]. While 3D rotational angiography (3DRA) enables detailed vas-

*Joint supervision.
†Corresponding author. Email: a.sarrami@manchester.ac.uk

cular imaging [17], automated segmentation remains challenging due to extreme size imbalance (<1% voxels), morphological similarity to normal vessels, and inter-scanner variability [13, 12].

A fundamental asymmetry exists between 2D and 3D visual representation learning. Self-supervised 2D vision foundation models (VFMs), particularly DINOv3, learn dense semantic features—object boundaries, part–whole relationships, and structural priors—through self-distillation on ∼1.7 billion natural images [16, 19]. These representations transfer effectively to medical modalities including CT and X-ray [14, 23]. Volumetric models, lacking comparable large-scale 3D pretraining due to data scarcity, struggle to learn equivalent priors [15]. Rather than pursuing expensive 3D pretraining, we leverage frozen 2D VFM features through a dual-path architecture: the frozen branch preserves pretrained knowledge while a trainable 3D branch learns domain-specific volumetric context, achieving foundation-scale benefits with minimal overhead.

Several works have adapted DINO for 2D medical imaging [23, 9, 5], yet effective 2D-to-3D transfer remains unresolved. Naïve slice-wise approaches fail to compete with optimised 3D architectures [14].

DINO-in-the-Room (DITR) projects calibrated RGB-D features onto 3D points [8]. 3DRA has no external views, camera, or point cloud, so its 2D-to-3D problem is distinct, requiring a dense semantic volume built from slice-wise tokens. These slice-wise features, however, lack inter-slice continuity, and the natural-to-medical domain mismatch risks destabilising hybrid optimisation.

This work presents DINO-3DRA with four primary contributions:

- **Lightweight cross-dimensional integration:** A dual-path architecture combining frozen DINOv3-Small with a 3D U-Net backbone [2] (∼120K additional parameters) achieves state-of-the-art aneurysm segmentation (Dice: 0.758; +13% over nnU-Net [6]) with only 5.72M trainable parameters.
- **Room-Lite spatial mixer:** A parameter-efficient module restoring volumetric coherence from slice-wise features via depth-aware positional encoding and depthwise-separable convolutions.
- **Calibrated heterogeneous fusion:** Residual injection with learnable scaling stabilises frozen–trainable feature alignment.
- **Systematic validation:** Ablations confirm structured cross-dimensional transfer drives gains; cross-dataset evaluation demonstrates improved robustness, eliminating catastrophic failures present in all baselines.

2 Methods

Figure 1 illustrates DINO-3DRA's dual-path architecture: a trainable 3D U-Net backbone [2] processing volumetric inputs and a frozen DINOv3-Small branch extracting semantic features from 2D axial slices, integrated via calibrated fusion modules.

Dual-Path Encoder. The *3D volumetric branch* follows the 3D U-Net architecture with four encoder stages, expanding channels from 32 to 256 while

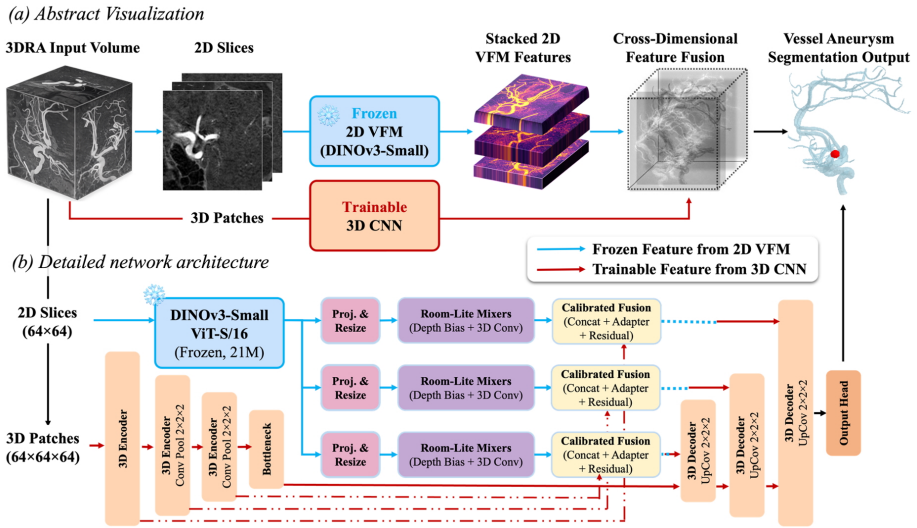

Fig. 1. Overview of DINO-3DRA. (a) Frozen 2D VFM (DINOv3-Small) features are fused with a trainable 3D U-Net backbone for vessel and aneurysm segmentation. (b) Multi-scale DINO features are refined by Room-Lite Mixers and fused into U-Net skip connections via Calibrated Fusion.

