

Aneumo: A Large-Scale Multimodal Aneurysm Dataset with Computational Fluid Dynamics Simulations and Deep Learning Benchmarks

Xigui Li^{1,2}, Yuanye Zhou², Feiyang Xiao^{1,2}, Xin Guo^{2*}, Chen Jiang², Tan Pan^{1,2}, Xingmeng Zhang², Cenyu Liu², Zeyun Miao^{1,2}, Jianchao Ge², Xiansheng Wang², Qimeng Wang², Yichi Zhang^{1,2}, Wenbo Zhang^{2,4}, Fengping Zhu³, Limei Han^{1,2}, Yuan Qi^{1,2}, Chensen Lin^{1,2*}, Yuan Cheng^{1,2*}

¹Artificial Intelligence Innovation and Incubation Institute, Fudan University, Shanghai, China

²Shanghai Academy of Artificial Intelligence for Science, Shanghai, China

³Huashan Hospital, Fudan University, Shanghai, China

⁴Human Phenome Institute, Fudan University, Shanghai, China

Abstract

Intracranial aneurysms (IAs) are serious cerebrovascular lesions found in approximately 5% of the general population. Their rupture may lead to high mortality. Current methods for assessing IA risk focus on morphological and patient-specific factors, but the hemodynamic influences on IA development and rupture remain unclear. While accurate for hemodynamic studies, conventional computational fluid dynamics (CFD) methods are computationally intensive, hindering their deployment in large-scale or real-time clinical applications. To address this challenge, we curated a large-scale, high-fidelity aneurysm CFD dataset to facilitate the development of efficient machine learning algorithms for such applications. Based on 427 real aneurysm geometries, we synthesized 10,660 3D shapes via controlled deformation to simulate aneurysm evolution. The authenticity of these synthetic shapes was confirmed by neurosurgeons. CFD computations were performed on each shape under eight steady-state mass flow conditions, generating a total of 85,280 blood flow dynamics data covering key parameters. Furthermore, the dataset includes segmentation masks, which can support tasks that use images, point clouds or other multimodal data as input. Additionally, we introduced a benchmark for estimating flow parameters to assess current modeling methods. This dataset aims to advance aneurysm research and promote data-driven approaches in biofluids, biomedical engineering, and clinical risk assessment. The code and dataset are available at: <https://github.com/Xigui-Li/Aneumo>.

1 Introduction

Biofluid mechanics is a key field for understanding physiological and pathological processes in living organisms [33]. Particularly in the circulatory system, the complex characteristics of blood flow and its interaction with the vessel walls (i.e., hemodynamics) profoundly influence the health and disease states of the cardiovascular and cerebrovascular systems [16]. In-depth research into hemodynamics is crucial for revealing the pathogenesis of cardiovascular and cerebrovascular diseases and guiding clinical intervention [20]. In this context, intracranial aneurysm stands out as a high-risk

*Corresponding authors: Xin Guo (guoxin@sais.com.cn), Chensen Lin (linchensen@fudan.edu.cn), and Yuan Cheng (cheng_yuan@fudan.edu.cn)

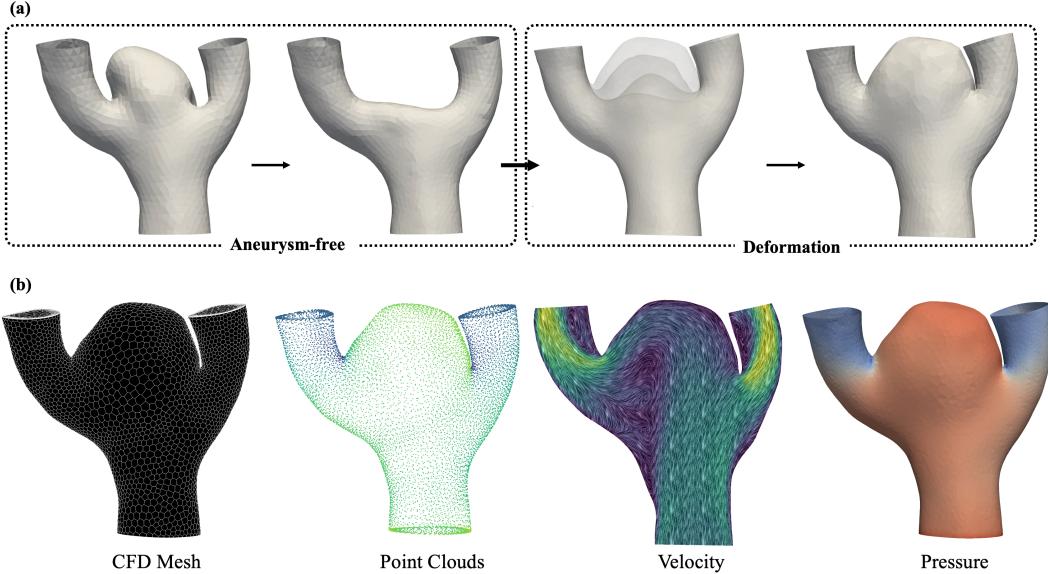


Figure 1: Workflow for deforming patient-specific aneurysm models and simulating vascular hemodynamics. (a) Patient-derived aneurysmal geometries are first processed to remove the aneurysm and recover a healthy vascular shape. Controlled geometric deformations are then applied to generate synthetic aneurysm models. (b) CFD meshes are created for the deformed geometries, followed by simulations of blood flow velocity and pressure fields for hemodynamic analysis.

cerebrovascular disease with a particularly strong link to hemodynamics. The formation, growth, and rupture risk of aneurysms are significantly influenced by the local blood flow environment. Although most IA are asymptomatic when detected, rupture leads to severe subarachnoid hemorrhage, resulting in extremely high mortality and disability rates [70]. Over the past decades, considerable effort has been devoted to elucidating the mechanisms of IA formation, development, and rupture [31]. Currently, clinical assessment of IA rupture risk primarily relies on patient-specific characteristics (such as age, sex, medical history, etc.) and morphological features of the aneurysm [12, 30, 57]. However, despite hemodynamic factors being theoretically considered crucial, their complex mechanisms are not yet fully elucidated, and related quantitative analysis methods (such as traditional computational fluid dynamics) have high barriers to application and are computationally expensive. Consequently, the value of hemodynamic information in clinical risk assessment has not been fully realized [61, 53].

Data-driven approaches can significantly reduce the time and resources required for traditional performance evaluation processes [9]. Traditionally, the process involves extracting vascular structures from medical images, constructing a three-dimensional (3D) model, generating a computational fluid dynamics mesh, setting boundary conditions, performing CFD solving, and performing post-processing [3]. By integrating and optimizing these tedious steps, the data-driven approach not only dramatically improves computational efficiency, but also accelerates the entire simulation process. This enables researchers and clinical experts to obtain accurate performance estimates in real time, which can then be used to explore the intrinsic connection between hemodynamic features and pathological changes, and ultimately drive the development of personalized diagnosis and precision treatment strategies [23]. Meanwhile, the latest machine learning advances [10, 8, 67] have demonstrated their ability to rapidly extract key flow field features from CFD simulation data, enabling efficient prediction of complex blood flow behaviors, and thus facilitating real-time and accuracy improvements in performance estimation. However, due to the lack of large-scale publicly available datasets, these methods are currently mostly focused on simple problems, limiting their generalization ability and application in a wider range of clinical scenarios.

Currently, most intracranial aneurysm datasets primarily focus on imaging data [38, 24, 59], while those in 3D modeling, especially integrating hemodynamic parameters [40, 68], remain limited. This hampers the in-depth study of hemodynamic mechanisms and restricts a more comprehensive understanding of the lesion. Although recent studies [54] have provided some data through steady-state blood flow analysis, these methods struggle to capture the complex and variable blood flow

environment and dynamic physiological characteristics due to the small sample size. In contrast, openly shared, large-scale multimodal datasets in fields like computer vision (e.g., ImageNet [22], COCO [44], LAION-5B[52], and KITTI [34]) have driven significant technological progress.

This study presents a more comprehensive dataset to address multiple challenges in aneurysm research. The dataset is based on a real aneurysm geometry model of AneuX [40], and high-quality 3D synthetic models covering the geometric evolution of the aneurysm at different stages of the aneurysm are generated by controlled deformation techniques. In addition, we provide segmentation masks similar to medical images and integrate hemodynamic parameters obtained from simulations under eight physiological flow conditions. This dataset lays a solid foundation for data-driven aneurysm modeling and analysis, and is expected to advance accurate simulation and prediction studies of intracranial aneurysms.

Our study makes the following contributions:

- **The First Large-Scale High-Fidelity Hemodynamic Dataset for Aneurysm Machine Learning:** We present 85,280 hemodynamic data samples (velocity/pressure fields) generated through CFD simulations under various physiological flow conditions (0.001–0.004 kg/s). This dataset bridges a critical gap in AI-driven hemodynamic modeling.
- **Diverse Aneurysm Geometry Evolution Data:** Based on 427 real aneurysm geometries, we generate 10,660 high-quality 3D models using controlled deformation techniques. These models comprehensively capture the geometric evolution of aneurysms at different stages, enabling quantitative modeling of rupture risks.
- **Multimodal Data Fusion Framework:** The dataset includes high-resolution binary segmentation mask images precisely aligned with CFD parameters . This enables multimodal learning tasks and facilitates multi-scale feature mapping in complex flow environments.

2 Related Work

The integration of computational fluid dynamics and deep learning is gaining traction in scientific research, with high-fidelity, standardized datasets serving as a foundation for methodological advances[29]. These datasets facilitate benchmarking and comparative evaluation of emerging algorithms. In recent years, several benchmarks have been introduced to capture key fluid dynamics phenomena. For instance, AirfRANS [7] provides RANS solutions for airfoil optimization; BubbleML [36] offers experimental data for multiphase flows and interface tracking; the Lagrangian Benchmark [60] includes particle-resolved simulations for turbulent diffusion; and BLASTNet [15] addresses multiphysics simulations involving reactive flows.

In addition, domain-specific datasets are increasingly being developed to meet industry needs. In aerospace, AircraftVerse [17] provides 27,714 parameterized aircraft designs with aerodynamic metrics, enabling generative design evaluation and efficient exploration of high-dimensional spaces via micro-geometric parameterization. In the automotive field, DrivAerNet++ [29] offers 8,000 high-fidelity CFD simulations with detailed pressure and velocity data, supporting advanced aerodynamic analysis.

In contrast to recent advancements in various fields such as computer vision and natural language processing, biomedical applications face a significant gap. In the specific study of cerebral aneurysms, hemodynamic modeling is fundamentally constrained by the lack of large-scale, multimodal datasets [41]. Specifically, such datasets should encompass advanced precise segmentation masks, and accurately reconstructed three-dimensional vascular geometries, in addition to high-resolution outputs derived from computational fluid dynamics simulations. This critical limitation fundamentally impedes progress in the application of advanced data-driven techniques such as generative modeling for data augmentation, the development of efficient surrogate models for rapid analysis , the acceleration of computationally intensive CFD simulations using machine learning, and ultimately, the accurate prediction of aneurysm rupture risk [43, 64]. The prohibitively high computational cost associated with traditional vascular simulations, which often require 1 to 4 hours per case even when executed on high-performance computing (HPC) clusters, further exacerbates this existing challenge and hinders the creation of the necessary large datasets[74].

