

000 RETHINKING INTRACRANIAL ANEURYSM VESSEL 001 SEGMENTATION: A PERSPECTIVE FROM COMPUTA- 002 TIONAL FLUID DYNAMICS APPLICATIONS 003

004 **Anonymous authors**
 005

006 Paper under double-blind review
 007

008 ABSTRACT 009

010 The precise segmentation of intracranial aneurysms and their parent vessels (IA-
 011 Vessel) is a critical step for hemodynamic analyses, which mainly depends on com-
 012 putational fluid dynamics (CFD). However, current segmentation methods pre-
 013 dominantly focus on image-based evaluation metrics, often neglecting their prac-
 014 tical effectiveness in subsequent CFD applications. To address this deficiency, we
 015 present the Intracranial Aneurysm Vessel Segmentation (IAVS) dataset, the first
 016 comprehensive, multi-center collection comprising 641 3D MRA images with 587
 017 annotations of aneurysms and IA-Vessels. In addition to image-mask pairs, IAVS
 018 dataset includes detailed hemodynamic analysis outcomes, addressing the limita-
 019 tions of existing datasets that neglect topological integrity and CFD applicability.
 020 To facilitate the development and evaluation of clinically relevant techniques, we
 021 construct two evaluation benchmarks including global localization of aneurysms
 022 (Stage I) and fine-grained segmentation of IA-Vessel (Stage II) and develop a sim-
 023 ple and effective two-stage framework, which can be used as a out-of-the-box
 024 method and strong baseline. For comprehensive evaluation of applicability of seg-
 025 mentation results, we establish a standardized CFD applicability evaluation sys-
 026 tem that enables the automated and consistent conversion of segmentation masks
 027 into CFD models, offering an applicability-focused assessment of segmentation
 028 outcomes. The data, code, and model will be made publicly available upon accep-
 029 tance.
 030

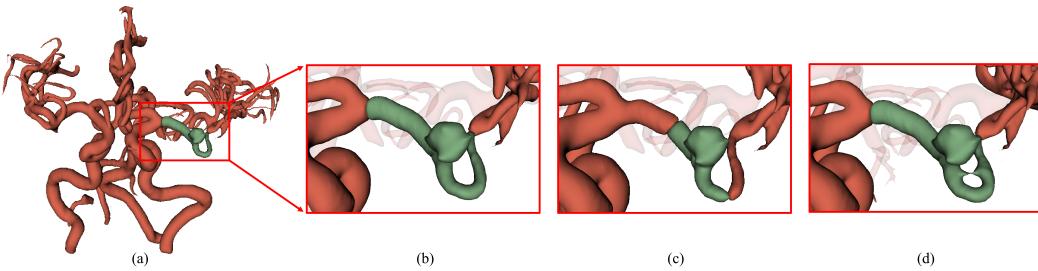
031 1 INTRODUCTION 032

033 Intracranial aneurysm (IA) is a pathological dilation of blood vessels, mainly occurring at the
 034 branches and bifurcations of arteries (Schiavink, 1997). IA is usually small and initially asymp-
 035 totic, but may gradually enlarge over time and lead to symptomatic manifestations, and even
 036 rupture in severe cases, resulting in a high incidence of morbidity and mortality (Cebral et al., 2005).
 037 Accurate assessment of rupture risk of IA is essential for medical intervention of neurovascular
 038 diseases (Etminan & Rinkel, 2016). Computational Fluid Dynamics (CFD) provides key biomechanical
 039 evidence for clinical decision-making by quantifying hemodynamic parameters such as wall shear
 040 stress and pressure distribution, which have been widely applied in various biomedical researches
 041 (Li et al., 2025; Morris et al., 2016; Wang et al., 2025).
 042

043 Magnetic resonance angiography (MRA) serves as a non-invasive, high-resolution imaging modality
 044 that facilitates the detailed visualization of aneurysms, enabling the identification of their anatomical
 045 characteristics, including location, size, and complex morphological features (Pierot et al., 2013).
 046 Accurate segmentation of intracranial aneurysm and parent vessels (IA-Vessel) from MRA is an
 047 important step for subsequent CFD analysis (Patel et al., 2023). As manual localization and delin-
 048 eation remain a labor-intensive and time-consuming procedure for radiologists (Jiao et al., 2023), it
 049 is highly desirable to develop automated segmentation methods in clinical applications. With the
 050 unprecedented developments of deep learning, state-of-the-art segmentation methods have achieved
 051 comparable results with inter-rater variability (Isensee et al., 2021). As deep learning-based methods
 052 require labeled data for training, high quality open-source datasets have become a crucial founda-
 053 tion for the development of segmentation algorithms for various modalities of medical imaging
 (Antonelli et al., 2022; Gatidis et al., 2022; Ji et al., 2022; Ma et al., 2022; Qu et al., 2023).

054 Table 1: Summary of existing 3D MRA datasets for intracranial aneurysm segmentation tasks.
055

056 Dataset	057 Volumes	058 IAs	059 IA-Vessel 060 Mask	061 STL	062 IA-Vessel 063 Centerline	064 Mesh	065 CFD 066 Results
067 ADAM	068 113	069 156	070 X	071 X	072 X	073 X	074 X
075 INSTED	076 191	077 68	078 X	079 X	080 X	081 X	082 X
083 Royal	084 63	085 85	086 ✓	087 ✓	088 X	089 X	090 X
091 IAVS(Ours)	092 641	093 587	094 ✓	095 ✓	096 ✓	097 ✓	098 ✓

075 Figure 1: (a) Whole intracranial vasculature and local parent vessels. (b) IA-Vessel ground truth.
076 (c) Despite the Dice score is relatively low (0.7648), no topological errors are present. (d) Although
077 the Dice similarity coefficient is high (0.9869), topological errors are present which is unusable for
078 CFD.
079

080 Despite the existence of several datasets for intracranial aneurysm segmentation, challenges persist
081 when applying these datasets to hemodynamic analysis. First, there are structural deficiencies in
082 the annotations of these datasets. Existing public datasets, such as ADAM (Timmins et al., 2021)
083 and Royal(de Nys et al., 2024), generally lack refined annotations of the parent vessels and geo-
084 metric validation labels. Additionally, they do not include records of hemodynamic results, which
085 makes it challenging to support the end-to-end analysis process from image segmentation to CFD
086 modeling. Second, the evaluation of segmentation results is limited. Most existing medical image
087 segmentation models are assessed using region overlap-based metrics, such as the Dice coefficient.
088 However, these metrics are insensitive to geometric topological abnormalities, including vessel ad-
089 hesion and surface irregularities. These abnormalities usually fail CFD validation because of issues
090 such as mesh generation failure or flow field distortion. Moreover, insufficient localization accu-
091 racy for small-sized aneurysms and the limited capability to maintain vascular connectivity further
092 exacerbate the challenges in transitioning from image segmentation to biomechanical modeling.

093 To address these challenges, this study presents a systematic solution for segmenting intracranial
094 aneurysms and vessels applicable to CFD, innovating across three sub-tasks: **dataset construction**,
095 **benchmark design**, and **evaluation system**. The main contributions are outlined as follows.

- 096 • We collect and curate a large-scale multi-centre **Intracranial Aneurysm Vessel**
097 **Segmentation (IAVS)** dataset, comprising 641 3D MRA images and 587 annotations of
098 aneurysms and IA-Vessels, including CFD analysis results. This dataset addresses the lim-
099 itations of previous datasets that lack topological integrity and CFD applicability.
- 100 • We establish an standardised CFD applicability evaluation system that enables standard-
101 ized estimation of CFD success probability given segmentation results. Additionally, we
102 introduce a novel evaluation metric, the **CFD-Applicability Score (CFD-AS)**, to facilitate a
103 more comprehensive assessment of segmentation results.
- 104 • We conduct two evaluation benchmarks including global localization of aneurysms (Stage
105 I) and fine-grained segmentation of IA-Vessel (Stage II) and develop a two-stage frame-
106 work as a strong baseline for the accurate detection and segmentation of IA-Vessel, which
107 significantly reduces geometric errors in segmentation masks and enhances CFD usability.

108
109

2 RELATED WORK

110
111
112
113
114
115
116
117
118
119
120
121
122
123

Intracranial Aneurysm Datasets. To accelerate the development of deep learning-based aneurysm and vessel segmentation, several segmentation datasets are evolved. However, existing public intracranial aneurysm datasets exhibit substantial limitations when applied to CFD studies. Regarding annotation completeness, the ADAM (Timmins et al., 2021) and INSTED (Chen et al., 2024) datasets offer 3D MRA images with aneurysm masks. However, they lack annotations of the parent vessels, which are essential for constructing CFD models. Conversely, the AneuX (Juchler et al., 2022) project provides preprocessed STL models for CFD but omits the original medical images and segmentation masks. In terms of anatomical accuracy, the Royal (de Nys et al., 2024) dataset includes both aneurysm outlines and vessel annotations. Nevertheless, several samples feature vessel adhesion, which undermine the validity of CFD boundary conditions. Similarly, the COSTA dataset (Mou et al., 2024) contains whole-brain vessel annotations, but suffers from adhesion errors in numerous distal branches of vessels, which inaccuracies directly impede the precision of CFD simulations. Overall, these works fail to provide a comprehensive database from image segmentation to CFD analysis, which underscores the necessity of developing application-oriented segmentation dataset.