downsampling from 64^3 to 8^3 voxels. Each stage comprises two $3\times3\times3$ convolutions with GroupNorm and GELU. The symmetric decoder uses transposed convolutions and skip connections at 64^3, 32^3, and 16^3 resolutions (∼5.6M parameters).

The *frozen 2D semantic branch* (DINOv3-Small, ViT-S/16, 384-dim embeddings, pretrained on LVD-1689M via self-supervised distillation) processes all 64 axial slices independently, each replicated to RGB and resized to 224×224. For each slice, the 196 patch tokens (14×14 grid) are globally averaged to a single 384-dimensional semantic embedding capturing the slice-level vascular context. These embeddings are broadcast to spatial supports at $\{64^2, 32^2, 16^2\}$, and projected via 1×1 convolutions to $\{24, 48, 96\}$ channels. Stacked features are converted to isotropic 3D volumes via trilinear interpolation.

Room-Lite Mixers and Calibrated Fusion. Since DINOv3 processes slices independently, the stacked volume lacks inter-slice continuity. Inspired by DITR [8], Room-Lite mixers restore volumetric coherence via a learnable depth-positional bias $B(z) \in \mathbb{R}^{C\times D\times1\times1}$ encoding per-slice anatomical position, and depthwise-separable 3D convolutions:

$$Y = X + \mathrm{GELU}(\mathrm{GN}_1(\mathrm{Conv}_{1\times1\times1}(\mathrm{Conv}_{3\times3\times3}(X + B(z))))) \,. \tag{1}$$

To align the statistically different DINO and U-Net feature distributions, calibrated residual fusion is applied at each skip level:

$$F_{\text{fused}} = F_U + \mathrm{GELU}(\mathrm{GN}_1(W_{1\times1\times1}\,[F_U;\, \mathrm{GN}_1(F_D)])) \,, \tag{2}$$

where F_U and F_D denote U-Net and DINOv3 features, W is the $1\times1\times1$ projection weight, and the residual connection preserves F_U when DINO features are unhelpful. Both equations use GN_1, GroupNorm with a single group, normalising over all channels and spatial positions jointly. The decoder upsamples via transposed convolutions, concatenates fused skip features, and outputs three logits (background, vessel, aneurysm).

Training Objective. Severe class imbalance—background (∼94%), vessels (∼5%), aneurysms (<1%)—necessitates a composite loss:

$$\mathcal{L} = 0.3\,\mathcal{L}_{\text{vessel}} + 0.5\,\mathcal{L}_{\text{aneurysm}} + 0.2\,\mathcal{L}_{\text{CE}}. \tag{3}$$

Vessel Dice operates on the union of vessel and aneurysm classes to preserve vascular continuity. Aneurysm Dice isolates lesion segmentation [20]. Class-weighted cross-entropy ($w = \{0.1, 1.0, 2.0\}$) stabilises per-voxel classification, using normalised inverse frequency with clamping to prevent gradient explosion [3].

3 Experiments

3.1 Dataset and Implementation

Training and evaluation used the multi-centre 3DRA @neurIST dataset [1] (233 patients, four institutions) with expert annotations for cerebral vessels and aneurysms, split 4:1:1 (train/val/test). Volumes were resampled to 0.35 mm isotropic spacing, partitioned into overlapping 64^3 patches (stride 32), and augmented with random rotations ($\pm15°$) and flips. Case-level cross-dataset evaluation used CADA [7] (n=45, randomly sampled cases, aneurysm-only) and SHINY-ICARUS [18] (n=30, randomly sampled cases, vessel+aneurysm) without fine-tuning; these datasets differ from @neurIST in scanner uniformity, acquisition protocols, and label completeness.