The comparison in Table 1 underscores the significant limitations of existing cerebral aneurysm datasets, which impede advancements in computational modeling, predictive analysis, and multimodal

Table 1: Overview of intracranial aneurysm datasets, highlighting key attributes including dataset scale, publication year, segmentation masks (suitable for generating 3D anatomical models), 3D geometric models, hemodynamic simulations, multimodal data (M : segmentation masks, PC : point clouds, $3D$: geometric models, C : CFD hemodynamic data), and availability as open-source resources. The symbol (🔒) indicates restricted access requiring registration or approval. * Indicates availability of 3D geometry without a CFD-compatible mesh, limiting direct use in computational simulations. †The dataset include hemodynamic simulations but only report global metrics (e.g., total pressure drop), without releasing field-level data (e.g., pressure or velocity fields in VTK/NumPy format), rendering them unsuitable for machine learning or data-driven analyses.

Dataset	Year	3D Model	Hemodynamic Data	Mask	Modalities	Open Source
ADAM [59]	2020	✗	✗	210	M	🔒
CHUV [25]	2022	✗	✗	157	M	✓
INSTED [13]	2022	✗	✗	64	M	🔒
RWS-MT [11]	2024	✗	✗	70	M	✓
GLIA-Net [6]	2021	✗	✗	1212	M	🔒
AneuX [40] *	2021	750	✗	✗	$3D$	✓
CADA [38] *	2020	131	✗	131	$M,3D$	🔒
RBWH [21] *	2024	63	✗	63	$M,3D$	🔒
IntrA [68]	2020	215	✗	✗	$M,3D,PC$	✓
Aneurisk [2] †	2012	103	8	23	$M,PC,3D,C$	✓
CMHA [54] †	2024	105	105	105	$M,PC,3D,C$	✓
Aneumo(Ours)	2025	10660	85280	10660	$M,PC,3D,C$	✓

research. While datasets such as ADAM [59], CHUV [25], and INSTED [13] provide segmentation masks, they lack critical components such as 3D models and hemodynamic parameters. These missing modalities are essential for constructing comprehensive computational frameworks and validating predictive models. Similarly, datasets like AneuX [40], which focus exclusively on 3D models, offer 750 samples but fail to include segmentation masks or CFD data, limiting their applicability for multimodal and integrative analyses. Datasets such as CADA [38] and RBWH [21] attempt to bridge this gap by including both segmentation masks and 3D models. However, the absence of hemodynamic parameters, which are indispensable for understanding blood flow dynamics and assessing aneurysm rupture risk, restricts their utility for advanced modeling. Even larger datasets, such as GLIA-Net [6], which contains 1212 segmentation masks, and IntrA [68], which includes 215 3D models and point clouds, fail to provide CFD data. This omission significantly limits their potential for high-fidelity simulations and predictive modeling tasks. Aneurisk [2] and CMHA [54] are among the few datasets that integrate segmentation masks, 3D models, point clouds, and CFD data. However, their limited scale, with only 23 and 105 segmentation masks respectively, renders them insufficient for large-scale data-driven research and machine learning applications. Furthermore, access restrictions in datasets such as ADAM [59] and GLIA-Net [6] introduce additional barriers, further constraining their usability in collaborative and open research environments.

These constraints underscore the urgent need for a high-quality, large-scale multimodal dataset that integrates segmentation masks, 3D geometric models, point clouds, and high-fidelity hemodynamic parameters. However, existing datasets often suffer from limited modality coverage, insufficient size, and incomplete annotations, falling short of the requirements for computational modeling, predictive analytics, and multimodal exploration in emerging research. The lack of such resources has long hampered our ability to fully elucidate the complex interplay among aneurysm morphology, hemodynamics, and rupture risk. To address this critical gap, we introduce **Aneumo**, a meticulously curated large-scale dataset designed to advance research in this domain. This dataset integrates rich segmentation masks, detailed 3D geometric models, dense point clouds, and high-resolution

hemodynamic parameters, offering unprecedented breadth and depth. By enabling comprehensive multimodal investigations, it lays the foundation for robust computational frameworks and accelerates progress in understanding the complex interplay of factors driving aneurysm behavior.

3 Dataset Presentation

Due to the extreme scarcity of comprehensive datasets required for intracranial aneurysm research, developing robust machine learning models presents significant challenges. To address this limitation, we leverage the AneuX[40] dataset to generate a diverse set of baseline aneurysm models. Each baseline model undergoes at least 20 randomized non-rigid stretching transformations to emulate the natural morphological variability of aneurysms. These transformations cover a wide range of shapes and sizes, ultimately forming a large-scale dataset comprising 10,660 unique 3D models. To achieve multimodal data representation, we convert high-fidelity 3D models into segmentation masks (serving as regions of interest (ROIs) rather than actual medical images) to enrich the multimodal dataset. Furthermore, under 8 steady-state physiological conditions, we perform CFD simulations for each deformed model, computing critical hemodynamic parameters such as pressure and velocity fields. This integrative dataset, combining geometric morphology with rich hemodynamic information, provides a multidimensional foundation for advancing the understanding of aneurysm initiation, progression, and treatment strategies, offering significant potential for both academic research and clinical applications. The dataset construction process and implementation details are described in the following sections, and the corresponding illustration is shown in Figure 1.

3.1 Geometry and Mask

Based on the real 3D model, we performed deaneurysm and deformation operations, including removal of the aneurysm region, stochastic deformation, and geometric optimization to ensure topological integrity and physiological plausibility. Important steps throughout the process were scrutinized by physicians and fluid mechanics experts to ensure scientific validity and data reliability. The processed model is then converted into a segmentation mask similar to medical imaging data, which enables spatial alignment of the hemodynamic simulation results with the imaging data. This step ensures spatial consistency and supports the integrated analysis of geometric and hemodynamic information. The operation steps and implementation are detailed in the supplementary material.

3.2 Computational Mesh

In order to fully analyze the hemodynamic characteristics of intracranial aneurysms, high-precision computational fluid dynamics simulations were performed on the processed 3D model. First, boundary conditions including the inlet, outlet, wall and fluid regions were defined to accurately reflect the physiological environment of the aneurysm. Subsequently, the fluid domain was discretized using an unstructured polyhedral mesh, a mesh form that excels in handling complex geometries and details of the boundary layer width high accuracy and computational efficiency [56]. Based on the mesh sensitivity analysis, the minimum mesh size of 0.15 mm was finally selected to achieve the best balance between numerical accuracy and computational cost. To ensure the reliability and scientific validity of the simulation results, the dataset generation process undergoes strict quality control, including boundary condition validation by domain experts and sensitivity analysis of multiple grid sizes. The implementation details are provided in the supplementary material.

3.3 Numerical Schemes

Boundary Conditions In this study, blood was modeled as an incompressible Newtonian fluid with a density of 1050 kg/m³ and a dynamic viscosity of 0.00345 Pa·s to represent the laminar flow behavior within an intracranial aneurysm [28]. The blood vessel wall, inlet, and outlet were defined as rigid no-slip [39], mass flow inlet [45], and zero-pressure outlet boundary conditions [18], respectively. The mass flow rates for steady-state simulations ranged from 0.0010 to 0.0040 kg/s, including values of 0.0010, 0.0015, 0.0020, 0.0025, 0.0030, 0.003, 0.00375, and 0.0040 kg/s.

CFD Simulation We use OpenFOAM v2312 [66] as a computational tool to solve the Navier-Stokes equations for a 3D incompressible Newtonian fluid based on the finite volume method (FVM).

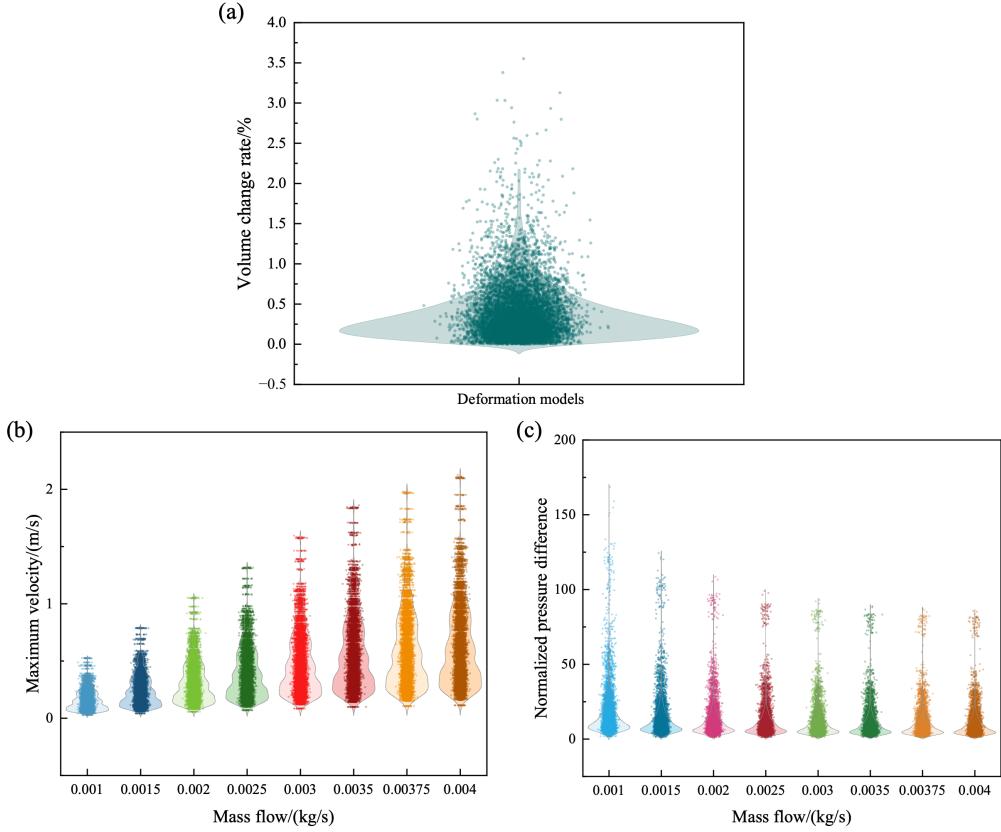


Figure 2: Simulation results of aneurysm deformation and hemodynamics under different flow rates, including (a) Volume change rate distribution, (b) Maximum blood flow velocity distribution, and (c) Normalized differential pressure distribution.

In the simulations, the `icoFoam` solver is employed, and the PISO (Pressure Implicit with Splitting of Operators) [37] algorithm is used to achieve highly accurate pressure–velocity coupling. During all numerical simulations, the Courant–Friedrichs–Lowy (CFL) number [19] is strictly controlled to be less than 1 to ensure numerical stability and accuracy. A total of 85,280 CFD simulations were completed for each 3D model in the Anuemo dataset. The simulation results generated key hemodynamic parameters, including velocity components (u, v, w) and pressure (p), which are important for studying the fluid dynamics inside intracranial aneurysms. These metrics not only contribute to an in-depth understanding of the hemodynamic environment within the aneurysm, but also provide an important basis for assessing its formation, progression, and rupture risk.