124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141

Aneurysm Vessel Segmentation. Before the advent of deep learning, aneurysm and vascular segmentation relied mainly on classical vesselness-based methods, most notably the multiscale Hessian filter Frangi et al. (1998), with later benchmarks Lamy et al. (2022) revealing their variability across anatomical regions. However, these approaches struggle with complex aneurysm-parent-vessel configurations and lack the geometric fidelity required for downstream hemodynamic analysis. Deep learning methods have shown excellent performance on several medical image segmentation tasks, yet aneurysm vessel segmentation presents unique challenges. Mainstream segmentation networks like 3D UNet (Çiçek et al., 2016) and nnUNet (Isensee et al., 2021) prioritize global voxel-wise accuracy but lack mechanisms for reliable small-target detection, essential for accurately segmenting both small aneurysms and fine vessels. Glia-Net (Bo et al., 2021) enhances aneurysm delineation via global context fusion but does not extend to parent-vessel segmentation. Object detection frameworks such as nnDetection (Baumgartner et al., 2021) achieve robust 3D lesion localization but falter on sub-voxel scale targets. Sphere-based detectors like CPM-Net (Song et al., 2020) and SCPM-Net (Luo et al., 2022) help stabilize small-object training dynamics but remain untested on vascular structures. Keypoint detection methods like MedLSAM (Lei et al., 2025) demonstrate promise for anatomical localization but have not been adapted for variable-size aneurysm center points. While AA-Seg (Yao et al., 2024a) pioneers joint aneurysm-vessel segmentation, it still permits vessel adhesion across the aneurysm neck, highlighting the ongoing need for methods that can accurately and jointly segment both structures while respecting anatomical boundaries.

142
143
144
145
146
147
148
149
150
151

Evaluation Metrics. Conventional segmentation metrics inadequately capture the requirements of downstream CFD analysis. The Dice similarity coefficient (DSC) quantifies volumetric overlap but is insensitive to topological errors such as spurious vessel connections. Boundary IoU Cheng et al. (2021) improves edge accuracy assessment yet remains blind to global connectivity flaws. Centerline-aware metrics (clDice) Shit et al. (2021) incorporate explicit topological constraints but do not directly reflect mesh-generation feasibility or flow-convergence behavior. While innovative research into differentiable CFD solvers Yao et al. (2024b) aims to integrate physical simulations directly into the training loop, these methods are not yet directly applicable to the clinical task of intracranial aneurysm analysis due to the complex geometries and the non-differentiable nature of the traditional high-fidelity meshing and simulation pipeline required for clinical validation.

152
153
154

3 IAVS DATASET

155
156
157
158
159
160
161

Motivation and Details. Existing intracranial aneurysm datasets have structural deficiencies in the annotation and lack applicability in CFD applications. To bridge this gap, our IAVS dataset contains 641 3D MRA images and 587 aneurysms and IA-Vessels annotations with CFD analysis results, which is adapted from three existing datasets including ADAM (Timmins et al., 2021), INSTED (Chen et al., 2024) and Royal (de Nys et al., 2024), and a new in-house dataset from [hidden for review]. An overview of IAVS dataset is shown in Figure 2. For public datasets, the original ADAM and INSTED datasets only provide annotations of IAs. Despite the Royal dataset contains IA-Vessel mask and STL models, several samples feature vessel adhesion and are not applicable for CFD

162

163

Table 2: Statistics of IAVS dataset including data source, number and diameter of IAs.

164

165

Dataset	No. of Images			No. of IAs per case				IAs	Diameter of IA		
	Total	Public	Private	0	1	2	≥ 3		<3mm	3-7mm	>7mm
Train	467	175	292	82	345	34	6	432	55	272	105
Set A	76	76	0	42	29	3	2	41	16	17	8
Set B	98	0	98	0	85	10	3	114	10	93	11

170

171

172

173

Table 3: Statistics of imaging parameters in the IAVS dataset.

174

175

176

Dataset Statistics	Min	Median	Max
Spacing (mm)	(0.21, 0.21, 0.30)	(0.36, 0.36, 0.50)	(0.47, 0.47, 1.20)
Volume Size (voxels)	(348, 384, 44)	(512, 512, 148)	(1024, 1024, 368)

177

178

179

180

analysis. In contrast, our dataset contains CFD applicable segmentation masks and CFD analysis results, including 3D MRA images (1), voxel-level segmentation masks (2)-(3), geometric models (4)-(6) and CFD analysis results (7).

181

Data Statistics. The IAVS dataset is partitioned meticulously to meet the practical requirements of clinical research. Statistics of the proposed IAVS dataset in Table 2 reveals that the distribution of aneurysm quantity and size across cases closely mirrors clinical epidemiological patterns. This congruence effectively guarantees the representativeness of the dataset, enhancing the generalizability of the research findings. Additionally, all data underwent strict anonymization procedures and were rigorously reviewed and approved by the hospital ethics committee, ensuring full compliance with ethical standards. We split the images into 467 cases for training and validation, 76 cases from public datasets as Set A for internal evaluation, and 98 cases from in-house dataset as Set B for evaluation in clinical scenarios. For the training of Stage I, 373 cases are used for training and 94 cases are used for validation. In Stage II, candidate patches cropped based on IA annotation are used for training of IA-Vessel segmentation network. Following the same split of MRA images in Stage I, 357 patches are used for training and 99 patches are used for validation.

182

183

Annotation. After integrating medical imaging resources from three public and private datasets, we conduct annotations of IA and IA-Vessels for CFD applicable segmentation. The annotation workflow can be observed in Figure 2. IA annotations from the existing datasets are used if available. For in-house dataset, the annotations are completed and checked by experienced radiologists. For annotation of parent vessels of IA, we first use a pre-trained model using COSTA (de Nys et al., 2024) to preliminary segment whole-brain vessels of all MRA images. The pre-trained coarse-vessel model achieves Dice of 0.9204 on the official COSTA test set. Subsequently, focusing on the aneurysm-related vessel regions, the parent vessels are cropped and refined from the coarse segmentation of whole-brain vessels using 3D Slicer. The model-generated vessel segmentation mask is refined and verified by one CFD specialist and one board-certified radiologist instead of directly used without human correction. The refinement process strictly adheres to clinical anatomical principles, including eliminating abnormal geometric features and implementing an adaptive truncation strategy based on vascular bifurcation topology. When the parent vessel extends to the bifurcation, if the length of this segment of the vessel exceeds its diameter, truncation processing is performed. This strategy effectively avoids adhesion issues of distal small branches while ensuring the learnability of vessel length for the model.

184

185

186

187

188

189

190

191

192

193

194

Quality Control. To further validate the CFD usability of annotations, other than conducting voxel-level segmentation masks, all cases are conducted vascular geometric annotations for CFD analysis, including STL files of cut inlet/outlet sections, vascular centerline data, and mesh grid files labeled with fluid boundary conditions. Besides, CFD applicability of the segmentation masks are evaluated to validate whether the pressure and velocity residuals in the blood flow dynamics analysis achieve convergence. We perform a rigorous quality check and screening, annotations not applicable for CFD are further refined and validated, or removed from the final dataset.

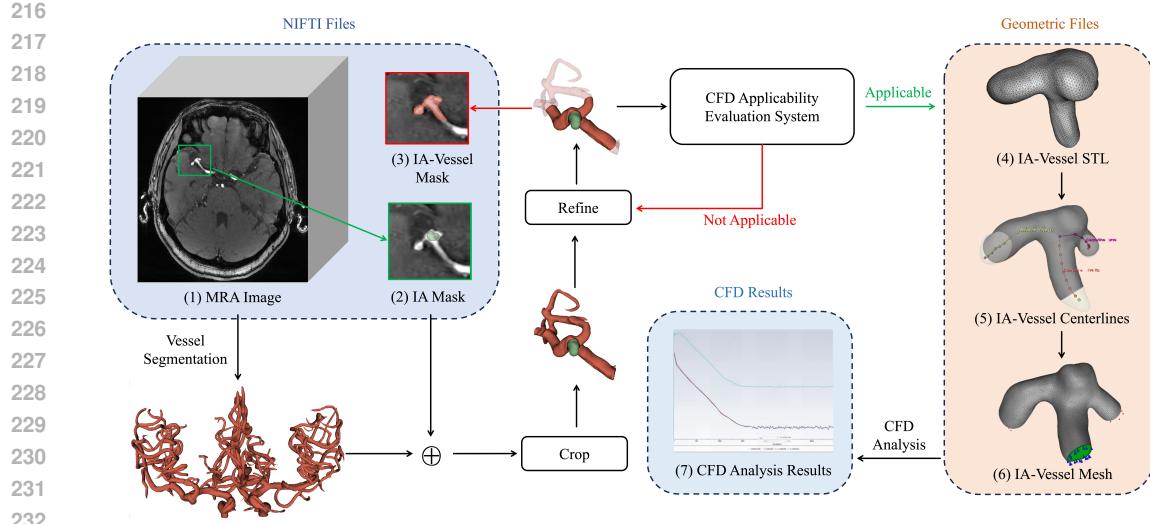


Figure 2: An overview of the IA-Vessel dataset and the annotation workflow. Each case encompasses seven types of standardized data: (1) whole-brain MRA images, (2) IA mask, (3) IA-Vessel mask, (4) STL models with cut inlets/outlets, (5) vascular centerlines, (6) mesh files with boundary annotations, (7) CFD analysis results.