Training used PyTorch 2.4.0 on two NVIDIA A100(80GB) GPUs with the Adam optimiser (lr=1×10^{-4}, weight decay=1×10^{-5}, batch size 4).

3.2 Preliminary Experiments

Before the main evaluation, we conducted controlled preliminary studies to select the loss function and fusion design. All runs used the same patch-level training/validation setting, with batch size 4, learning rate 1×10^{-4}, 60 epochs, and model selection by macro-averaged Dice. As shown in Tables 1 and 2, the separate vessel/aneurysm Dice loss with cross-entropy achieved the best overall performance and was adopted in all subsequent experiments. For architecture design, naïve DINO feature concatenation degraded performance, whereas Room-Lite with calibrated fusion produced the strongest total and aneurysm Dice, supporting the need for both volumetric coherence restoration and feature-distribution alignment when transferring frozen 2D foundation features to 3D segmentation.

Post-Processing. Overlapping patch predictions are aggregated via majority voting. The aneurysm threshold (0.005) was selected by grid search over

Table 1. Loss function comparison on the baseline 3D U-Net. Best results in **bold**.

ID	Loss Type	Dice (tot)	Dice (ves)	Dice (aneur)
A1	Separate Enhanced	**0.900**	0.927	**0.874**
A2	Simple Loss	0.874	0.919	0.828
A3	Asymmetric Basic	0.885	0.932	0.839
A4	Asymmetric Focal	0.857	**0.934**	0.780
A5	Asymmetric Adaptive	0.897	0.930	0.864

Table 2. Architectural ablation. All DINO variants use frozen DINOv3-small. Best results in **bold**.

ID	Configuration	Dice (tot)	Dice (ves)	Dice (aneur)
a1	Baseline 3D U-Net	0.896	0.922	0.871
a2	+ DINO (naive)	0.885	0.919	0.850
a3	+ Calib + Gating	0.891	0.924	0.858
a4	+ Room-Lite + Calib	**0.914**	**0.930**	**0.899**
b1	Multi-Dir (learn)	0.859	0.921	0.798
b2	Multi-Dir (concat)	0.865	0.921	0.810
b3	FAPM Projection	0.866	0.923	0.809

$[0, 0.1]$ on the baseline and applied uniformly, maximising aneurysm Dice with >90% voxel-level recall; vessel threshold follows the standard 0.5. Candidates are scored by $S = P_{\max} + \log_{10}(V)$ and filtered by minimum size ($\geq$40 voxels) and vessel proximity.

3.3 Comparison with State-of-the-Art

Table 3 compares DINO-3DRA with established baselines (paired Wilcoxon tests, $p < 0.05$). DINO-3DRA achieves the strongest aneurysm performance (Dice: 0.758, +13% over nnU-Net) with notably lower variance ($\sigma = 0.033$ vs. 0.097–0.138 for baselines). Qualitative results (Fig. 2) show clearer aneurysm boundaries and improved vascular continuity frequently missed by competing methods.

3.4 Ablation Study

We quantified the individual contributions of semantic transfer, volumetric coherence restoration, and calibrated fusion under identical training settings (Table 4). Compared configurations: (i) Baseline 3D U-Net (5.60M); (ii) Full DINO-3DRA (5.72M); (iii) Random-feature control, replacing pretrained DINOv3 with a ViT-S/16 using standard initialisation (truncated normal, std=0.02); (iv) w/o Room-Lite, directly stacking 2D features without spatial mixing; (v) w/o Calibration, using naïve concatenation in place of calibrated residual fusion; and (vi) DINOv2-small substitution. Figure 3 visualises vessel confidence maps for a challenging case. Variants w/o Room-Lite, w/o Calibration, and Random Features all exhibit low confidence across the aneurysm dome (Dice: 0.385–0.628),