Computational Cost The high-fidelity CFD simulation of the Aneumo dataset consumes a large amount of computational resources. The simulation is run on the Fudan CFFF intelligent supercomputing platform, utilizing MPI [26] parallelization acceleration. Each CFD case is computed using 24 CPU cores, with the computation time for a single case ranging from about 45 minutes to 2 hours. Simultaneously, Slurm [69] and Kubernetes [58] are used to realize massively parallel computation of multiple CFD cases, with a maximum parallel processing capacity of 10,000 CPU cores. The complete dataset (including log files, etc.) occupies 23.1 TB of storage space, and the total number of CFD files generated reaches 7.12×10^6 . The processed dataset made publicly available is approximately 4 TB. The entire simulation process took about 2×10^6 CPU hours to complete.

3.4 Data Analysis

Volume Change Rate Firstly, to further assess the reasonableness of the geometric deformation, we analyzed the distribution of the volume change rate after deformation. The volume change rate is defined as $\Delta V = \frac{V_{\text{def}} - V_{\text{ref}}}{V_{\text{ref}}}$, where V_{def} denotes the volume of the deformed model and V_{ref} represents

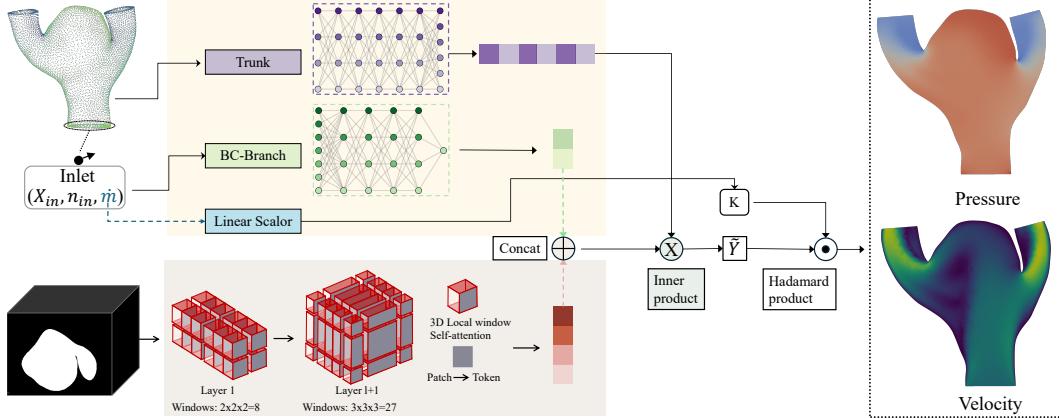


Figure 3: Schematic illustration of the DeepONet-SwinT model architecture for predicting aneurysm hemodynamic parameters.

the reference volume of the aneurysm-free model. As shown in Figure 2(a), the volume change rate was mainly distributed between 0 and 1, with only a small number of samples exceeding this range. This result is consistent with the theoretical expectation, indicating that the volume change rate is generally within a reasonable range, and the geometric volume change is effectively controlled within the expected physical constraints. However, a few anomalies were found to have volume change rates that significantly exceeded 1, and in extreme cases even reached 3.5. Further analysis showed that these anomalies were mainly due to the significant expansion of the aneurysm region after multiple stretching manipulations, whereas the vascular region connected to it was relatively small, thus amplifying the overall volume change rate.

Max Velocity To further validate the reliability of the CFD data, we analyze the distribution of the maximum velocity (V_{\max}) across all cases under different mass flow rates. The maximum velocity is defined as $V_{\max} = \max(\sqrt{u^2 + v^2 + w^2})$. As shown in Figure 2(b), the maximum velocity values in the dataset are typically distributed between 0 and 2 m/s, and are mainly concentrated in the range of 0 to 1 m/s. The figure also reveals a clear trend: as the mass flow rate increases, the maximum velocity also increases, and the range of the velocity distribution broadens significantly. This expansion in the distribution range reflects the increasing complexity of the flow field at higher flow rates, indicating enhanced flow instability and variability in velocity characteristics. The wider range of observed velocities highlights the increasing diversity and complexity of dynamic behaviors within the flow field as flow rates increase.

Normalized Differential Pressure Similarly, Figure 2(c) illustrates the distribution of normalized pressure differences (ΔP^*) under different flow conditions. The normalized pressure is defined as $P^* = \frac{P}{(1/2)*\rho*V_{\max}^2}$, where P is the pressure and ρ is the fluid density. The corresponding normalized pressure difference is $\Delta P^* = P_{\max}^* - P_{\min}^*$. This normalization process eliminates the direct influence of flow rate and velocity on absolute pressure values, thereby highlighting the relative variations in pressure more clearly. The results demonstrate that, across different flow conditions, the normalized pressure difference for nearly all cases falls within the range of 0 to 150, with the majority concentrated between 0 and 75. Moreover, the distribution density is notably higher in the lower range, particularly between 0 and 50, indicating that the relative pressure variations within the flow field are generally moderate under most conditions. Additionally, as the inlet mass flow rate increases, the distribution of the normalized pressure difference tends to shift toward the 0-60 range, suggesting a more uniform pressure distribution under higher flow conditions. This characteristic of the normalized pressure distribution provides intuitive insights into the flow field dynamics and further validates the reliability and accuracy of the CFD simulation results.

4 Benchmark for Aneurysm Hemodynamics

Accurate and efficient prediction of blood flow hemodynamics within aneurysms using Scientific Machine Learning (SciML) methods is critical for assessing their development and potential rupture

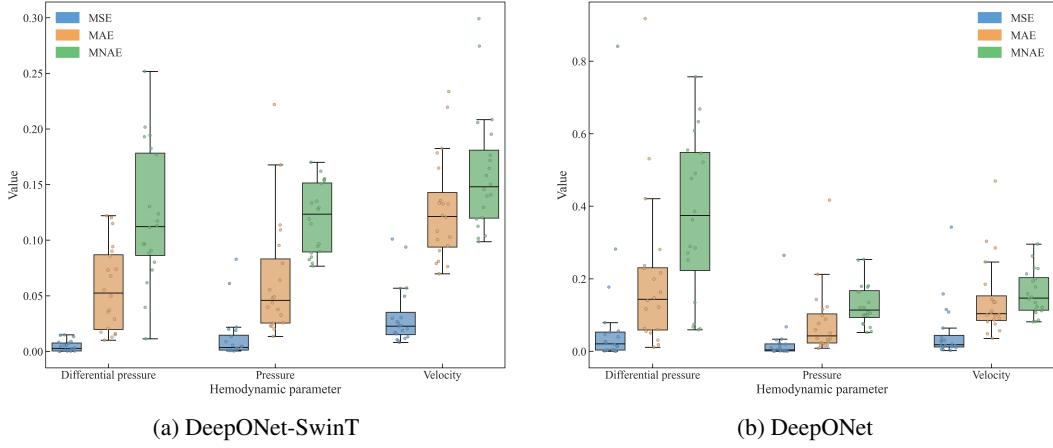


Figure 4: Performance comparison of DeepONet and DeepONet-SwinT on the test set.

risk. To advance research in this area, we propose a dedicated benchmark for this task. On this benchmark, we systematically compare the performance of the standard DeepONet architecture [47] against our developed DeepONet-SwinT model. The architecture of the DeepONet-SwinT model, illustrated in Figure 3, innovatively combines the operator learning paradigm of DeepONet with the strengths of Swin Transformer [46] in capturing complex spatial features, aiming to leverage multi-modal inputs more effectively for high-precision prediction of key hemodynamic parameters: pressure, velocity, and pressure difference (Δp). The benchmark dataset is rigorously constructed, with all data ensuring geometric independence across training, validation, and test sets. Both the training and validation sets are sourced from the Aneumo dataset: the training set includes 1280 CFD simulation cases from 160 different aneurysm geometries under 8 distinct flow conditions; the validation set includes 80 CFD cases from 40 independent geometries (non-overlapping with the training set) under two typical flow conditions. To strictly evaluate generalization capability on unseen data, the final test set is derived from the completely independent real aneurysm dataset, Aneux [40], including 20 CFD cases from 10 geometries under high and low flows. Details of the training procedure and parameters are provided in the Appendix.

On the test set, we employ Mean Normalized Absolute Error (MNAE), Mean Squared Error (MSE), and Mean Absolute Error (MAE) as final evaluation metrics[49, 32, 29]. Results presented in Figure 4 and the visualized comparison in Figure 5 clearly indicate that DeepONet-SwinT exhibits a significantly lower median error and a more concentrated error distribution across all predicted hemodynamic parameters. As visually demonstrated in Figure 5, DeepONet-SwinT’s predictions for both pressure and velocity fields show a higher degree of fidelity to the CFD ground truth compared to the standard DeepONet, particularly in capturing intricate flow features. These findings powerfully demonstrate DeepONet-SwinT’s superior prediction accuracy and robustness, particularly on unseen real data. In terms of computational efficiency, while DeepONet-SwinT incurs a higher training cost (approx. 88 hours for 5000 epochs on four NVIDIA A100 80GB GPUs, compared to about 3 hours for DeepONet), both models show extremely high efficiency in single-GPU inference time (DeepONet \sim 0.002 s/case, DeepONet-SwinT \sim 0.01 s/case), satisfying the requirements for rapid inference in practical applications.

5 Conclusion

In this study, a new large-scale open-source multimodal computational fluid dynamics dataset is constructed based on 427 real cerebral aneurysm geometries with the aim of advancing the application of data-driven methods in cerebral aneurysm hemodynamic studies. The dataset covers eight different flow conditions and simulates the evolution of aneurysms from small to large by systematically and stochastically deforming the aneurysm geometry to capture key hemodynamic features. High-fidelity CFD simulations ensure the accuracy and reliability of the dataset. The dataset is released in a common open-source format and contains 10,660 aneurysm geometry models and their corresponding segmentation mask binary image files, along with 85,280 CFD data sets generated under eight time-averaged flow conditions.