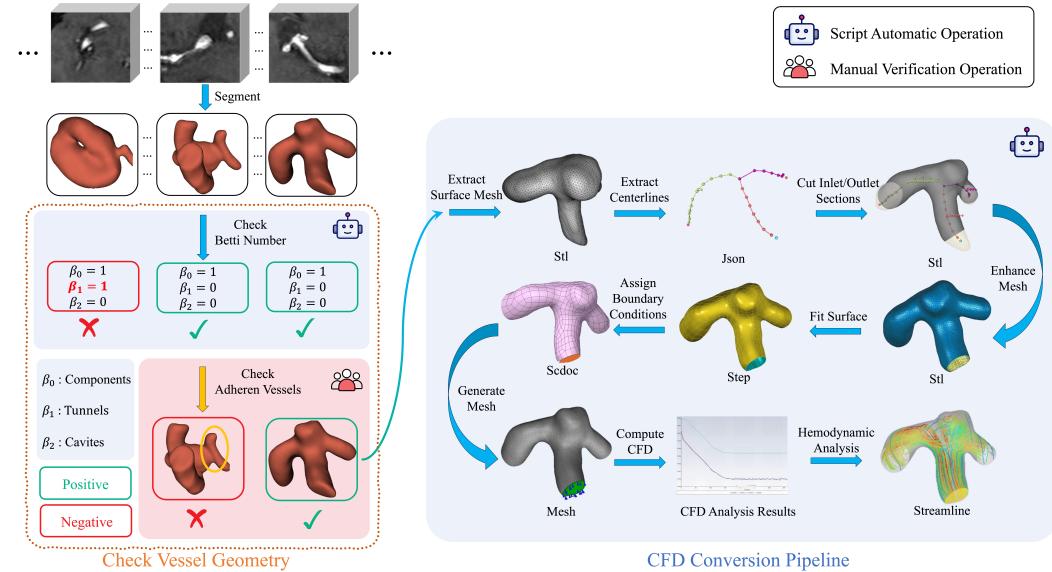


Figure 3: Overview of our conversion pipeline from segmentation masks to CFD models, which realizes the entire chain process from medical imaging to flow field simulation. The pipeline consists of following steps, including vascular topology inspection, morphological preprocessing, geometric model conversion, centerline generation, end face cutting, mesh enhancement, surface fitting, boundary labeling, mesh generation, and CFD computation.

4 CFD APPLICABILITY EVALUATION SYSTEM

To achieve automated and standardized conversion from segmentation mask to CFD model, we establish a standardised CFD applicability evaluation system as shown in Figure 3. The pipeline consists

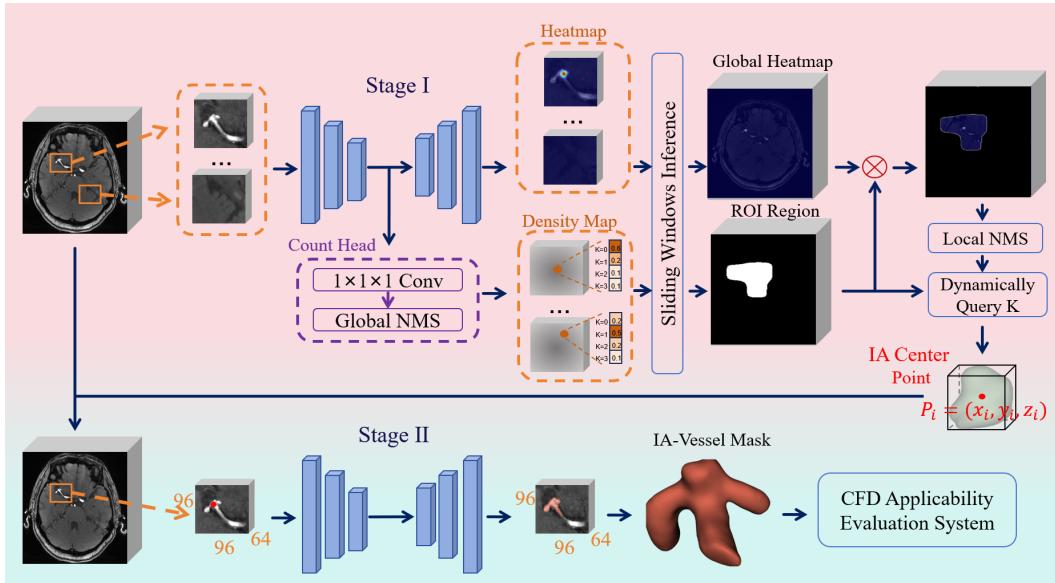


Figure 4: Our proposed two-stage framework for IA-Vessel segmentation. Stage I utilizes a detection network for global localization of aneurysms. After cropping out candidate patches, Stage II utilizes a topological-aware segmentation network for IA-Vessel segmentation to reduce topology errors.

of the following steps, including vascular topology inspection, morphological preprocessing, geometric model conversion, centerline generation, end face cutting, mesh enhancement, surface fitting, boundary labeling, mesh generation, and CFD computation. The detailed procedure for each step is shown in the Appendix A.

Based on the evaluation system, we propose a novel applicability-based evaluation metric entitled CFD applicability score (CFD-AS) to enable more comprehensive evaluation of segmentation results, which is defined as follows:

$$AS_{CFD} = \frac{\widehat{TP}}{TP + FP + FN} \quad (1)$$

$$\widehat{TP} = \sum_{i=1}^N (y = 1) \wedge (\hat{y} = 1) \wedge (AE(\hat{y}) = 1) \quad (2)$$

$$AE_{\hat{y}} = \begin{cases} 1, & \text{if } (VTA_{\hat{y}}) = 1 \wedge (MGA_{\hat{y}}) = 1 \wedge (BFA_{\hat{y}}) = 1 \\ 0, & \text{if } (VTA_{\hat{y}}) = 0 \vee (MGA_{\hat{y}}) = 0 \vee (BFA_{\hat{y}}) = 0 \end{cases} \quad (3)$$

where \widehat{TP} represents true positive cases that can be successfully applicable for CFD analysis. Specifically, $VTA_{\hat{y}} \in \{0, 1\}$, $MGA_{\hat{y}} \in \{0, 1\}$, and $BFA_{\hat{y}} \in \{0, 1\}$ represent the vascular topology availability, mesh generation availability, and blood flow availability of the segmentation mask \hat{y} , indicating, respectively, whether there are geometric topological abnormalities in the vessels, whether geometric errors occur during the conversion process that interrupt subsequent operations, and whether the generated mesh file can be successfully used for CFD analysis. The computation of all three indices above can be automated via scripts.

5 BENCHMARK DESIGN

To demonstrate the utility of the IAVS dataset, we adopt two benchmark tasks that mirror the clinical workflow from raw MRA to simulation-ready geometry: **Stage I for global localization of**

324 **aneurysms and Stage II for fine-grained IA-Vessel segmentation.** These benchmarks are designed
 325 to evaluate methods on clinically meaningful tasks. In Stage I, a global localization step identifies
 326 regions of interest containing aneurysms, setting the stage for more precise analysis. Stage II then
 327 focuses on fine-grained, topology-aware segmentation of the IA-Vessel within these localized
 328 regions. This approach is specifically designed to evaluate segmentation performance in the context
 329 of topological consistency and CFD applicability. As illustrated in Figure 4, we develop a simple and
 330 effective two-stage framework, which can be used as a out-of-the-box method and strong baseline
 331 for the benchmark.

332 **Stage I: Aneurysm Localization.** To overcome the difficulty of directly segmenting small
 333 aneurysms from full MRA volumes, we first use a detection network to pinpoint their locations.
 334 We use a counting-guided heatmap formulation to substantially reduce false positives by constraining
 335 the predicted count. Specifically, the network is designed to simultaneously predict a heatmap,
 336 indicating the probability of an aneurysm center, and a density map, estimating the number of
 337 aneurysms. The training loss is shown in Formula 4, which consists of two parts. The first part
 338 is the heatmap loss, inspired by the focal loss used for centroid prediction in Zhou et al. (2019). Due
 339 to the extreme sparsity of positive voxels (aneurysm centers), each ground truth center point is su-
 340 pervised using a 3D Gaussian heatmap t_{xyz} with a peak value of 1. To address the severe foreground-
 341 background imbalance, a weighting scheme is applied. For positive voxels ($t_{xyz} \geq 0.9$), the loss is
 342 $(1 - p_{xyz})^\alpha \cdot \log(p_{xyz})$. For all other voxels (negatives), the loss is $(1 - t_{xyz})^\beta \cdot p_{xyz}^\alpha \cdot \log(1 - p_{xyz})$. Here,
 343 p_{xyz} is the predicted heatmap value, α is a focusing parameter that down-weights easily classified
 344 examples, and the $(1 - t_{xyz})^\beta$ term for negatives places more emphasis on ambiguous regions near
 345 the Gaussian boundaries. The total heatmap loss is normalized by the number of positive voxels
 346 N_{pos} . The second part is a standard cross-entropy loss for the aneurysm count classification, where
 347 the number of aneurysms per case is treated as a classification problem with classes ranging from 0
 348 to 5.

$$\mathcal{L}_{\text{Stage I}} = -\frac{1}{N_{\text{pos}}} \underbrace{\left[\sum_{x,y,z} \begin{cases} (1 - p_{xyz})^\alpha \log(p_{xyz}) & \text{if } t_{xyz} \geq 0.9 \\ (1 - t_{xyz})^\beta p_{xyz}^\alpha \log(1 - p_{xyz}) & \text{otherwise} \end{cases} \right]}_{\text{Heatmap Loss}} + \underbrace{\mathcal{L}_{\text{CE}}(C_{\text{pred}}, C_{\text{true}})}_{\text{Count Classification Loss}}, \quad (4)$$

355 During inference, candidate center points are extracted from aggregated heatmaps and density maps.
 356 We employ a dynamic selection mechanism where the number of candidates is adaptively deter-
 357 mined by the connected components in the density map, effectively reducing false positives. This
 358 point-based detection is less sensitive to variations in aneurysm size compared to standard segmen-
 359 tation or bounding-box detection.