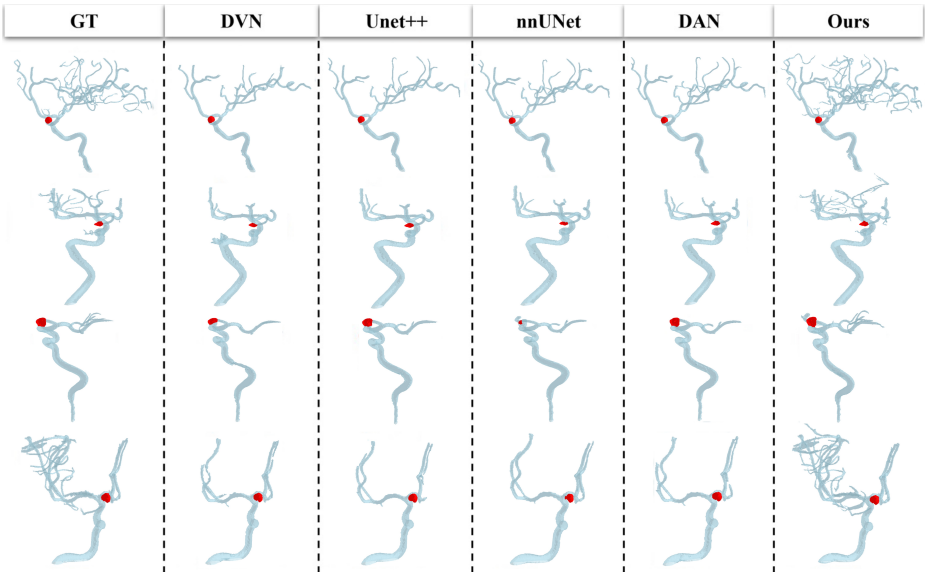

Fig. 2. Qualitative comparison (Blue: Vessel / Red: Aneurysm). Left to right: Ground truth (GT), DeepVesselNet (DVN), U-Net++, nnU-Net, Dual Attention (DAN), DINO-3DRA (Ours).

Table 3. Quantitative comparison on @neurIST dataset. p-values vs. Ours (paired Wilcoxon signed-rank test). Best results in **bold**.

Method	Structure	Dice	Jaccard	Vol. Sim.	p
U-Net++ (∼9M) [24]	Vessel	0.820±0.109	0.827±0.123	0.946±0.100	<0.01
	Aneurysm	0.598±0.340	0.504±0.322	0.658±0.337	<0.01
Dual Attention (∼18M) [4]	Vessel	0.889±0.138	0.818±0.145	0.945±0.095	<0.01
	Aneurysm	0.595±0.337	0.499±0.321	0.663±0.334	<0.01
DeepVesselNet (∼2M) [21]	Vessel	0.892±0.097	0.815±0.117	0.946±0.074	<0.01
	Aneurysm	0.156±0.235	0.108±0.190	0.262±0.277	<0.01
nnU-Net (∼31M) [6]	Vessel	0.869±0.120	0.782±0.137	0.913±0.107	<0.01
	Aneurysm	0.669±0.319	0.575±0.306	0.732±0.307	<0.01
Ours (∼5.72M)	Vessel	**0.897**±0.033	**0.815**±0.053	**0.962**±0.036	—
	Aneurysm	**0.758**±0.234	**0.661**±0.253	**0.838**±0.239	—

failing to recognise vascular tissue. Baseline U-Net and DINOv2 achieve reasonable detection but display fragmented confidence, treating aneurysms as isolated entities. Full DINO-3DRA produces homogeneous confidence across the entire aneurysm (Dice: 0.758), recognising aneurysms as pathological vessel dilations rather than separate anomalies—anatomical consistency that provides a better geometric starting point for downstream surface meshing and hemodynamic modeling [10, 11].

Table 4. Ablation results (paired Wilcoxon signed-rank test; HD95 in mm).

Configuration	Aneurysm			Vessel		
	Dice	p	HD95	Dice	p	HD95
Ours (Full)	**0.758**±0.234	—	**2.75**	**0.897**±0.033	—	**3.74**
w/o Room-Lite	0.385±0.310	<.001	22.08	0.857±0.043	<.001	5.64
w/o Calibration	0.492±0.301	<.001	10.12	0.857±0.045	<.001	6.71
Random Features	0.628±0.279	<.001	6.15	0.885±0.039	.042	4.41
DINOv2-small	0.590±0.252	<.001	6.50	0.868±0.050	<.001	4.30
Baseline U-Net	0.711±0.240	.030	4.20	0.873±0.046	.002	5.41