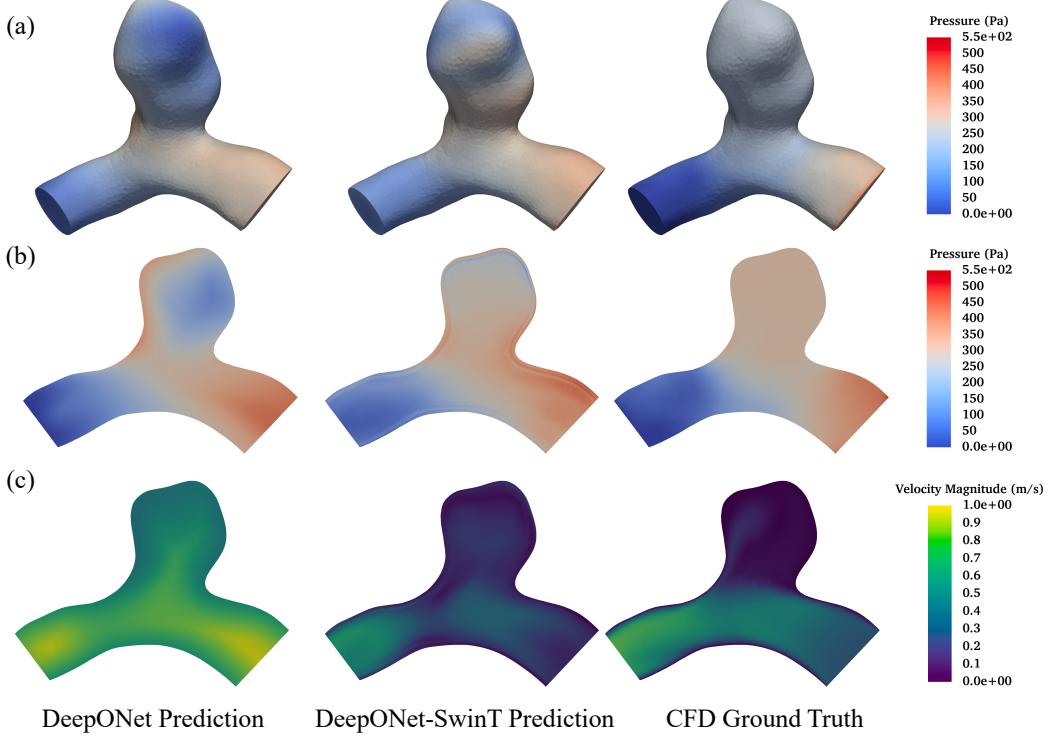


Figure 5: Comparison of predicted pressure and velocity fields by DeepONet and DeepONet-SwinT with CFD ground truth, including (a) Wall pressure contour plots, (b) Internal pressure contour plots, and (c) Internal velocity contour plots.

This dataset provides extensive support for a variety of machine learning tasks, including alternative modeling of hemodynamic performance, acceleration of CFD simulations, data-driven optimization, generative artificial intelligence, and 3D shape reconstruction. Initial benchmarking validates the effectiveness of the DeepONet model and the DeepONet-SwinT framework in pressure and velocity prediction tasks. We hope that this open-source dataset will accelerate the development of data-driven versus physically-driven machine learning methods in the field of hemodynamic prediction of cerebral aneurysms, provide an important resource to support related research, and open up new possibilities for future research explorations.

6 Limitations and Future Work

Although the high-precision, large-scale dataset constructed in this study has significant advantages over existing work, it also has limitations. The dataset contains only steady-state CFD simulations, which do not capture dynamic blood flow characteristics during the cardiac cycle, thus limiting the study of aneurysm dynamic behavior. While the DeepONet-SwinT framework that we developed using this dataset demonstrated promising initial performance in pressure and velocity prediction tasks, current models may not be sufficient to fully capture the fine-grained interactions between complex geometries and hemodynamic features. Therefore, this dataset also provides an opportunity for the community to encourage the development and application of more advanced models that can fully characterize these complex hemodynamic properties.

Future work will introduce transient CFD simulations to more fully characterize dynamic blood flow properties and extend the dataset’s geometric diversity and multimodal information (e.g., vessel wall biomechanical properties) to support more complex machine learning tasks. These enhancements will further increase the scientific value of the dataset and provide enhanced support for aneurysm research and clinical diagnosis.

Acknowledgments and Disclosure of Funding

This paper was supported by the National Natural Science Foundation of China (Grant No. 82394432 and 92249302), the Shanghai Municipal Science and Technology Major Project (Grant No. 2023SHZDZX02). The computations in this research were performed using the CFFF platform of Fudan University.

References

- [1] M. AL-Rawi, D. Belkacemi, and A. M. Al-Jumaily. Mesh independency analysis for aorta geometry using a computational modelling approach. In *ASME International Mechanical Engineering Congress and Exposition*, volume 87622, page V005T06A040. American Society of Mechanical Engineers, 2023.
- [2] Aneurisk-Team. AneuriskWeb project website, <http://ecm2.mathcs.emory.edu/aneuriskweb>. Web Site, 2012.
- [3] L. Antiga, M. Piccinelli, L. Botti, B. Ene-Iordache, A. Remuzzi, and D. A. Steinman. An image-based modeling framework for patient-specific computational hemodynamics. *Medical & biological engineering & computing*, 46:1097–1112, 2008.
- [4] M. Auer and T. C. Gasser. Reconstruction and finite element mesh generation of abdominal aortic aneurysms from computerized tomography angiography data with minimal user interactions. *IEEE transactions on medical imaging*, 29(4):1022–1028, 2010.
- [5] O. Aycan, A. Topuz, and L. Kadem. Evaluating uncertainties in cfd simulations of patient-specific aorta models using grid convergence index method. *Mechanics Research Communications*, 133:104188, 2023.
- [6] Z.-H. Bo, H. Qiao, C. Tian, Y. Guo, W. Li, T. Liang, D. Li, D. Liao, X. Zeng, L. Mei, et al. Toward human intervention-free clinical diagnosis of intracranial aneurysm via deep neural network. *Patterns*, 2(2):100197, 2021.
- [7] F. Bonnet, J. Mazari, P. Cinnella, and P. Gallinari. Airfrans: High fidelity computational fluid dynamics dataset for approximating reynolds-averaged navier–stokes solutions. *Advances in Neural Information Processing Systems*, 35:23463–23478, 2022.
- [8] K. A. Boster, S. Cai, A. Ladrón-de Guevara, J. Sun, X. Zheng, T. Du, J. H. Thomas, M. Nedergaard, G. E. Karniadakis, and D. H. Kelley. Artificial intelligence velocimetry reveals in vivo flow rates, pressure gradients, and shear stresses in murine perivascular flows. *Proceedings of the National Academy of Sciences*, 120(14):e2217744120, 2023.
- [9] S. L. Brunton, B. R. Noack, and P. Koumoutsakos. Machine learning for fluid mechanics. *Annual review of fluid mechanics*, 52(1):477–508, 2020.
- [10] S. Cai, H. Li, F. Zheng, F. Kong, M. Dao, G. E. Karniadakis, and S. Suresh. Artificial intelligence velocimetry and microaneurysm-on-a-chip for three-dimensional analysis of blood flow in physiology and disease. *Proceedings of the National Academy of Sciences*, 118(13):e2100697118, 2021.
- [11] W. Cao, X. Chen, J. Lv, L. Shao, and W. Si. Semi-supervised intracranial aneurysm segmentation via reliable weight selection. *The Visual Computer*, pages 1–13, 2024.
- [12] N. Chalouhi, B. L. Hoh, and D. Hasan. Review of cerebral aneurysm formation, growth, and rupture. *Stroke*, 44(12):3613–3622, 2013.
- [13] H. Chen, X. Zhao, R. Li, H. Li, H. Sun, Z. Xu, H. Wei, Y. Li, J. Dou, and X. Li. Intracranial aneurysm and intracranial artery stenosis detection and segmentation challenge, Apr. 2024.
- [14] H. Chen, X. Zhou, Z. Feng, and S.-J. Cao. Application of polyhedral meshing strategy in indoor environment simulation: Model accuracy and computing time. *Indoor and Built Environment*, 31(3):719–731, 2022.
- [15] W. T. Chung, B. Akoush, P. Sharma, A. Tamkin, K. S. Jung, J. Chen, J. Guo, D. Brouzet, M. Talei, B. Savard, et al. Turbulence in focus: Benchmarking scaling behavior of 3d volumetric super-resolution with blastnet 2.0 data. *Advances in Neural Information Processing Systems*, 36:77430–77484, 2023.
- [16] J. A. Claassen, D. H. Thijssen, R. B. Panerai, and F. M. Faraci. Regulation of cerebral blood flow in humans: physiology and clinical implications of autoregulation. *Physiological reviews*, 101(4):1487–1559, 2021.

- [17] A. Cobb, A. Roy, D. Elenius, F. Heim, B. Swenson, S. Whittington, J. Walker, T. Bapty, J. Hite, K. Ramani, et al. Aircraftverse: a large-scale multimodal dataset of aerial vehicle designs. *Advances in Neural Information Processing Systems*, 36:44524–44543, 2023.
- [18] M. Conijn and G. J. Krings. Computational analysis of the pulmonary arteries in congenital heart disease: a review of the methods and results. *Computational and mathematical methods in medicine*, 2021(1):2618625, 2021.
- [19] R. Courant, K. Friedrichs, and H. Lewy. On the partial difference equations of mathematical physics. *IBM journal of Research and Development*, 11(2):215–234, 1967.
- [20] E. Cuadrado-Godia, P. Dwivedi, S. Sharma, A. O. Santiago, J. R. Gonzalez, M. Balcells, J. Laird, M. Turk, H. S. Suri, A. Nicolaides, et al. Cerebral small vessel disease: a review focusing on pathophysiology, biomarkers, and machine learning strategies. *Journal of stroke*, 20(3):302, 2018.
- [21] C. M. de Nys, E. S. Liang, M. Prior, M. A. Woodruff, J. I. Novak, A. R. Murphy, Z. Li, C. D. Winter, and M. C. Allenby. time-of-flight mra of intracranial aneurysms with interval surveillance, clinical segmentation and annotations. *Scientific Data*, 11(1):555, 2024.
- [22] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei. Imagenet: A large-scale hierarchical image database. In *2009 IEEE conference on computer vision and pattern recognition*, pages 248–255. Ieee, 2009.
- [23] D. Dey, P. J. Slomka, P. Leeson, D. Comaniciu, S. Shrestha, P. P. Sengupta, and T. H. Marwick. Artificial intelligence in cardiovascular imaging: Jacc state-of-the-art review. *Journal of the American College of Cardiology*, 73(11):1317–1335, 2019.
- [24] T. Di Noto, G. Marie, S. Tourbier, Y. Alemán-Gómez, O. Esteban, G. Saliou, M. B. Cuadra, P. Hagmann, and J. Richiardi. Towards automated brain aneurysm detection in tof-mra: open data, weak labels, and anatomical knowledge. *Neuroinformatics*, 21(1):21–34, 2023.
- [25] T. Di Noto, G. Marie, S. Tourbier, Y. Alemán-Gómez, O. Esteban, G. Saliou, M. B. Cuadra, P. Hagmann, and J. Richiardi. Towards automated brain aneurysm detection in tof-mra: open data, weak labels, and anatomical knowledge. *Neuroinformatics*, 21(1):21–34, 2023.
- [26] O. M. Documentation. *mpirun / mpiexec*, 2024. Accessed: 2024-05-26.
- [27] C. S. Drapaca. Poiseuille flow of a non-local non-newtonian fluid with wall slip: A first step in modeling cerebral microaneurysms. *Fractal and fractional*, 2(1):9, 2018.
- [28] G. Duan, N. Lv, J. Yin, J. Xu, B. Hong, Y. Xu, J. Liu, and Q. Huang. Morphological and hemodynamic analysis of posterior communicating artery aneurysms prone to rupture: a matched case–control study. *Journal of NeuroInterventional Surgery*, 8(1):47–51, 2016.
- [29] M. Elrefaei, F. Morar, A. Dai, and F. Ahmed. Drivaernet++: A large-scale multimodal car dataset with computational fluid dynamics simulations and deep learning benchmarks. *Advances in Neural Information Processing Systems*, 37:499–536, 2024.
- [30] N. Etminan, K. Beseoglu, D. L. Barrow, J. Bederson, R. D. Brown Jr, E. S. Connolly Jr, C. P. Derdeyn, D. Hänggi, D. Hasan, S. Juvela, et al. Multidisciplinary consensus on assessment of unruptured intracranial aneurysms: proposal of an international research group. *Stroke*, 45(5):1523–1530, 2014.
- [31] N. Etminan and G. J. Rinkel. Unruptured intracranial aneurysms: development, rupture and preventive management. *Nature Reviews Neurology*, 12(12):699–713, 2016.
- [32] E. Ferdian, D. Marlevi, J. Schollenberger, M. Aristova, E. R. Edelman, S. Schnell, C. A. Figueroa, D. Nordsletten, and A. A. Young. Cerebrovascular super-resolution 4d flow mri—sequential combination of resolution enhancement by deep learning and physics-informed image processing to non-invasively quantify intracranial velocity, flow, and relative pressure. *Medical Image Analysis*, 88:102831, 2023.
- [33] Y.-c. Fung. *Biomechanics: mechanical properties of living tissues*. Springer Science & Business Media, 2013.
- [34] A. Geiger, P. Lenz, and R. Urtasun. Are we ready for autonomous driving? the kitti vision benchmark suite. In *2012 IEEE conference on computer vision and pattern recognition*, pages 3354–3361. IEEE, 2012.
- [35] F. J. Gijsen, F. N. van de Vosse, and J. Janssen. The influence of the non-newtonian properties of blood on the flow in large arteries: steady flow in a carotid bifurcation model. *Journal of biomechanics*, 32(6):601–608, 1999.