360 **Stage II: IA-Vessel Segmentation.** Using the center points from Stage I, we crop candidate patches
 361 to focus the segmentation task. In Stage II, we apply a topology-aware segmentation network built
 362 upon the robust nnUNet (Isensee et al., 2021) backbone. To ensure the resulting vessel geometry is
 363 suitable for CFD analysis, we incorporate a loss function that preserves vascular connectivity. As
 364 shown in Equation 5, the total loss combines a standard segmentation loss (Dice and cross-entropy)
 365 with a cIDice loss term. The cIDice component enhances the model’s sensitivity to vascular topol-
 366 ogy by explicitly supervising on centerline connectivity, which is critical for preventing spurious
 367 connections or breaks in the vessel structure.

$$\mathcal{L}_{\text{Stage II}} = -\underbrace{\frac{2 \sum_i p_i g_i}{\sum_i p_i + \sum_i g_i} + \left(-\sum_i g_i \log p_i \right)}_{\text{Segmentation Loss}} + \underbrace{\lambda \left(-\frac{2 \sum_i \mathcal{T}(p_i) \mathcal{T}(g_i)}{\sum_i \mathcal{T}(p_i) + \sum_i \mathcal{T}(g_i)} \right)}_{\text{cIDice Loss}} \quad (5)$$

373 6 EXPERIMENTS

374 We systematically evaluate the proposed framework compared with existing state-of-the-art methods
 375 on the IAVS dataset, including the evaluation of aneurysm detection for Stage I, the evaluation of
 376 IA-Vessel segmentation for Stage II, and the comprehensive evaluation of end-to-end segmentation

378 Table 4: Comparison of different strategies for aneurysm localization in Stage I.
379

380 Method	381 Set A			
	382 PR ↑	383 RE ↑	384 ACC ↑	385 F1 ↑
386 nnDetection	387 0.3737 ± 0.4122	388 0.9250 ± 0.4926	389 0.3627 ± 0.4131	390 0.5324 ± 0.4225
391 SwinUNETR	392 0.3472 ± 0.4718	393 0.6098 ± 0.5004	394 0.2841 ± 0.4644	395 0.4425 ± 0.4642
396 nnUNet	397 0.5778 ± 0.4754	398 0.6341 ± 0.4853	399 0.4333 ± 0.4668	400 0.6047 ± 0.4650
401 Ours	402 0.8286 ± 0.4182	403 0.7073 ± 0.4238	404 0.6170 ± 0.4222	405 0.7632 ± 0.4102
406 Method	407 Set B			
	408 PR ↑	409 RE ↑	410 ACC ↑	411 F1 ↑
412 nnDetection	413 0.5440 ± 0.3036	414 0.9292 ± 0.1751	415 0.5224 ± 0.3017	416 0.6863 ± 0.2389
417 SwinUNETR	418 0.5145 ± 0.3956	419 0.7807 ± 0.3792	420 0.4495 ± 0.3884	421 0.6202 ± 0.3650
422 nnUNet	423 0.6942 ± 0.4046	424 0.7368 ± 0.4033	425 0.5563 ± 0.4008	426 0.7149 ± 0.3846
427 Ours	428 0.8785 ± 0.2523	429 0.8246 ± 0.2708	430 0.7402 ± 0.2871	431 0.8507 ± 0.2491

393
394
395 with CFD applicability score. We split the test set into Set A for evaluation from multiple public
396 datasets, and Set B for evaluation in clinical scenarios from our private dataset. More experimental
397 and implementation details are shown in the Appendix B.

398 Table 5: Ablation experiments of topological-aware loss for IA-Vessel segmentation in Stage II.

400 Model	401 Topological-aware Loss	402 Set A			
		403 Dice ↑	404 HD95 ↓	405 clDice ↑	406 BIoU ↑
407 Zig-RiR	408 \times	409 0.7069 ± 0.1366	410 7.6029 ± 4.2749	411 0.6975 ± 0.1352	412 0.5626 ± 0.1568
413 nnUNet	414 \times	415 0.8533 ± 0.0840	416 3.2187 ± 2.7283	417 0.8555 ± 0.1113	418 0.7527 ± 0.1211
419 nnUNet	420 Skeleton Recall Loss	421 0.8401 ± 0.0958	422 3.5820 ± 3.4290	423 0.8447 ± 0.1269	424 0.7350 ± 0.1330
425 nnUNet	426 clDice Loss	427 0.8563 ± 0.0878	428 3.2809 ± 3.0868	429 0.8629 ± 0.1175	430 0.7576 ± 0.1214
431 Model	432 Topological-aware Loss	433 Set B			
		434 Dice ↑	435 HD95 ↓	436 clDice ↑	437 BIoU ↑
438 Zig-RiR	439 \times	440 0.7536 ± 0.1353	441 6.0273 ± 4.9289	442 0.7425 ± 0.8521	443 0.6216 ± 0.7363
444 nnUNet	445 \times	446 0.8363 ± 0.1307	447 4.2557 ± 5.6929	448 0.8538 ± 0.1502	449 0.7368 ± 0.1652
450 nnUNet	451 Skeleton Recall Loss	452 0.8296 ± 0.1452	453 4.1835 ± 5.7645	454 0.8516 ± 0.1565	455 0.7303 ± 0.1769
456 nnUNet	457 clDice Loss	458 0.8368 ± 0.1368	459 4.2134 ± 5.5734	460 0.8616 ± 0.1524	461 0.7388 ± 0.1693

438

6.1 EVALUATION OF ANEURYSM DETECTION

439
440 To address the challenge of localizing small aneurysms, we conduct a comprehensive evaluation
441 of proposed method with existing approaches. Specifically, we use three different task settings to
442 achieve the localization of aneurysms, including utilizing state-of-the-art detection model nnDetection
443 (Baumgartner et al., 2021), and segmentation models SwinUNETR (Hatamizadeh et al., 2021)
444 and nnUNet (Isensee et al., 2021), where the segmentation results is processed to generate the center
445 point of output targets. As illustrated in Table 4, our proposed method stands out with remarkable
446 performance in multiple metrics. We achieve a PR of 0.8286 and 0.8785, ACC of 0.6170 and 0.7402,
447 and F1-scores of 0.7632 and 0.8507 in Set A and Set B, respectively. Although our method exhibits a
448 slightly lower RE compared to nnDetection, our innovative dynamic candidate point selection
449 mechanism plays a crucial role. This mechanism effectively controls the false positive rate, preventing
450 the generation of an excessive number of false detections. As a result, it alleviates the computational
451 burden and complexity of subsequent processing stages, providing a more efficient and reliable
452 solution for small aneurysm localization. Although heatmap regression itself is a standard technique
453 in keypoint detection, our modification does not aim to introduce a new general localization
454 theory. Instead, to explicitly address the false-positive issue, we design a loss function that combines a
455 heatmap regression loss and a count classification loss. Compared with segmentation-based localiza-
456 tion and bounding-box detection, our design is an engineering optimization tailored to this specific
457 task to substantially reduce false positives. Overall, our method significantly outperforms the exist-
458 ing methods.

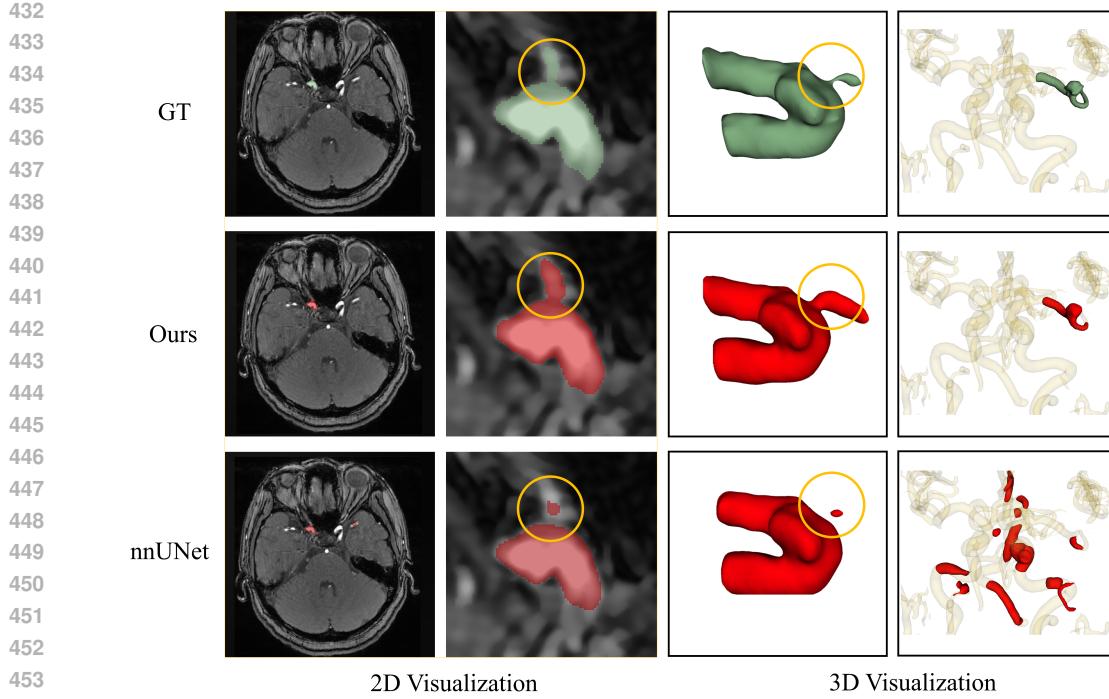


Figure 5: Visualization of IA-Vessel segmentation results of different methods.

ing detection and segmentation methods, demonstrating its strong competitiveness and potential for practical applications in medical imaging analysis.