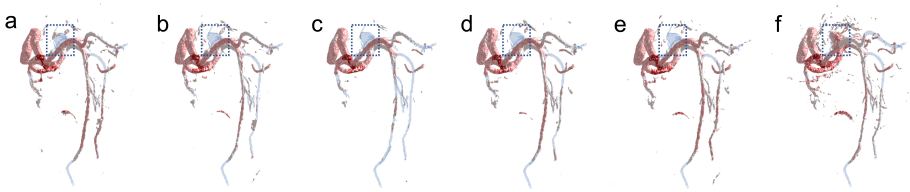

Fig. 3. Vessel confidence maps for ablation variants. Ground truth (semi-transparent blue), predictions (red; darker = higher confidence). (a) Baseline U-Net, (b) Random Features, (c) w/o Room-Lite, (d) w/o Calibration, (e) DINOv2-small, (f) DINO-3DRA (Ours).

3.5 Cross-Dataset Generalization

Models trained on @neurIST were evaluated on external datasets without fine-tuning: CADA (n=45) and SHINY-ICARUS (n=30). Results are summarised in Table 5 and Fig. 4. On SHINY-ICARUS, DINO-3DRA achieved Dice +10.4% (p <0.001) with HD95 reduced by 75.8% (15.52→3.76 mm). On CADA, mean Dice improved modestly (p=0.607) but variance halved (σ: 0.186→0.093). Critically, DINO-3DRA eliminated all failures (Dice < 0.5) on both datasets (CADA: 11.1%→0%; SI: 3.3%→0%), reducing the coefficient of variation by 52–70%. Case

Table 5. Cross-dataset generalization results. Failures: Dice<0.5; SI denotes SHINY-ICARUS. Best results in **bold**.

Dataset	Model	Dice	Jaccard	Vol. Sim.	HD95	Failures
CADA	Baseline	0.711±0.186	0.577±0.186	0.813±0.128	5.16	11.1%
	DINO-3DRA	**0.739**±0.093	**0.594**±0.118	0.807±0.109	**3.58**	**0%**
SI	Baseline	0.718±0.147	0.574±0.129	0.808±0.089	15.52	3.3%
	DINO-3DRA	**0.793**±0.049	**0.660**±0.065	0.812±0.056	**3.76**	**0%**

analysis reveals that DINO-3DRA tends to under-segment large or morphologically atypical aneurysms, reflecting the model's strong vessel prior: ambiguous boundary voxels at the dome periphery (where intensity gradients are subtle)

are conservatively assigned to the vessel class. Two factors amplify this effect: (1) training data consisting predominantly of compact saccular aneurysms, while CADA includes fusiform morphologies with inherently ambiguous boundaries; and (2) large aneurysms spanning multiple patches, where boundary consistency degrades. This trade-off favours clinical utility: reliable localisation and reduced failure rates over volumetric completeness, where consistent detection matters more than precise boundary delineation.

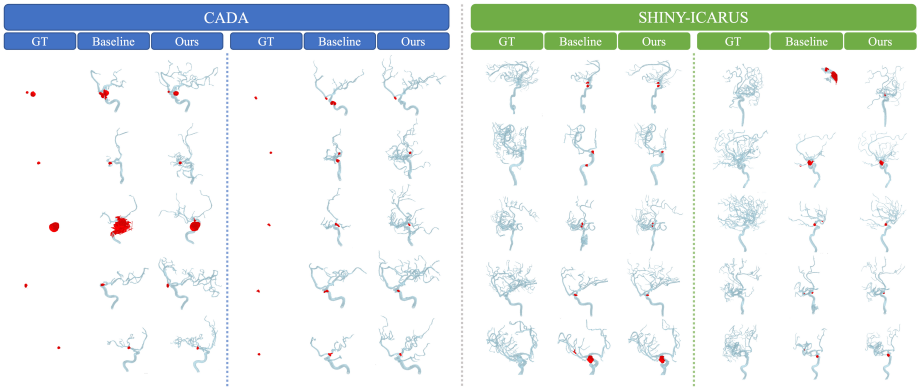

Fig. 4. Cross-dataset generalization validation on CADA (left) and SHINY-ICARUS (right). Blue: vessel; Red: aneurysm. Columns show ground truth, baseline, and DINO-3DRA predictions.