- [36] S. M. S. Hassan, A. Feeney, A. Dhruv, J. Kim, Y. Suh, J. Ryu, Y. Won, and A. Chandramowlishwaran. Bubbleml: A multiphase multiphysics dataset and benchmarks for machine learning. *Advances in Neural Information Processing Systems*, 36:418–449, 2023.
- [37] R. I. Issa. Solution of the implicitly discretised fluid flow equations by operator-splitting. *Journal of computational physics*, 62(1):40–65, 1986.
- [38] M. Ivantsits, L. Goubergrits, J.-M. Kuhnigk, M. Huellebrand, J. Bruening, T. Kossen, B. Pfahringer, J. Schaller, A. Spuler, T. Kuehne, et al. Detection and analysis of cerebral aneurysms based on x-ray rotational angiography-the cada 2020 challenge. *Medical image analysis*, 77:102333, 2022.
- [39] F. Joly, G. Soulez, D. Garcia, S. Lessard, and C. Kauffmann. Flow stagnation volume and abdominal aortic aneurysm growth: Insights from patient-specific computational flow dynamics of lagrangian-coherent structures. *Computers in biology and medicine*, 92:98–109, 2018.
- [40] N. Juchler, S. Schilling, P. Bijlenga, V. Kurtcuoglu, and S. Hirsch. Shape trumps size: image-based morphological analysis reveals that the 3d shape discriminates intracranial aneurysm disease status better than aneurysm size. *Frontiers in Neurology*, 13:809391, 2022.
- [41] M. Kadem, L. Garber, M. Abdelkhalek, B. K. Al-Khzraji, and Z. Keshavarz-Motamed. Hemodynamic modeling, medical imaging, and machine learning and their applications to cardiovascular interventions. *IEEE Reviews in Biomedical Engineering*, 16:403–423, 2022.
- [42] Y. Karnam, F. Mut, A. K. Yu, B. Cheng, S. Amin-Hanjani, F. T. Charbel, H. H. Woo, M. Niemelä, R. Tulamo, B. R. Jahromi, et al. Description of the local hemodynamic environment in intracranial aneurysm wall subdivisions. *International journal for numerical methods in biomedical engineering*, 40(8):e3844, 2024.
- [43] D. Kochkov, J. A. Smith, A. Alieva, Q. Wang, M. P. Brenner, and S. Hoyer. Machine learning-accelerated computational fluid dynamics. *Proceedings of the National Academy of Sciences*, 118(21):e2101784118, 2021.
- [44] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick. Microsoft coco: Common objects in context. In *Computer vision–ECCV 2014: 13th European conference, zurich, Switzerland, September 6–12, 2014, proceedings, part v 13*, pages 740–755. Springer, 2014.
- [45] Y. Liu, L. Cai, Y. Chen, and Q. Chen. Icpinn: Integral conservation physics-informed neural networks based on adaptive activation functions for 3d blood flow simulations. *Computer Physics Communications*, 311:109569, 2025.
- [46] Z. Liu, Y. Lin, Y. Cao, H. Hu, Y. Wei, Z. Zhang, S. Lin, and B. Guo. Swin transformer: Hierarchical vision transformer using shifted windows. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 10012–10022, 2021.
- [47] L. Lu, P. Jin, G. Pang, Z. Zhang, and G. E. Karniadakis. Learning nonlinear operators via deeponet based on the universal approximation theorem of operators. *Nature machine intelligence*, 3(3):218–229, 2021.
- [48] A. Marzo, P. Singh, I. Larrabide, A. Radaelli, S. Coley, M. Gwilliam, I. D. Wilkinson, P. Lawford, P. Reymond, U. Patel, et al. Computational hemodynamics in cerebral aneurysms: the effects of modeled versus measured boundary conditions. *Annals of biomedical engineering*, 39:884–896, 2011.
- [49] E. Pajaziti, J. Montalt-Tordera, C. Capelli, R. Sivera, E. Sauvage, M. Quail, S. Schievano, and V. Muthurangu. Shape-driven deep neural networks for fast acquisition of aortic 3d pressure and velocity flow fields. *PLoS Computational Biology*, 19(4):e1011055, 2023.
- [50] P. K. Paritala, H. Anbananthan, J. Hautaniemi, M. Smith, A. George, M. Allenby, J. B. Mendieta, J. Wang, L. MacLachlan, E. Liang, et al. Reproducibility of the computational fluid dynamic analysis of a cerebral aneurysm monitored over a decade. *Scientific Reports*, 13(1):219, 2023.
- [51] Y. Qiu, J. Wang, J. Zhao, T. Wang, T. Zheng, and D. Yuan. Association between blood flow pattern and rupture risk of abdominal aortic aneurysm based on computational fluid dynamics. *European Journal of Vascular and Endovascular Surgery*, 64(2-3):155–164, 2022.
- [52] C. Schuhmann, R. Beaumont, R. Vencu, C. Gordon, R. Wightman, M. Cherti, T. Coombes, A. Katta, C. Mullis, M. Wortsman, et al. Laion-5b: An open large-scale dataset for training next generation image-text models. *Advances in neural information processing systems*, 35:25278–25294, 2022.
- [53] D. M. Sforza, C. M. Putman, and J. R. Cebral. Hemodynamics of cerebral aneurysms. *Annual review of fluid mechanics*, 41(1):91–107, 2009.

- [54] M. Song, S. Wang, Q. Qian, Y. Zhou, Y. Luo, and X. Gong. Intracranial aneurysm cta images and 3d models dataset with clinical morphological and hemodynamic data. *Scientific data*, 11(1):1213, 2024.
- [55] M. Spiegel, T. Redel, Y. J. Zhang, T. Struffert, J. Hornegger, R. G. Grossman, A. Doerfler, and C. Karmonik. Tetrahedral and polyhedral mesh evaluation for cerebral hemodynamic simulation—a comparison. In *2009 Annual international conference of the IEEE engineering in medicine and biology society*, pages 2787–2790. IEEE, 2009.
- [56] M. Spiegel, T. Redel, Y. J. Zhang, T. Struffert, J. Hornegger, R. G. Grossman, A. Doerfler, and C. Karmonik. Tetrahedral vs. polyhedral mesh size evaluation on flow velocity and wall shear stress for cerebral hemodynamic simulation. *Computer methods in biomechanics and biomedical engineering*, 14(01):9–22, 2011.
- [57] T. Steiner, S. Juvela, A. Unterberg, C. Jung, M. Forsting, and G. Rinkel. European stroke organization guidelines for the management of intracranial aneurysms and subarachnoid haemorrhage. *Cerebrovascular diseases*, 35(2):93–112, 2013.
- [58] The Kubernetes Authors. Kubernetes - production-grade container orchestration, 2025. Accessed: 2025-05-11.
- [59] K. M. Timmins, I. C. van der Schaaf, E. Bennink, Y. M. Ruigrok, X. An, M. Baumgartner, P. Bourdon, R. De Feo, T. Di Noto, F. Dubost, et al. Comparing methods of detecting and segmenting unruptured intracranial aneurysms on tof-mras: The adam challenge. *Neuroimage*, 238:118216, 2021.
- [60] A. Toshev, G. Galletti, F. Fritz, S. Adami, and N. Adams. Lagrangebench: A lagrangian fluid mechanics benchmarking suite. *Advances in Neural Information Processing Systems*, 36:64857–64884, 2023.
- [61] A. S. Turjman, F. Turjman, and E. R. Edelman. Role of fluid dynamics and inflammation in intracranial aneurysm formation. *Circulation*, 129(3):373–382, 2014.
- [62] K. Valen-Sendstad, M. Piccinelli, R. KrishnankuttyRema, and D. A. Steinman. Estimation of inlet flow rates for image-based aneurysm cfd models: where and how to begin? *Annals of biomedical engineering*, 43:1422–1431, 2015.
- [63] P. Venugopal, D. Valentino, H. Schmitt, J. P. Villablanca, F. Vinuela, and G. Duckwiler. Sensitivity of patient-specific numerical simulation of cerebral aneurysm hemodynamics to inflow boundary conditions. *Journal of neurosurgery*, 106(6):1051–1060, 2007.
- [64] R. Vinuesa and S. L. Brunton. Enhancing computational fluid dynamics with machine learning. *Nature Computational Science*, 2(6):358–366, 2022.
- [65] A. M. Vukicevic, S. Çimen, N. Jagic, G. Jovicic, A. F. Frangi, and N. Filipovic. Three-dimensional reconstruction and nurbs-based structured meshing of coronary arteries from the conventional x-ray angiography projection images. *Scientific reports*, 8(1):1711, 2018.
- [66] H. G. Weller, G. Tabor, H. Jasak, and C. Fureby. A tensorial approach to computational continuum mechanics using object-oriented techniques. *Computers in physics*, 12(6):620–631, 1998.
- [67] K. Wiśniewski, P. Reorowicz, Z. Tyfa, B. Price, A. Jian, A. Fahlström, D. Obidowski, D. J. Jaskólski, K. Jóźwik, K. Drummond, et al. Computational fluid dynamics; a new diagnostic tool in giant intracerebral aneurysm treatment. *Computers in Biology and Medicine*, 181:109053, 2024.
- [68] X. Yang, D. Xia, T. Kin, and T. Igarashi. Intra: 3d intracranial aneurysm dataset for deep learning. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 2656–2666, 2020.
- [69] A. B. Yoo, M. A. Jette, and M. Grondona. Slurm: Simple linux utility for resource management. In *Workshop on job scheduling strategies for parallel processing*, pages 44–60. Springer, 2003.
- [70] M. Zakeri, A. Atef, M. Aziznia, and A. Jafari. A comprehensive investigation of morphological features responsible for cerebral aneurysm rupture using machine learning. *Scientific Reports*, 14(1):15777, 2024.
- [71] J.-m. Zhan, T.-d. Lu, Z.-y. Yang, W.-q. Hu, and W. Su. Influence of the flow field and vortex structure of patient-specific abdominal aortic aneurysm with intraluminal thrombus on the arterial wall. *Engineering Applications of Computational Fluid Mechanics*, 16(1):2100–2122, 2022.
- [72] Y. Zhang, Y. Bazilevs, S. Goswami, C. L. Bajaj, and T. J. Hughes. Patient-specific vascular nurbs modeling for isogeometric analysis of blood flow. *Computer methods in applied mechanics and engineering*, 196(29-30):2943–2959, 2007.