In addition, we conduct new experiments on the publicly available GLIA-Net (Bo et al., 2021) dataset for aneurysm detection. The experimental results are presented in Appendix C. The experiments demonstrate that our method achieves optimal performance across all evaluation metrics.

6.2 EVALUATION OF IA-VESSEL SEGMENTATION

To evaluate the effectiveness of proposed topological-aware segmentation framework, we conduct ablations of cIDice loss for the segmentation. For the training of Stage II, input patches cropped based on ground truth IA are utilized for localization to avoid error accumulation in Stage I. As observed in Table 5, the results show that the introduction of cIDice loss significantly enhances the vascular topology maintenance ability, which improves the cIDice performance from 0.8555 to 0.8629 on Set A and 0.8538 to 0.8616 on Set B.

6.3 EVALUATION OF CFD APPLICABILITY

To make a comprehensive evaluation of our framework for CFD applicable IA-Vessel segmentation from MRA images, we integrate the localization results of Stage I with the segmentation procedure of Stage II to enable end-to-end segmentation. In comparison, we conduct direct end-to-end IA-Vessel segmentation using state-of-the-art nnUNet (Isensee et al., 2021) as baseline performance, and ground truth aneurysm localization for patch cropping as an upperbound comparison. As shown in Table 6, end-to-end nnUNet segmentation yields Dice coefficients of 0.1548 on Set A and 0.4557 on Set B. Due to an excessive number of false positives over 120 and fewer than 10 true positives, we conclude that the end-to-end segmentation approach is not suitable for the segmentation task. As shown in Table 7, among comparison of two-stage frameworks, we observe that our method achieve a high applicability score of 57.45% and 54.76%, significantly outperforms other comparing methods by a large margin. As shown in Figure 5, we observe that proposed method can generate mask predictions align more accurately with ground truth masks with less topologic errors and false positive predictions of background vessels.

486
487 Table 6: Performance of different methods for end-to-end IA-Vessel segmentation from MRA im-
488 ages.

489 Framework	490 Set A			
	491 Dice↑	492 HD95↓	493 cIDice↑	494 BIoU↑
495 nnUNet Baseline	0.1548 ± 0.2520	48.8495 ± 39.9534	0.1552 ± 0.2468	0.1088 ± 0.1927
496 Stage I nnDetection + Stage II	0.4285 ± 0.3753	15.9663 ± 17.8063	0.4311 ± 0.3862	0.3477 ± 0.3236
497 Stage I nnUNet + Stage II	0.4864 ± 0.4070	50.5446 ± 90.9340	0.4943 ± 0.4124	0.4174 ± 0.3686
498 Stage I Ours + Stage II	0.6324 ± 0.3630	27.8342 ± 65.1445	0.6361 ± 0.3682	0.5482 ± 0.3356
499 Stage I GT + Stage II	0.8563 ± 0.0878	3.2809 ± 3.0868	0.8629 ± 0.1175	0.7576 ± 0.1214
500 Framework	501 Set B			
	502 Dice↑	503 HD95↓	504 cIDice↑	505 BIoU↑
506 nnUNet Baseline	0.4557 ± 0.2898	37.2055 ± 17.9024	0.4323 ± 0.2853	0.3395 ± 0.2496
507 Stage I nnDetection + Stage II	0.6611 ± 0.2252	19.6505 ± 17.8079	0.6846 ± 0.2334	0.5344 ± 0.2476
508 Stage I nnUNet + Stage II	0.6186 ± 0.3499	65.7056 ± 112.9016	0.6477 ± 0.3556	0.5286 ± 0.3264
509 Stage I Ours + Stage II	0.7442 ± 0.2406	15.8149 ± 36.6947	0.7706 ± 0.2533	0.6391 ± 0.2500
510 Stage I GT + Stage II	0.8368 ± 0.1368	4.2134 ± 5.5734	0.8616 ± 0.1524	0.7388 ± 0.1693

506 Table 7: Evaluation of CFD Applicability Score of different IA-Vessel segmentation masks.

507 Framework	508 Set A					509 Set B				
	510 TP	511 FP	512 FN	513 \bar{TP}	514 AS_{CFD}	515 TP	516 FP	517 FN	518 \bar{TP}	519 AS_{CFD}
520 Stage I nnDetection + Stage II	37	62	4	30	29.13%	105	88	9	74	36.63%
521 Stage I nnUNet + Stage II	26	19	15	23	38.33%	84	37	30	65	43.05%
522 Stage I Ours + Stage II	29	6	12	27	57.45%	94	12	20	69	54.76%
523 Stage I GT + Stage II	41	0	0	35	85.37%	114	0	0	88	77.19%
524 IA-Vessel GT	41	0	0	41	100.00%	114	0	0	114	100.00%

525

7 DISCUSSION AND CONCLUSION

526 In this work, we introduce a systematic solution for CFD-applicable IA-Vessel segmentation. To
527 overcome the limitations of existing datasets, we construct IAVS, a large-scale multi-centre dataset
528 with comprehensive annotations and CFD analysis results, providing a solid foundation for sub-
529 sequent research. Our proposed two-stage framework for detection and segmentation effectively
530 reduces geometric errors and enhances the CFD usability of segmentation masks, making a break-
531 through in improving the accuracy and reliability of segmentation. Additionally, the establishment
532 of a standardized CFD applicability evaluation system, along with the introduction of the CFD appli-
533 cability score, enables a more comprehensive and standardized evaluation of segmentation results.
534 Experimental results demonstrate that our proposed method achieves a high CFD applicability score
535 of 57.45% and 54.76% on different test sets, which is significantly higher than that of existing state-
536 of-the-art methods, verifying its clinical applicability in CFD analysis, so as to assist in clinical
537 decision-making.

538 **Limitations.** Firstly, our framework employs independent training procedure of each stage, which
539 may limit further performance improvement of the model. Future work could explore an end-to-
540 end joint training mechanism. By sharing encoder-layer features and jointly optimizing the loss
541 functions, the tasks of localization and segmentation could be synergistically enhanced. Besides,
542 existing loss functions for training segmentation models primarily rely on image-based segmenta-
543 tion metrics, which have a semantic gap with the CFD applicability. Future work could focus on
544 utilizing the applicability-based evaluation for optimization of segmentation networks to enhance
545 the applicability of segmentation results for CFD applications.

546 Currently, CFD validation needs to be performed independently of the segmentation process. Fu-
547 ture research could introduce physics-informed neural networks to build an end-to-end predictive
548 model from segmentation results to hemodynamic parameters, achieving a closed-loop optimization
549 between segmentation and simulation.(Lu et al., 2019; Yao et al., 2024b).

540
541
ETHICS STATEMENT542
543
544
The authors of this paper have read and adheres to the ICLR Code of Ethics. This research involves
the use of sensitive medical data and aims for a direct clinical application; therefore, we have taken
several steps to ensure our work is conducted with the highest ethical standards.545
546
547
Human Subjects and Data Privacy: Our study utilizes 3D MRA images from both publicly available
and private, in-house clinical datasets. The collection and use of the in-house patient data were
conducted in full compliance with institutional and national ethical guidelines. The study protocol,
including data collection and anonymization procedures, received formal approval from the relevant
hospital’s Institutional Review Board (IRB) / Ethics Committee. All data were fully anonymized
prior to their use in this research, with all personally identifiable information (PII) removed to pro-
tect patient privacy and confidentiality.552
553
554
Dataset Curation and Release: We are committed to scientific transparency and reproducibility.
Upon acceptance, the curated IAVS dataset, along with our code and models, will be made publicly
available. We will ensure that the released data is thoroughly de-identified to prevent any potential
for re-identification of individuals, thereby responsibly contributing a valuable resource to the
research community while upholding our duty to protect patient privacy.555
556
557
Potential for Societal Impact and Misuse: The primary goal of this research is to contribute pos-
itively to human well-being by improving the accuracy and applicability of intracranial aneurysm
segmentation for hemodynamic analysis. This can ultimately aid clinicians in assessing aneurysm
rupture risk and making more informed treatment decisions. However, we acknowledge that any au-
tomated medical analysis tool carries the risk of misuse if not properly validated and deployed. Our
proposed framework is intended to be used as a decision-support tool to assist trained medical pro-
fessionals (such as radiologists and neurosurgeons) and is not designed to replace clinical expertise
or serve as a standalone diagnostic system.560
561
562
563
564
Bias and Fairness: Our IAVS dataset is compiled from multiple centers, which helps to mitigate
biases associated with a single institution’s population or imaging hardware. Nonetheless, the de-
mographic distribution (e.g., race, age, sex) of the patient data may not fully represent the global
population. This could potentially lead to performance disparities when the model is applied to un-
derrepresented groups. We acknowledge this as a limitation and advocate for future work to validate
and fine-tune our models on more diverse and larger-scale datasets to ensure equitable and robust
performance across all patient populations.572
573
REPRODUCIBILITY STATEMENT
574575
576
577
To ensure the reproducibility of our research, we have provided comprehensive details throughout
the paper and its appendices. The construction, annotation workflow, and statistical breakdown
of our proposed **IAVS** dataset are thoroughly described in Section 3. A detailed description of
our proposed two-stage framework, including the network architectures and loss functions for both
detection and segmentation stages, is provided in Section 5. All experimental settings, including data
preprocessing, training hyperparameters (e.g., optimizer, learning rate, patch sizes), and the specific
evaluation metrics used, are detailed in Appendix B. The procedure for our novel CFD applicability
evaluation system, which automates the conversion from segmentation masks to CFD models, is
outlined step-by-step in Section 4 and Appendix A. As stated in the abstract, we are committed to
transparency and will make our source code, the complete IAVS dataset, and the pre-trained models
publicly available upon acceptance of this manuscript to facilitate verification and further research
in the community.587
588
REFERENCES589
590
Michela Antonelli, Annika Reinke, Spyridon Bakas, Keyvan Farahani, Annette Kopp-Schneider,
Bennett A Landman, Geert Litjens, Bjoern Menze, Olaf Ronneberger, Ronald M Summers, et al.
591
The medical segmentation decathlon. *Nature communications*, 13(1):4128, 2022.592
593
Michael Baumgartner, Paul F Jäger, Fabian Isensee, and Klaus H Maier-Hein. nnDetection: a
self-configuring method for medical object detection. In *Medical Image Computing and Com-*

594 *puter Assisted Intervention–MICCAI 2021: 24th International Conference, Strasbourg, France,*
 595 *September 27–October 1, 2021, Proceedings, Part V 24*, pp. 530–539. Springer, 2021.
 596

597 Zi-Hao Bo, Hui Qiao, Chong Tian, Yuchen Guo, Wuchao Li, Tiantian Liang, Dongxue Li, Dan
 598 Liao, Xianchun Zeng, Leilei Mei, et al. Toward human intervention-free clinical diagnosis of
 599 intracranial aneurysm via deep neural network. *Patterns*, 2(2), 2021.