4 Conclusion

This paper presents DINO-3DRA, a framework that transfers frozen 2D foundation model features to 3D medical segmentation via Room-Lite spatial mixers and calibrated residual fusion. Ablation studies confirm that proper bridging mechanisms are essential for cross-dimensional semantic transfer, yielding anatomically coherent predictions in which aneurysms are recognised as pathological vessel dilations. The method achieves state-of-the-art aneurysm segmentation (Dice: 0.758, +13% over nnU-Net) and vessel segmentation (Dice: 0.897) with substantially reduced prediction variance. Cross-dataset evaluation eliminates all failures cases (Dice < 0.5) while preserving fine vascular structures and vessel-aneurysm continuity, suggesting improved robustness across heterogeneous protocols, to be confirmed on larger external cohorts.

Acknowledgements. AFF acknowledges support from the Royal Academy of Engineering under the RAEng Chair in Emerging Technologies (INSILEX CiET191919), ERC Advanced Grant–UKRI Frontier Research Guarantee (INSILICO EPY0304941), the UK Centre of Excellence on in-silico Regulatory

Science and Innovation (UK CEiRSI) (10139527), the National Institute for Health and Care Research (NIHR) Manchester Biomedical Research Centre (BRC) (NIHR203308), the BHF Manchester Centre of Research Excellence (RE2430017), and the CRUK RadNet Manchester (C1994A28701).

Disclosure of Interests. AF is Co-founder and shareholder of OculomeX Health Ltd, adsilico Ltd, and Synaptive Consulting Ltd.

References

1. Benkner, S., Arbona, A., Berti, G., Chiarini, A., Dunlop, R., Engelbrecht, G., Frangi, A.F., Friedrich, C.M., Hanser, S., Hasselmeyer, P., et al.: @neurist: Infrastructure for advanced disease management through integration of heterogeneous data, computing, and complex processing services. IEEE Transactions on Information Technology in Biomedicine **14**(6), 1365–1377 (2010)
2. Çiçek, Ö., Abdulkadir, A., Lienkamp, S.S., Brox, T., Ronneberger, O.: 3d u-net: Learning dense volumetric segmentation from sparse annotation. In: Medical Image Computing and Computer-Assisted Intervention (MICCAI). Lecture Notes in Computer Science, vol. 9901, pp. 424–432. Springer (2016)
3. Cui, Y., Jia, M., Lin, T.Y., Song, Y., Belongie, S.: Class-balanced loss based on effective number of samples. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 9268–9277 (2019)
4. Fu, J., Liu, J., Tian, H., Li, Y., Bao, Y., Fang, Z., Lu, H.: Dual attention network for scene segmentation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 3146–3154 (2019)
5. Gao, Y., Li, H., Yuan, F., Wang, X., Gao, X.: Dino u-net: Exploiting high-fidelity dense features from foundation models for medical image segmentation. arXiv preprint arXiv:2508.20909 (2025)
6. Isensee, F., Jaeger, P.F., Kohl, S.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)
7. Ivantsits, M., Goubergrits, L., Kuhnigk, J.M., Huellebrand, M., Brüning, J., Kossen, T., Pfahringer, B., Schaller, J., Spuler, A., Kuehne, T., et al.: Cerebral aneurysm detection and analysis challenge 2020. In: International Workshop on Cerebral Aneurysm Detection. pp. 3–17. Springer (2020)
8. Knaebel, K., Yilmaz, K., de Geus, D., Hermans, A., Adrian, D., Linder, T., Leibe, B.: Dino in the room: Leveraging 2d foundation models for 3d segmentation. arXiv preprint arXiv:2503.18944 (2025)
9. Li, Y., Wu, Y., Lai, Y., Hu, M., Yang, X.: Meddinov3: How to adapt vision foundation models for medical image segmentation? arXiv preprint arXiv:2509.02379 (2025)
10. Lin, F., Xia, Y., Deo, Y., MacRaild, M., Dou, H., Liu, Q., Wu, K., Ravikumar, N., Frangi, A.F.: Unsupervised domain adaptation for brain vessel segmentation through transwarp contrastive learning. In: 2024 IEEE International Symposium on Biomedical Imaging (ISBI). pp. 1–5. IEEE (2024)
11. Lin, F., Xia, Y., MacRaild, M., Deo, Y., Dou, H., Liu, Q., Cheng, N., Ravikumar, N., Frangi, A.F.: Gs-ema: Integrating gradient surgery exponential moving average with boundary-aware contrastive learning for enhanced domain generalization in aneurysm segmentation. arXiv preprint arXiv:2402.15239 (2024)