- [73] M. Zhao, S. Amin-Hanjani, S. Ruland, A. Curcio, L. Ostergren, and F. Charbel. Regional cerebral blood flow using quantitative mr angiography. *American Journal of Neuroradiology*, 28(8):1470–1473, 2007.
- [74] S. Zheng, J. Wenbin, and W. Shuyan. A recurrent neural network for computational physiology and its application to intracoronary image-derived fractional flow reserve. *Engineering Applications of Artificial Intelligence*, 146:110309, 2025.

A Aneumo Dataset Generation

A.1 Geometry and Mask Generation

3D Model Deformation To systematically expand our dataset and generate a diverse collection of synthetic aneurysm cases based on real patient anatomy, we developed a specific processing pipeline targeting original 3D vascular models derived from clinical data. This pipeline’s core process comprises two principal stages: the precise removal of the real aneurysm region from the original model, and the subsequent introduction of a synthetic aneurysm deformation at that specific anatomical location.

The initial stage involves the import and preprocessing of the real aneurysm’s 3D model. The model, typically represented as a triangular mesh, is imported into Geomagic Wrap 2021 software for processing. Upon import, a crucial step is to leverage the integrated Mesh Doctor tool to execute comprehensive quality inspection and automated repair of the geometric mesh. This procedure is designed to address various common mesh defects that could impede subsequent processing and numerical simulations, including, but not limited to, non-manifold edges, self-intersections, acute spikes, small tunnels, and holes. Ensuring the mesh’s topological integrity and overall quality at this juncture is paramount for obtaining a physically viable model suitable for subsequent deformation and high-fidelity numerical simulations, such as CFD. Following mesh repair, the original, patient-specific aneurysm region is precisely identified and digitally removed from the vascular model. An initial local repair and surface smoothing operation, typically employing the Ramesh function within the software, is then applied to the remaining vascular wall segment encircling the site of aneurysm removal, preparing the area for the introduction of the synthetic geometry.

Subsequent to the removal of the aneurysm region and initial local processing, the synthetic aneurysm deformation is implemented on the vascular wall corresponding to the previously removed area. This deformation is actualized by selecting the continuous surface region and applying the software’s polygon offset command. This command generates a new surface by displacing the selected polygons along their local normal direction (either positive or negative) by a specified distance. To introduce morphological variability, this offset distance, denoted as d , is randomly sampled from the interval $[0.5, 1.0]$ (with units consistent with the original model’s measurement system). A critical aspect of this operation, particularly for preserving surface continuity and mesh quality, is the automatic generation of a transition band around the perimeter of the deformed area. This band comprises additional triangular mesh elements engineered to facilitate a smooth blend between the synthetically deformed surface and the surrounding original vascular wall. To optimize the mesh quality within this crucial blending zone and effectively mitigate the creation of sharp geometric transitions, which can detrimentally affect numerical stability and accuracy in simulations, the width of this inserted polygon transition band feature was explicitly set to its maximum permissible value within the software. Following the synthetic deformation, a final, global optimization pass, typically utilizing the Ramesh function again, is executed on the entire updated 3D vascular model to further enhance overall mesh smoothness and quality prior to export. Finally, the optimized 3D vascular model containing the synthetic aneurysm is exported in STL format, a widely supported standard suitable for subsequent analysis and processing tools.

3D Model Conversion to Segmentation Mask For subsequent multi-modal data analysis and, specifically, to facilitate the integrated analysis of hemodynamic simulation outcomes with medical imaging data, the exported STL format 3D vascular geometry model is converted into a NII.GZ format segmentation mask utilizing the open-source medical image processing software 3D Slicer. Within the 3D Slicer environment, the imported 3D surface model undergoes a voxelization process, transforming the continuous surface representation into a discrete label map volume data structure. In the resulting label map, voxels situated within the vascular model’s boundaries are assigned a value of 1, while voxels representing the background are assigned a value of 0. This voxelized segmentation

output is subsequently saved as a standard medical imaging format file, specifically a NII.GZ file. This conversion is pivotal: it enables the precise mapping and registration of computationally derived data, such as CFD simulation results (e.g., velocity fields, pressure distributions), onto this unified image space. This capability facilitates intuitive visualization and quantitative analysis of hemodynamic parameters directly within the imaging context. Furthermore, transforming the 3D geometric information into a medical image-compatible segmentation mask format establishes a foundational basis for conducting advanced cross-modal data analysis and constructing sophisticated machine learning models capable of effectively integrating geometric features, simulated hemodynamic parameters, and imaging characteristics.

The entirety of this processing pipeline, especially the critical steps involving the judgment and removal of the aneurysm region, the design and implementation of the synthetic aneurysm deformation, and the subsequent conversion from 3D model to segmentation mask, has undergone rigorous review and validation by clinical physicians and fluid dynamics experts. This validation process ensures the physiological plausibility of the synthetic models and the scientific reliability of the data processing methodology.

A.2 CFD Simulation Preprocessing and Mesh Generation

Geometric Model Processing The high-quality initial three-dimensional vascular models were first imported into Geomagic Wrap 2021 software for subsequent geometric refinement and optimization. The core technical step at this stage involved leveraging its powerful Fit Surface function to convert the original discrete point cloud or polygonal mesh models into a set of high-precision Non-Uniform Rational B-Spline (NURBS) surfaces. This process establishes the model's accurate Boundary Representation (B-rep). NURBS, as an industry-standard mathematical surface representation, is renowned for its excellent geometric fidelity, inherent continuity and smoothness, and precise local control over curvature [72]. This transformation from a discrete representation to a precise boundary representation substantially enhances the continuity and expressive accuracy of the vascular lumen geometry. This is critically important for accurately imposing boundary conditions on the wall and precisely calculating local surface normal vectors, thereby laying a solid foundation for obtaining physically accurate fluid dynamics simulation results [65].

Following the construction of the high-precision NURBS surface models, the corresponding geometric data was exported as STEP format files. The purpose of using the STEP format is to ensure the precise and lossless exchange of geometric data between different Computer-Aided Design (CAD) and Computer-Aided Engineering (CAE) software platforms. This format guarantees the integrity and accuracy of the model's boundary description, providing a reliable and consistent geometric input for subsequent computational domain definition, boundary condition setup, and mesh generation.

Computational Domain and Boundary Condition Setup The optimized STEP format geometric models were imported into the ANSYS SpaceClaim software environment for processing. Within this computational preprocessing platform, we first precisely extracted the internal volume of the vascular lumen as the computational fluid domain. Furthermore, the exterior boundary surfaces of this fluid domain were functionally partitioned and named, clearly identifying the fluid inflow interface (Inlet), fluid outflow interfaces (Outlet), and the fixed boundary in contact with the vessel wall (Wall). The internal volume of the fluid domain itself was designated as the Fluid Region.

Accurately defining the computational fluid domain and its boundary conditions is the foremost prerequisite for obtaining physically realistic and credible CFD simulation results. Given that the AneuX [40] dataset used in this study does not include original imaging directly associated with hemodynamic measurement data (e.g., Phase-Contrast Magnetic Resonance Imaging), it is therefore difficult to directly determine the specific inlet vascular interface. In such cases lacking complete flow data, existing studies in computational fluid dynamics simulations commonly adopt methods based on 3D model geometric features, particularly selecting the vascular cross-section with the largest cross-sectional area as the inlet and applying idealized velocity or flow boundary conditions on this section [62, 48, 63]. Following this common method, we determined the inlet interface based on the morphological characteristics of the vascular tree: through meticulous anatomical examination and diameter measurements, the cross-sectional plane with the largest area among the main vessels connected to the aneurysm was designated as the fluid inlet. This purely morphology-driven inlet selection strategy underwent independent evaluation and cross-validation by a seasoned clinical

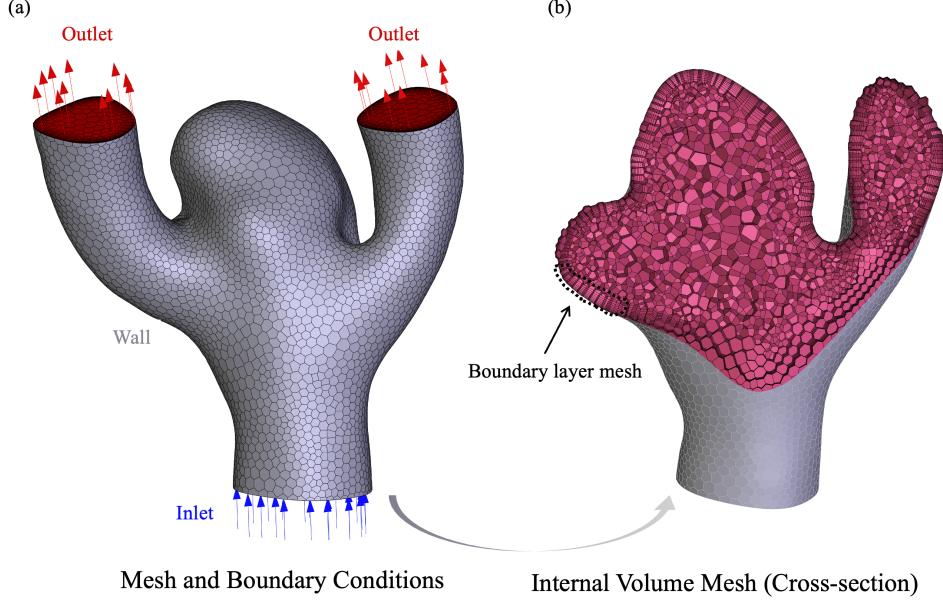


Figure 6: Visualization of computational mesh structure and boundary conditions for intracranial aneurysm models.

medical expert and an experienced computational fluid dynamics specialist to ensure its anatomical plausibility and hemodynamic representativeness. All remaining open ends of the vascular tree, other than the designated inlet interface, were set as fluid outflow interfaces. Upon completion of the computational domain construction and precise boundary condition setup, the model was saved in SCDOC format files.