600 Juan R Cebral, Marcelo Adrián Castro, Sunil Appanaboyina, Christopher M Putman, Daniel Millan,
 601 and Alejandro F Frangi. Efficient pipeline for image-based patient-specific analysis of cerebral
 602 aneurysm hemodynamics: technique and sensitivity. *IEEE transactions on medical imaging*, 24
 603 (4):457–467, 2005.
 604

605 Huijun Chen, Xihai Zhao, Rui Li, Haokun Li, Haozhong Sun, Ziming Xu, Haining Wei, Yan Li,
 606 Jiaqi Dou, and Xueyan Li. Intracranial aneurysm and intracranial artery stenosis detection and
 607 segmentation challenge, 2024. URL <https://zenodo.org/records/10990482>.

608 Bowen Cheng, Ross Girshick, Piotr Dollár, Alexander C Berg, and Alexander Kirillov. Boundary
 609 IoU: Improving object-centric image segmentation evaluation. In *Proceedings of the IEEE/CVF*
 610 *conference on computer vision and pattern recognition*, pp. 15334–15342, 2021.
 611

612 Özgün Çiçek, Ahmed Abdulkadir, Soeren S Lienkamp, Thomas Brox, and Olaf Ronneberger. 3D U-
 613 Net: learning dense volumetric segmentation from sparse annotation. In *Medical Image Comput-
 614 ing and Computer-Assisted Intervention–MICCAI 2016: 19th International Conference, Athens,
 615 Greece, October 17-21, 2016, Proceedings, Part II 19*, pp. 424–432. Springer, 2016.

616 Chloe M. de Nys, Ee Shern Liang, Marita Prior, Maria A. Woodruff, James I. Novak, Ashley R. Mur-
 617 phy, Zhiyong Li, Craig D. Winter, and Mark C. Allenby. Royal brisbane TOFMRA intracranial
 618 aneurysm database, 2024.

619 Nima Etminan and Gabriel J Rinkel. Unruptured intracranial aneurysms: development, rupture and
 620 preventive management. *Nature Reviews Neurology*, 12(12):699–713, 2016.

621 Alejandro F Frangi, Wiro J Niessen, Koen L Vincken, and Max A Viergever. Multiscale vessel
 622 enhancement filtering. In *International conference on medical image computing and computer-
 623 assisted intervention*, pp. 130–137. Springer, 1998.

624 Sergios Gatidis, Tobias Hepp, Marcel Früh, Christian La Fougère, Konstantin Nikolaou, Christina
 625 Pfannenberg, Bernhard Schölkopf, Thomas Küstner, Clemens Cyran, and Daniel Rubin. A whole-
 626 body FDG-PET/CT dataset with manually annotated tumor lesions. *Scientific Data*, 9(1):601,
 627 2022.

628 Ali Hatamizadeh, Vishwesh Nath, Yucheng Tang, Dong Yang, Holger R Roth, and Daguang Xu.
 629 Swin unetr: Swin transformers for semantic segmentation of brain tumors in mri images. In
 630 *International MICCAI brainlesion workshop*, pp. 272–284. Springer, 2021.

631 Fabian Isensee, Paul F Jaeger, Simon AA Kohl, Jens Petersen, and Klaus H Maier-Hein. nnU-
 632 Net: a self-configuring method for deep learning-based biomedical image segmentation. *Nature
 633 methods*, 18(2):203–211, 2021.

634 Yuanfeng Ji, Haotian Bai, Chongjian Ge, Jie Yang, Ye Zhu, Ruimao Zhang, Zhen Li, Lingyan
 635 Zhanng, Wanling Ma, Xiang Wan, et al. Amos: A large-scale abdominal multi-organ benchmark
 636 for versatile medical image segmentation. *Advances in neural information processing systems*,
 637 35:36722–36732, 2022.

638 Rushi Jiao, Yichi Zhang, Le Ding, Bingsen Xue, Jicong Zhang, Rong Cai, and Cheng Jin. Learning
 639 with limited annotations: a survey on deep semi-supervised learning for medical image segmen-
 640 tation. *Computers in Biology and Medicine*, pp. 107840, 2023.

641 Norman Juchler, Sabine Schilling, Philippe Bijlenga, Vartan Kurtcuoglu, and Sven Hirsch. Shape
 642 trumps size: image-based morphological analysis reveals that the 3D shape discriminates intracra-
 643 nial aneurysm disease status better than aneurysm size. *Frontiers in Neurology*, 13:809391, 2022.

648 Jonas Lamy, Odyssée Merveille, Bertrand Kerautret, and Nicolas Passat. A benchmark framework
 649 for multiregion analysis of vesselness filters. *IEEE Transactions on Medical Imaging*, 41(12):
 650 3649–3662, 2022.

651

652 Wenhui Lei, Wei Xu, Kang Li, Xiaofan Zhang, and Shaoting Zhang. MedLSAM: Localize and
 653 segment anything model for 3D CT images. *Medical Image Analysis*, 99:103370, 2025.

654

655 Xigui Li, Yuanye Zhou, Feiyang Xiao, Xin Guo, Yichi Zhang, Chen Jiang, Jianchao Ge, Xiansheng
 656 Wang, Qimeng Wang, Taiwei Zhang, et al. Aneumo: A large-scale comprehensive synthetic
 657 dataset of aneurysm hemodynamics. *arXiv preprint arXiv:2501.09980*, 2025.

658

659 Lu Lu, Pengzhan Jin, and George Em Karniadakis. Deeponet: Learning nonlinear operators for iden-
 660 tifying differential equations based on the universal approximation theorem of operators. *arXiv*

661

662 preprint *arXiv:1910.03193*, 2019.

663

664 Xiangde Luo, Tao Song, Guotai Wang, Jieneng Chen, Yinan Chen, Kang Li, Dimitris N Metaxas,
 665 and Shaoting Zhang. SCPM-Net: An anchor-free 3D lung nodule detection network using sphere
 666 representation and center points matching. *Medical image analysis*, 75:102287, 2022.

667

668 Jun Ma, Yao Zhang, Song Gu, Cheng Zhu, Cheng Ge, Yichi Zhang, Xingle An, Congcong Wang,
 669 Qiyuan Wang, Xin Liu, Shucheng Cao, Qi Zhang, Shangqing Liu, Yunpeng Wang, Yuhui Li, Jian
 670 He, and Xiaoping Yang. AbdomenCT-1K: Is abdominal organ segmentation a solved problem?
 671 *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(10):6695–6714, 2022.

672

673 Paul D Morris, Andrew Narracott, Hendrik von Tengg-Kobligk, Daniel Alejandro Silva Soto, Sarah
 674 Hsiao, Angela Lungu, Paul Evans, Neil W Bressloff, Patricia V Lawford, D Rodney Hose, et al.
 675 Computational fluid dynamics modelling in cardiovascular medicine. *Heart*, 102(1):18–28, 2016.

676

677 Lei Mou, Qifeng Yan, Jinghui Lin, Yifan Zhao, Yonghuai Liu, Shaodong Ma, Jiong Zhang, Wen-
 678 hao Lv, Tao Zhou, Alejandro F Frangi, et al. COSTA: A multi-center TOF-MRA dataset and a
 679 style self-consistency network for cerebrovascular segmentation. *IEEE transactions on medical*

680

681 imaging, 2024.

682

683 Tatsat R Patel, Aakash Patel, Sriharan S Veeturi, Munjal Shah, Muhammad Waqas, Andre Mon-
 684 teiro, Ammad A Baig, Nandor Pinter, Elad I Levy, Adnan H Siddiqui, et al. Evaluating a 3D
 685 deep learning pipeline for cerebral vessel and intracranial aneurysm segmentation from computed
 686 tomography angiography–digital subtraction angiography image pairs. *Neurosurgical Focus*, 54

687

688 (6):E13, 2023.

689

690 Laurent Pierot, Christophe Portefaix, Christine Rodriguez-Régent, Sophie Gallas, Jean-François
 691 Meder, and Catherine Oppenheim. Role of MRA in the detection of intracranial aneurysm in the
 692 acute phase of subarachnoid hemorrhage. *Journal of Neuroradiology*, 40(3):204–210, 2013.

693

694 Chongyu Qu, Tiezheng Zhang, Hualin Qiao, Yucheng Tang, Alan L Yuille, Zongwei Zhou, et al.
 695 Abdomenatlas-8k: Annotating 8,000 CT volumes for multi-organ segmentation in three weeks.
 696 *Advances in Neural Information Processing Systems*, 36:36620–36636, 2023.