12. Lin, F., Xia, Y., Ravikumar, N., Liu, Q., MacRaild, M., Frangi, A.F.: Adaptive semi-supervised segmentation of brain vessels with ambiguous labels. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 106–116. Springer (2023)
13. Lin, F., Xia, Y., Song, S., Ravikumar, N., Frangi, A.F.: High-throughput 3dra segmentation of brain vasculature and aneurysms using deep learning. Computer Methods and Programs in Biomedicine **230**, 107355 (2023)
14. Liu, C., Chen, Y., Shi, H., Lu, J., Jian, B., Pan, J., Cai, L., Wang, J., Yu, J., Gao, Z., et al.: Does dinov3 set a new medical vision standard? benchmarking 2d and 3d classification, segmentation, and registration. arXiv preprint arXiv:2509.06467 (2025)
15. Müller-Franzes, G., Khader, F., Siepmann, R., Han, T., Kather, J.N., Nebelung, S., Truhn, D.: Medical slice transformer for improved diagnosis and explainability on 3d medical images with dinov2. Scientific Reports **15**(1), 23979 (2025)
16. Oquab, M., Darcet, T., Moutakanni, T., Vo, H., Szafraniec, M., Khalidov, V., Fernandez, P., Haziza, D., Massa, F., El-Nouby, A., et al.: Dinov2: Learning robust visual features without supervision. arXiv preprint arXiv:2304.07193 (2023)
17. van Rooij, W.J., Sprengers, M.E., de Gast, A.N., Peluso, J.P.P., Sluzewski, M.: 3d rotational angiography: the new gold standard in the detection of additional intracranial aneurysms. American Journal of Neuroradiology **29**(5), 976–979 (2008)
18. SHINY-ICARUS Organizers: Shiny-icarus: Segmentation over three dimensional rotational angiography of internal carotid artery with aneurysm. ISBI 2023 Challenge, IEEE International Symposium on Biomedical Imaging (2023). https://doi.org/10.17023/z7f7-kz51, dataset with 3DRA vessel and aneurysm annotations
19. Siméoni, O., Vo, H.V., Seitzer, M., Baldassarre, F., Oquab, M., Jose, C., Khalidov, V., Szafraniec, M., Yi, S., Ramamonjisoa, M., et al.: Dinov3. arXiv preprint arXiv:2508.10104 (2025)
20. Sudre, C.H., Li, W., Vercauteren, T., Ourselin, S., Jorge Cardoso, M.: Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations. In: International Workshop on Deep Learning in Medical Image Analysis. pp. 240–248. Springer (2017)
21. Tetteh, G., Efremov, V., Forkert, N.D., Schneider, M., Kirschke, J., Weber, B., Zimmer, C., Piraud, M., Menze, B.H.: Deepvesselnet: Vessel segmentation, centerline prediction, and bifurcation detection in 3-d angiographic volumes. Frontiers in Neuroscience **14**, 592352 (2020)
22. Vlak, M.H., Algra, A., Brandenburg, R., Rinkel, G.J.: Prevalence of unruptured intracranial aneurysms, with emphasis on sex, age, comorbidity, country, and time period: a systematic review and meta-analysis. The Lancet Neurology **10**(7), 626–636 (2011)
23. Xu, G., Udupa, J.K., Lu, W., Zhang, Y.: Exploiting dinov3-based self-supervised features for robust few-shot medical image segmentation. arXiv preprint arXiv:2601.08078 (2026)
24. Zhou, Z., Siddiquee, M.M.R., Tajbakhsh, N., Liang, J.: Unet++: Redesigning skip connections to exploit multiscale features in image segmentation. IEEE Transactions on Medical Imaging **39**(6), 1856–1867 (2019)