Meshing Strategy Constructing a high-quality computational mesh is a critical step in ensuring both the efficiency and accuracy of CFD simulations. Given the complex geometric features of the intracranial aneurysm vascular system, we utilized ANSYS Fluent Meshing 2023 R1 to generate an unstructured polyhedral mesh for the defined fluid domain. Polyhedral elements, due to their high connectivity (i.e., each vertex is connected to significantly more neighboring cells than in tetrahedral meshes), facilitate rapid and uniform propagation of numerical information across the computational domain. This property significantly reduces numerical diffusion errors and enhances the convergence and stability of velocity and pressure field solutions [55, 5, 56]. Compared to commonly used tetrahedral meshes, polyhedral meshes offer greater flexibility in capturing local curvature variations, and typically require 30%–60% fewer elements to achieve comparable accuracy, thereby reducing memory usage and computational time [56, 5, 14]. In contrast to the limitations of structured meshes in handling complex bifurcations and highly tortuous segments, unstructured polyhedral meshes exhibit superior topological adaptability and greater robustness when simulating vascular bifurcations, tortuous segments, and aneurysm sacs [1, 4].

To accurately capture the sharp velocity gradient that forms near the vessel wall due to the no-slip boundary condition, transitioning from zero velocity at the wall to the core flow velocity, we generated specialized Boundary Layer Mesh (Inflation Layers) in the region immediately adjacent to the vessel wall, as shown in Figure 6(b). These boundary layer meshes are typically composed of multiple layers of high-quality prismatic cells extruded along the wall normal direction. Specifically, we generated a total of 10 boundary layers on the vessel wall surface. The thickness of each boundary layer cell increased outwards from the wall in the normal direction with a constant growth rate of 1.2[50]. This layered, progressive mesh refinement strategy ensures sufficiently fine mesh resolution in the near-wall region, enabling accurate resolution of the fluid's velocity profile. This is critically important for precisely calculating the velocity field and pressure distribution within the vascular lumen and is a necessary technical guarantee for obtaining high-accuracy fluid dynamics simulation results.

Mesh Sensitivity Analysis To rigorously verify the independence of the numerical simulation results from the mesh resolution, effectively exclude numerical discretization error as the dominant source of error in the analysis results, and determine the optimal mesh parameter configuration before large-scale computation execution, we conducted a systematic Mesh Sensitivity Analysis.

As illustrated in Figure 7, by systematically varying the global Minimum Element Size parameter, we constructed and tested five computational domain discretization schemes with significantly different mesh densities. The tested minimum mesh size gradients were 0.25 mm, 0.20 mm, 0.15 mm, 0.10 mm, and 0.05 mm. The corresponding total fluid domain cell counts were approximately 150,000, 250,000, 330,000, 530,000, and 880,000, respectively. For each mesh density, the same physical model setup and boundary conditions were applied for simulation solving. By comparing and analyzing key hemodynamic indicators obtained at different mesh resolutions, including the velocity distribution characteristics inside and at the neck of the aneurysm, the velocity profiles in the main channels, and the volume-averaged pressure within the aneurysm lumen, we quantitatively evaluated the degree of influence of mesh density on the simulation results.

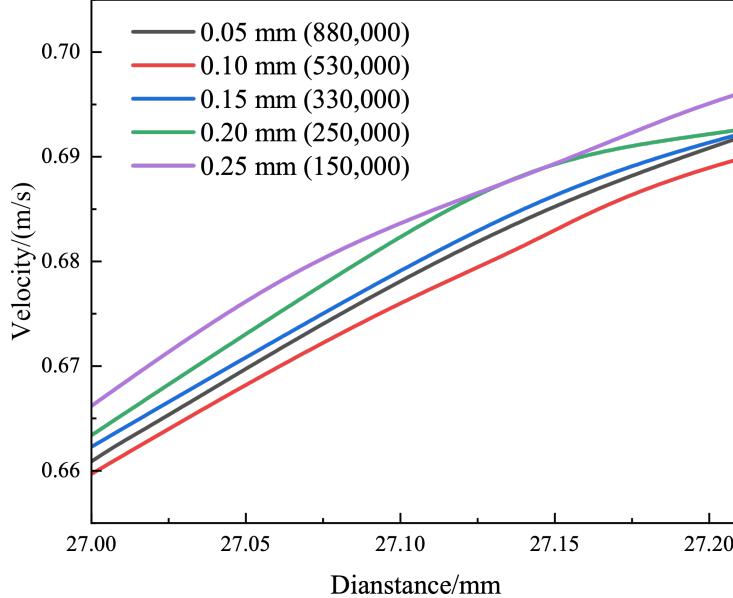


Figure 7: Grid independence analysis.

The analysis results clearly revealed that as the mesh size was progressively reduced (i.e., mesh resolution increased), the calculated values of the monitored flow field parameters exhibited significant convergence characteristics, indicating that the numerical solution was asymptotically approaching the true physical solution. A key finding was that when the minimum mesh size reached 0.15 mm, the primary flow field parameter results obtained were in excellent agreement with those obtained using the finer 0.10 mm mesh. Quantitative comparison showed that the relative difference between the two sets of results, for the monitored core flow field parameters, was below 0.01%. This result clearly indicates that at a mesh resolution of 0.15 mm, the numerical solution had essentially reached a state of mesh convergence, and further increasing the mesh density had reduced the impact on the final results to a negligible level. Considering the balance between the sufficient computational accuracy verified at this resolution and the required computational resources and efficiency, we determined 0.15 mm to be the optimal global minimum size configuration for generating meshes for all subsequent vascular models in this study.

A.3 Numerical Simulation

This study employs Computational Fluid Dynamics methods to simulate the hemodynamic characteristics within intracranial aneurysms. The numerical scheme adopted, including the governing equations, fluid model, boundary condition settings, solver configuration, and convergence criteria, is detailed below.

Governing Equations Blood flow is treated as a three-dimensional, incompressible Newtonian fluid, governed by the Navier-Stokes equations, which include the continuity equation (mass conservation) and the momentum conservation equation. For an incompressible fluid, the continuity equation is simplified to:

$$\nabla \cdot \mathbf{u} = 0 \quad (1)$$

Where \mathbf{u} represents the velocity vector. The momentum conservation equation is expressed as:

$$\rho \left(\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla) \mathbf{u} \right) = -\nabla p + \mu \nabla^2 \mathbf{u} \quad (2)$$

Where, ρ is the fluid density, t is time, p is pressure, and μ is the dynamic viscosity of the fluid. The left side of the equation represents the inertia terms, including local acceleration and convective acceleration terms; the right side includes the pressure gradient term and the viscous diffusion term. This study focuses on steady-state hemodynamic results, which are obtained by running transient simulations for a sufficiently long duration until the flow field reaches a stable state, at which point $\frac{\partial \mathbf{u}}{\partial t} \approx 0$.

Fluid Constitutive Model and Flow Characteristics In this study, blood is modeled as an incompressible Newtonian fluid with a density $\rho = 1050 \text{ kg/m}^3$ and a dynamic viscosity $\mu = 0.00345 \text{ Pa} \cdot \text{s}$ [33, 50]. These parameters are widely accepted physiological averages. Although blood can exhibit non-Newtonian properties at low shear rates, the Newtonian fluid model is a reasonable simplification for larger blood vessels such as intracranial arteries, especially under the flow conditions set in this study where high shear rates prevail [35]. This model effectively captures the main hemodynamic features while avoiding the additional complexity of non-Newtonian models. The constant mass flow rate at the inlet is set in the range of 0.001 kg/s to 0.004 kg/s. This range is intended to ensure that the blood flow within the aneurysm remains in the laminar regime.

Boundary Condition Settings To accurately simulate the hemodynamic environment within the aneurysm, specific physical conditions were applied to the boundaries of the computational domain. The vessel walls, including the aneurysm sac, were defined as rigid no-slip boundaries, meaning the fluid velocity at the wall is $\mathbf{u} = (0, 0, 0)$. This setup, which neglects the elastic deformation of the vessel wall, is a common simplification in hemodynamic simulations, suitable for studies focusing on blood flow shear stress and pressure distribution [27, 42].

At the inlet cross-section of the computational domain, a mass flow rate inlet condition was applied. The specific constant mass flow rates imposed include 0.0010 kg/s, 0.0015 kg/s, 0.0020 kg/s, 0.0025 kg/s, 0.0030 kg/s, 0.00375 kg/s, and 0.0040 kg/s. The choice of constant flow rates over more physiologically realistic pulsatile flow was primarily based on the following considerations: first, unsteady inlet profiles significantly increase data dimensionality, leading to excessively large datasets for subsequent model training; second, patient-specific, precise inlet velocity data are not routinely available in clinical practice and their profiles vary with individuals and aneurysm location [3, 73]. Therefore, this study utilizes multiple discrete steady flow rates for simulations to represent, to some extent, different physiological conditions or different phases within a cardiac cycle.

At the outlet cross-section of the computational domain, a zero relative pressure outlet condition was set, i.e., $p = 0 \text{ Pa}$. This is a standard outlet treatment that allows the fluid to exit the domain naturally in a fully developed state and minimizes adverse effects of the outlet boundary on the upstream flow field in the aneurysm region [51, 71].

Numerical Solution Method and Simulation Parameters All CFD simulations were performed using the open-source software platform OpenFOAM (Open Field Operation and Manipulation) v2312[66]. The governing equations were discretized using the Finite Volume Method. The icoFoam solver was selected to solve the Navier-Stokes equations for three-dimensional incompressible Newtonian fluid flow. This solver is designed for unsteady, incompressible, laminar flow problems. Although the objective is to obtain steady-state solutions, transient computations are performed, ensuring that the simulation runs for a sufficient duration (advancing in pseudo-time) until the flow field parameters no longer change significantly, thereby achieving the final steady-state results.

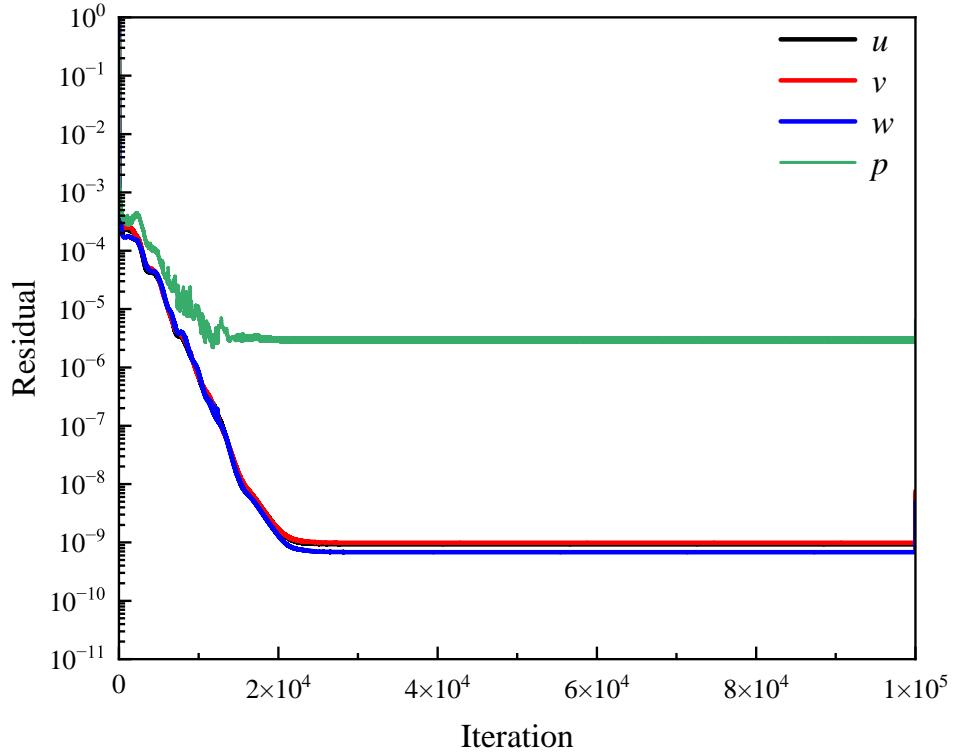


Figure 8: Residual convergence curve of velocity and pressure.