697

698 Wouter I Schievink. Intracranial aneurysms. *New England Journal of Medicine*, 336(1):28–40,
 699 1997.

700

701 Suprosanna Shit, Johannes C Paetzold, Anjany Sekuboyina, Ivan Ezhov, Alexander Unger, An-
 702 drey Zhylka, Josien PW Pluim, Ulrich Bauer, and Bjoern H Menze. cIDice-a novel topology-
 703 preserving loss function for tubular structure segmentation. In *Proceedings of the IEEE/CVF*

704 conference on computer vision and pattern recognition

705 pp. 16560–16569, 2021.

706

707 Tao Song, Jieneng Chen, Xiangde Luo, Yechong Huang, Xinglong Liu, Ning Huang, Yinan Chen,
 708 Zhaoxiang Ye, Huaqiang Sheng, Shaoting Zhang, et al. CPM-Net: A 3D center-points matching
 709 network for pulmonary nodule detection in CT scans. In *Medical Image Computing and Computer*

710 *Assisted Intervention–MICCAI 2020: 23rd International Conference, Lima, Peru, October 4–8,*
 711 *2020, Proceedings, Part VI 23*, pp. 550–559. Springer, 2020.

702 Kimberley M. Timmins, Irene C. van der Schaaf, Edwin Bennink, Ynte M. Ruigrok, Xingle An,
 703 Michael Baumgartner, Pascal Bourdon, Riccardo De Feo, Tommaso Di Noto, Florian Dubost,
 704 Augusto Fava-Sanches, Xue Feng, Corentin Giroud, Inteneural Group, Minghui Hu, Paul F.
 705 Jaeger, Juhana Kaiponen, Micha Klimont, Yuexiang Li, Hongwei Li, Yi Lin, Timo Loehr, Jun
 706 Ma, Klaus H. Maier-Hein, Guillaume Marie, Bjoern Menze, Jonas Richiardi, Saifeddine Rjiba,
 707 Dhaval Shah, Suprosanna Shit, Jussi Tohka, Thierry Urruty, Urszula Waliska, Xiaoping Yang,
 708 Yunqiao Yang, Yin Yin, Birgitta K. Velthuis, and Hugo J. Kuijf. Comparing methods of detect-
 709 ing and segmenting unruptured intracranial aneurysms on TOF-MRAs: The ADAM challenge.
 710 *NeuroImage*, 238:118216, 2021. ISSN 1053-8119. doi: <https://doi.org/10.1016/j.neuroimage.2021.118216>. URL <https://www.sciencedirect.com/science/article/pii/S1053811921004936>.

713 Jing Wang, Fengtao Liu, Jingjie Ge, Yimin Sun, Yilin Tang, Xiaoniu Liang, Yixin Zhao, Wenbo Yu,
 714 Jianjun Wu, Chensen Lin, et al. Nasal septum deviation correlates with asymmetry in parkinsons
 715 disease. *medRxiv*, pp. 2025–02, 2025.

716 Henry G Weller, Gavin Tabor, Hrvoje Jasak, and Christer Fureby. A tensorial approach to com-
 717 putational continuum mechanics using object-oriented techniques. *Computers in physics*, 12(6):
 718 620–631, 1998.

719 Linlin Yao, Dongdong Chen, Xiangyu Zhao, Manman Fei, Zhiyun Song, Zhong Xue, Yiqiang Zhan,
 720 Bin Song, Feng Shi, Qian Wang, et al. AASeG: Artery-aware global-to-local framework for
 721 aneurysm segmentation in head and neck CTA images. *IEEE Transactions on Medical Imaging*,
 722 2024a.

723 Tina Yao, Endrit Pajaziti, Michael Quail, Silvia Schievano, Jennifer Steeden, and Vivek Muthurangu.
 724 Image2Flow: A proof-of-concept hybrid image and graph convolutional neural network for rapid
 725 patient-specific pulmonary artery segmentation and CFD flow field calculation from 3d cardiac
 726 MRI data. *PLOS Computational Biology*, 20(6):e1012231, 2024b.

727 Xingyi Zhou, Dequan Wang, and Philipp Krähenbühl. Objects as points. *arXiv preprint*
 728 *arXiv:1904.07850*, 2019.

729
 730
 731
 732
 733
 734
 735
 736
 737
 738
 739
 740
 741
 742
 743
 744
 745
 746
 747
 748
 749
 750
 751
 752
 753
 754
 755

756 A PROCEDURE FOR CFD APPLICABILITY EVALUATION
757758
759 This study establishes a standardized workflow for transforming medical images into computational
760 fluid dynamics models consists of following steps.761 **Vascular Topology Inspection** The vascular topology of the segmentation results of intracranial
762 aneurysms and their associated vessels is first screened to detect geometric defects such as abnormal
763 adhesion, holes, indentations, and protrusions. These voxel-level segmentation errors, although not
764 affecting traditional segmentation metrics such as the Dice coefficient, can significantly impact the
765 integrity of vascular geometry and subsequently cause flow field distortions in CFD analysis. Issues
766 such as discontinuities, holes, and partial adhesions can be identified by computing Betti numbers,
767 which can detect most topological problems, while a minority of other issues still require manual
768 verification.769 **Preprocessing of Segmentation Results** Morphological optimization operations are performed using
770 3D Slicer software, including removal of stretched regions, filling of small holes, and smoothing
771 of details. A median filter with a kernel size of 1 mm is uniformly applied for surface smoothing
772 to eliminate discrete segmentation artifacts while ensuring reproducibility. The largest connected
773 component is extracted after smoothing to exclude isolated noise structures. It should be noted
774 that approximately 1% of samples may experience abnormal adhesion due to excessive smoothing,
775 which requires manual correction using the segmentation tools in 3D Slicer.776 **Conversion to Geometric Models** The voxel-represented NIFTI image data is converted into a
777 three-dimensional geometric STL model, providing a geometric basis for subsequent CFD analysis.778 **Generation of Vascular Inlet and Outlet Endpoints and Centerlines** Based on the generated
779 STL model, the VMTK toolkit is used to automatically identify the topological endpoints of vascular
780 inlets and outlets, and to generate vascular centerlines accordingly. When automatic detection
781 deviates, interactive corrections are made using 3D Slicer.782 **Cutting of Inlet and Outlet Cross-sections** The ptvista library is used to cut vascular cross-sections
783 based on the normal vectors of the centerlines. The cutting plane is uniformly set at the 1/5 end
784 position (when the candidate cutting point radius is less than 0.3 mm, it automatically retracts to a
785 proximal position that meets the radius requirement). This approach retains the complete vascular
786 structure while avoiding morphological distortion caused by excessive cutting. Experiments have
787 shown that approximately 10% of samples fail automatic cutting due to insufficient centerline length,
788 requiring manual intervention using Geomagic Wrap 2021.789 **Mesh Enhancement** Geomagic Wrap is used to perform mesh optimization processes: first, the
790 mesh doctor is used to repair non-manifold edges, self-intersections, and highly refractive edges,
791 followed by mesh re-meshing, refinement, optimization, and enhancement operations. All parameters
792 are set to the software's default values to ensure consistency in processing.793 **Fitting of Surface Geometry Files** The STL mesh file is reconstructed into a CAD model with
794 precise geometric definitions and topological relationships, i.e., a STEP format file. Based on the
795 STL mesh file, surface patches are constructed, a grid is built, and the surface is fitted to generate
796 the STEP file. The number of surface patches is set to 1000. At this point, less than 1% of the data
797 may detect intersecting grids during grid construction, which can be manually repaired by moving
798 surface patch vertices to eliminate concave polygons. For cases where the surface patches are too
799 large, the patches can be subdivided to resolve the issue.800 **Boundary Condition Annotation** The STEP model is imported into ANSYS SpaceClaim for
801 boundary condition definition, including the precise annotation of inlets, outlets, walls, and fluid
802 regions. The end faces are automatically identified using the previously generated endpoints and
803 centerline information, and the final model is saved in SCDOC format.804 **Mesh Generation** Fluent Meshing is used to generate unstructured polyhedral meshes, with mesh
805 quality and computational stability ensured through CFL number control and residual monitoring
806 mechanisms.808 **CFD Calculation** Blood flow field simulation is performed using the incompressible Newtonian
809 fluid model. The Navier-Stokes equations are solved using the icoFoam solver in Open Field
Operation and Manipulation (OpenFOAM (Weller et al., 1998) combined with the PISO algorithm,

810 calculating the velocity field, pressure field, and wall shear stress distribution under a mass flow rate
 811 range of 0.00100.0040 kg/s (only steady-state calculations are performed).

812 Prior to this, there is no fully automated workflow for converting binary segmentation masks to
 813 computational fluid dynamics models. The alternate use of multiple industrial software packages, as
 814 well as the cumbersome and repetitive nature of the operational process, significantly increases the
 815 labor and time costs associated with the annotation process. Moreover, the subjective variability in-
 816 troduced by manual cutting of vascular inlet and outlet cross-sections directly affects the objectivity of
 817 CFD-AS calculations.

819 B EXPERIMENTAL SETTINGS

820 **Implementation Details.** All of our experiments are implemented in Python with PyTorch, using an
 821 NVIDIA A100 GPU. We use the SGD optimizer with an initial learning rate of 0.01, a weight decay
 822 of 3e-5 and a momentum of 0.99 to update the network parameters with the maximum epoch number
 823 set to 1000. In Stage I, the original images are resampled to a voxel spacing of $0.34 \times 0.34 \times 0.55$
 824 mm^3 , and then are cropped into patches of size $224 \times 224 \times 48$, with the batch size set to 2. In
 825 Stage II, the patch size is $96 \times 96 \times 64$ and the batch size set to 9. During the training stage, random
 826 cropping, flipping and rotation are used to enlarge the training set and avoid over-fitting. In the
 827 inference stage, the final segmentation results are obtained using a sliding-window strategy. For
 828 other comparing methods, we follow the official implementations as in (Hatamizadeh et al., 2021;
 829 Baumgartner et al., 2021).

830 **Evaluation Metrics.** For aneurysm detection, four metrics include precision (PR), recall (RE),
 831 average precision (AP), and the F1 score are used for evaluation. These metrics are defined as
 832 follows:

$$834 \quad PR = \frac{TP}{TP + FP}$$

$$837 \quad RE = \frac{TP}{TP + FN}$$

$$839 \quad Acc = \frac{TP}{TP + FP + FN}$$

$$842 \quad F1 = 2 \times \frac{PR \times RE}{PR + RE}$$

844 where TP, FP, TN, FN represent true positive (correct detection of an aneurysm), false positive
 845 (incorrect detection of an aneurysm in a healthy case), and false negative (missed detection of an
 846 aneurysm), respectively.

847 For IA-Vessel segmentation, four metrics including the Dice similarity coefficient (Dice), 95% Haus-
 848 dorff Distance (HD95), the centerline Dice (clDice), and the boundary IoU (BIoU) are used for
 849 evaluation. These metrics are defined as follows:

$$851 \quad Dice = \frac{2 \times |A \cap B|}{|A| + |B|}$$

$$855 \quad HD95 = \inf \{d \geq 0 \mid S_A \subseteq \mathcal{N}_d(S_B) \text{ and } S_B \subseteq \mathcal{N}_d(S_A)\}$$

$$857 \quad clDice = \frac{2 \times |C_A \cap C_B|}{|C_A| + |C_B|}$$

$$860 \quad BIoU = \frac{|\partial A \cap \partial B|}{|\partial A \cup \partial B|}$$

863 where A and B represent the predicted segmentation mask and ground truth. $\mathcal{N}_d(\cdot)$ denotes the
 864 d -neighborhood around a set, and S_{pred}/S_{gt} are the predicted/ground truth boundaries. C_A/C_B and

864 $\partial A / \partial B$ represent the centerline and the boundary pixels of predicted segmentation mask and ground
 865 truth, respectively.
 866

867 C ADDITIONAL EXPERIMENTS 868

870 We further demonstrate that the heatmap cascading strategy of detection can effectively enhance seg-
 871 mentation performance. When the encoder pretrained in our framework are used to initialize the
 872 encoder of nnUNet for aneurysm segmentation, the Dice coefficient of the segmentation model in-
 873 creases from 0.4841 to 0.5676 as shown in Figure 8, confirming the position constraint effect of
 874 localization on aneurysms. Ablation experiments show that fixing the encoder of keypoint detec-
 875 tion for feature transfer outperforms the direct cascading of heatmaps (Dice: 0.5676 vs. 0.5148),
 876 indicating that semantic consistency in the feature space is crucial for segmentation accuracy.
 877

878 Table 8: Impact of keypoint detection on segmentation performance

879 Model	880 Set A				881 Set B			
	882 Dice \uparrow	883 HD95 \downarrow	884 PR \uparrow	885 RE \uparrow	886 Dice \uparrow	887 HD95 \downarrow	888 PR \uparrow	889 RE \uparrow
SwinUNETR	0.2911	28.82	0.6298	0.4358	0.4780	41.04	0.6374	0.4611
nnUNet	0.3915	43.09	0.6029	0.5056	0.5490	58.51	0.5904	0.5734
Heatmap Cascade	0.4776	48.01	0.6079	0.4736	0.5772	50.69	0.6455	0.5830
Fixed nnUNet encoder	0.3874	33.68	0.6587	0.5763	0.5765	40.32	0.6363	0.6008
Fixed keypoint encoder	0.5041	33.01	0.6901	0.5586	0.5999	32.56	0.6902	0.5945

887 To evaluate the generalization ability of our framework, we conduct extensive experiments on the
 888 publicly available GLIA-Net (Bo et al., 2021) dataset for aneurysm detection. The internal dataset
 889 includes 1338 3D CTA images/1489 IAs from 6 institutions. The external dataset includes 138 3D
 890 CTA images/101 IAs from 2 institutions. After locally retraining nnUNet and our detector, and using
 891 the publicly released segmentation weights from GLIA-Net, the experimental results summarized in
 892 Table 9 show that our method outperforms existing approaches across all evaluation metrics on the
 893 external test set A, external test set B, and internal test set as defined by the GITA-Net official split.
 894

895 D ADDITIONAL DISCUSSION 896

897 D.1 RATIONALE FOR LOCALIZED HEMODYNAMIC ANALYSIS 898

899 Our decision to focus the CFD simulations on the aneurysm and its adjacent parent vessels, rather
 900 than the entire segmented vascular network, represents a principled balance between clinical relevance
 901 and computational feasibility. This approach aligns with established standards in 3D aneurysm
 902 research, where focusing on local geometry is the mainstream methodology. This is because
 903 aneurysm rupture risk assessment primarily depends on local hemodynamic parameters such as
 904 pressure, velocity, and Wall Shear Stress (WSS) rather than the global flow dynamics of the complete
 905 network. Numerous studies have validated that such local vessel models are sufficient for
 906 revealing the key hemodynamic mechanisms essential for pathological analysis. Furthermore, simu-
 907 lating the entire cerebrovascular network, which contains thousands of arterial branches, imposes a
 908 prohibitive computational burden that typically requires weeks of supercomputing resources. Such
 909 demands render global simulations infeasible for large-scale datasets or clinical translation. By ad-
 910hering to this focused approach, we ensure that our evaluation pipeline remains computationally
 911 efficient, reducing simulation times to hours while maintaining the hemodynamic fidelity necessary
 912 for robust risk assessment and the development of physics-informed metrics.
 913

914 D.2 NUMERICAL IMPLEMENTATION AND SENSITIVITY ANALYSIS 915

916 To ensure the reliability and reproducibility of our CFD results, we adhered to strict engineering as-
 917 sumptions and conducted comprehensive sensitivity analyses regarding mesh generation and solver
 918 convergence. We utilized the PISO algorithm within the OpenFOAM framework for pressure-
 919 velocity coupling, enforcing a strict Courant-Friedrichs-Lowy (CFL) number below 1 to ensure
 920

918
919
920
921
922
923
924
925
926
927
Table 9: Comparison of the three test sets in the GLIA-Net dataset.

Metric	External test A			External test B			Internal test		
	nnUNet	GLIA-Net	Ours	nnUNet	GLIA-Net	Ours	nnUNet	GLIA-Net	Ours
PR\uparrow	0.0631	0.3008	0.3866	0.0354	0.3544	0.3981	0.0681	0.4076	0.4925
RE\uparrow	0.5200	0.7400	0.9200	0.2941	0.5490	0.8039	0.4444	0.7698	0.7857
ACC\uparrow	0.0596	0.2721	0.3740	0.0326	0.2745	0.3628	0.0628	0.3633	0.4342
F1\uparrow	0.1126	0.4277	0.5444	0.0632	0.4308	0.5325	0.1181	0.5330	0.6055

928 numerical computation stability and convergence. Convergence was defined by residuals for ve-
929 locity components (u, v, w) stabilizing at 10^{-6} and pressure (p) residuals below 10^{-6} . Furthermore,
930 we performed a grid independence study to validate our meshing strategy, testing schemes with
931 minimum element sizes ranging from 0.30 mm to 0.10 mm. Our analysis demonstrated a clear con-
932 vergence trend; specifically, refining the mesh from 0.15 mm to 0.10 mm resulted in a negligible
933 relative difference in Wall Average Shear Stress of approximately 0.25%. Consequently, we adopted
934 the 0.15 mm scheme as the standard for our dataset, confirming that our hemodynamic outputs are
935 grid-independent and not artifacts of discretization errors.

936
937

D.3 ASSUMPTIONS AND LIMITATIONS OF BOUNDARY CONDITIONS

938 While our simulation pipeline is rigorous, we acknowledge certain limitations inherent to the re-
939 spective nature of the dataset, particularly regarding patient-specific boundary conditions. Ide-
940 ally, inlet flow rates should be derived from individual phase-contrast MRI or Doppler ultrasound
941 measurements. However, since our dataset is constructed from standard non-invasive MRA scans,
942 patient-specific inlet velocity data was not available. To address this, we adopted a standardized
943 average inlet flow rate as the inlet boundary condition, combined with zero-pressure outlet and no-
944 slip wall conditions. Although this simplification prevents the precise replication of an individual
945 patient’s absolute flow dynamics, it is a necessary and widely accepted approximation in large-scale
946 computational studies where non-invasive data collection is prioritized. This approach remains ro-
947 bust for evaluating relative hemodynamic patterns and training deep learning models to learn gener-
948 alizable physics-based features.

949
950

E USE OF LARGE LANGUAGE MODELS

951 In accordance with the conference guidelines, we disclose the use of Large Language Models
952 (LLMs) in the preparation of this manuscript.

953 The role of LLMs is strictly limited to that of an assistive tool for language editing and proofreading.
954 Specifically, LLMs were utilized to improve grammar, correct spelling, refine sentence structure for
955 better clarity, and suggest alternative word choices to enhance the overall readability of the text.

956 The LLMs were not used for research ideation, generation of the core scientific content, method-
957 ology design, data analysis, or drawing conclusions. All conceptual and scientific contributions
958 presented in this paper are exclusively the work of the human authors. The authors have carefully
959 reviewed and edited all text and take full responsibility for the scientific accuracy and integrity of
960 the final manuscript.

961
962
963
964
965
966
967
968
969
970
971