In the solution process, pressure-velocity coupling was handled using the PISO algorithm [37]. The PISO algorithm, by introducing prediction-correction steps, can efficiently and stably manage pressure-velocity coupling. To ensure numerical stability and accuracy, a virtual time step (Δt) of 1×10^{-5} seconds was used for all simulations. The CFL number was strictly controlled to be below 1. The CFL number ($CFL = \frac{|u|\Delta t}{\Delta x}$, where Δx is the characteristic mesh size) is an important parameter that measures the relationship between time step, flow velocity, and mesh size; maintaining it at a low level is key to ensuring stable convergence of the numerical solution. The total physical time for each simulation was 1 second, corresponding to 10^5 iterations. This simulation duration was sufficient for the flow field to reach a steady state. For each three-dimensional model in the Anuemo dataset, a total of 85,280 independent CFD simulations were completed under the various flow conditions described above.

Convergence Criteria and Result Validation The reliability of the simulation results was assessed by monitoring the change in residuals of various physical quantities with the number of iterations. A simulation was considered converged when the residuals of the velocity components (u, v, w) and pressure (p) decreased to preset convergence standards and remained stable. As recorded (e.g., as might be illustrated in a Figure 8, not provided here), the residuals of the velocity components rapidly decreased from an order of 10^{-3} to an order of 10^{-9} within approximately the first 20,000 iterations, eventually stabilizing at this level. This indicates that the velocity field achieved a high-precision converged solution. In contrast, the pressure residuals decreased more gradually from an order of 10^{-3} , with slight fluctuations, eventually stabilizing below 10^{-5} . Although the absolute magnitude of the pressure residuals was higher than that of the velocity residuals, this convergence level is considered acceptable accuracy for the pressure field. This residual convergence behavior validates the appropriateness of the chosen mesh resolution and ensures the accuracy and computational efficiency of the simulation results, providing a reliable basis for subsequent hemodynamic parameter analysis (including velocity components (u, v, w) and pressure p). These parameters are crucial for a deeper understanding of the blood flow environment within intracranial aneurysms and its role in aneurysm formation, development, and rupture risk assessment.

A.4 Dataset description

This dataset has been carefully curated to facilitate the development of advanced machine learning approaches for cerebral aneurysm research, with a particular focus on modeling and predicting hemodynamic behavior. It provides a comprehensive suite of data outputs that support the analysis of blood flow dynamics, which are critical to understanding aneurysm initiation, progression, and rupture risk. To enable a wide range of modeling tasks, we include multiple data representations: binary region-of-interest (ROI) images in NIfTI format (`.nii.gz`) serving as segmentation masks for vascular structures; 3D surface mesh models in STL format (`.stl`) capturing the vascular geometry; computational fluid dynamics meshes in MSH format (`.msh`) for simulation purposes; and VTK files in both `.vti` and `.vtu` formats representing pressure and velocity fields across different domains and boundary conditions. Additionally, to ensure seamless integration into machine learning pipelines, we provide NumPy-formatted (`.npy`) versions of these data fields for each flow case. All data types are preprocessed and organized to be compatible with popular machine learning frameworks such as PyTorch and TensorFlow, as well as standard CFD analysis tools. The dataset supports a broad spectrum of learning and inference tasks, including vascular segmentation, 3D geometry reconstruction, simulation of blood flow, and prediction of hemodynamic quantities such as pressure and velocity fields. To promote reproducibility and ease of use, all files follow standardized formats and are accompanied by detailed documentation. A summary of the dataset structure, file types, and their purposes is provided in Table 2.

Table 2: Dataset File Categories. *flow* represents different mass flow rates ranging from 0.001 to 0.004 kg/s, and *case_id* represents unique geometry models ranging from 1 to 10660.

Output File	Description
Mask/ <i>case_id</i> .nii.gz	Segmented vascular ROI region, similar to a segmentation mask file, used to identify vascular regions.
Stl/ <i>case_id</i> .stl	Surface mesh of the vascular geometry.
Mesh/ <i>case_id</i> .msh	Computational grid file used for fluid simulation.
VTK/ <i>flow</i> /inlet.vtp	Fluid characteristics at the inlet boundary.
VTK/ <i>flow</i> /internal.vtu	Fluid characteristics within the internal volume.
VTK/ <i>flow</i> /outlet.vtp	Fluid characteristics at the outlet boundary.
VTK/ <i>flow</i> /wall.vtp	Fluid characteristics at the wall boundary.
NPY/ <i>flow</i> /array_inlet_ <i>case_id</i> .npy	Numerical data for inlet boundary conditions.
NPY/ <i>flow</i> /array_internal_ <i>case_id</i> .npy	Numerical data for internal volume conditions.
NPY/ <i>flow</i> /array_outlet_ <i>case_id</i> .npy	Numerical data for outlet boundary conditions.
NPY/ <i>flow</i> /array_wall_ <i>case_id</i> .npy	Numerical data for wall boundary conditions.

B Machine Learning Evaluation

B.1 Model Architectures and Training Configurations

DeepONet is a deep learning architecture designed to learn operators (i.e., mappings from one function space to another). Its core components include a Branch Net and a Trunk Net. The Branch Net processes the input function; in this study, it receives encoding vectors representing the aneurysm geometry or fluid boundary conditions and extracts key features. The Trunk Net handles query

point information in the output domain, specifically 3D spatial coordinates and the Signed Distance Function (SDF) from the point to the aneurysm surface. By combining the outputs of the Branch Net and the Trunk Net, DeepONet can predict physical quantities such as pressure and velocity at the query points. In standard DeepONet implementations, both the Branch Net and Trunk Net typically employ Multilayer Perceptrons (MLPs). The main advantage of DeepONet lies in its end-to-end learning paradigm, which eliminates the need for mesh generation required by traditional CFD. It can efficiently predict full-field physical quantities and exhibits good generalization potential to new input conditions (such as boundary conditions and geometric shapes).

Our proposed DeepONet-SwinT model is an extension and improvement of the DeepONet framework. Its key innovation is the introduction of a geometric encoder (or geometric feature extraction branch) specifically designed to process 3D image inputs of aneurysms, which adopts the Swin Transformer architecture. As shown in 3, the overall architecture of DeepONet-SwinT includes: the Swin-T-based geometric encoder, a Boundary Condition Branch (BC-Branch) processing boundary conditions, a Trunk Net processing spatial coordinates and SDF, and a bypass linear scaler related to mass flow rate. Except for the geometric encoder adopting Swin-T, the MLP structure of the BC-Branch and Trunk Net, as well as the function of the bypass linear scaler, remain consistent with the configurations in the baseline DeepONet experiments. The core advantage of Swin-T lies in its shifted window mechanism (as shown in 3), which enables it to efficiently capture multi-scale spatial context features from 3D images. This is crucial for accurately encoding complex aneurysm geometric morphology, thereby significantly enhancing the model’s ability to process image inputs.

B.2 Dataset Construction and ML Evaluation Experimental Design

For model training, both DeepONet and DeepONet-SwinT are trained end-to-end using the Adam optimizer. The training objective is to minimize the total loss function between the model predictions and the high-fidelity CFD simulation reference solutions. We use the Sum of Square Error (SSE) as the basic metric, defined as follows:

$$SSE = \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (3)$$

where y_i is the ground truth value, \hat{y}_i is the predicted value, and n is the number of data points in the case. The total loss function \mathcal{L} is a weighted sum of the SSEs for the flow field (pressure and the three velocity components u, v, w), summed over all cases in the training set:

$$\mathcal{L} = \sum_{j=1}^m (\lambda_p SSE_{p,j} + \lambda_u SSE_{u,j} + \lambda_v SSE_{v,j} + \lambda_w SSE_{w,j}) \quad (4)$$

where m is the total number of cases in the training set, $SSE_{p,j}, SSE_{u,j}, SSE_{v,j}, SSE_{w,j}$ are the Sum of Square Errors for pressure and the u, v, w velocity components, respectively, for the j -th case, and $\lambda_p, \lambda_u, \lambda_v, \lambda_w$ are the corresponding weight factors. In this study, we set $\lambda_p = 1 \times 10^5$ and $\lambda_u = \lambda_v = \lambda_w = 1 \times 10^3$. During training, we used a fixed learning rate of 1×10^{-4} .

To comprehensively evaluate the performance of the DeepONet and DeepONet-SwinT models on the aneurysm hemodynamics prediction task and to validate the effectiveness of our constructed Aneumo dataset, we designed a series of systematic machine learning evaluation experiments. All evaluations are based on subsets sampled from the Aneumo dataset.

In most evaluation experiments, we adopted a geometry-based split strategy. Specifically, for each unique fundamental geometric deformation, we randomly sampled 8 CFD simulation cases to form the training set, and the remaining 2 cases were used as the validation set. This split ensures the independence of the training and validation sets in terms of geometric shape, allowing for a more reliable evaluation of the models’ generalization ability to unseen geometries. However, in experiments specifically exploring the models’ performance and scaling laws under large-scale data, we used a different split ratio: 70% of the cases from each geometric deformation were used for training, while still reserving 20% of the cases for the validation set.

We defined the metrics used to evaluate model prediction performance as follows:

The Mean Relative L2 Norm ($\overline{L2}$) is defined as:

$$\overline{L2} = \frac{1}{m} \sum_{j=1}^m \frac{\sqrt{\sum_{i=1}^n (y_{i,j} - \hat{y}_{i,j})^2}}{\sqrt{\sum_{i=1}^n y_{i,j}^2}} \quad (5)$$

The Mean Normalized Absolute Error (MNAE) is defined as:

$$MNAE = \frac{1}{m} \sum_{j=1}^m \frac{1}{n} \sum_{i=1}^n \frac{|y_{i,j} - \hat{y}_{i,j}|}{y_j^{\max} - y_j^{\min}} \quad (6)$$

where $y_{i,j}$ and $\hat{y}_{i,j}$ represent the ground truth and predicted values, respectively, for the i -th data point in the j -th case, n is the number of data points in the case, m is the total number of cases, and y_j^{\max} and y_j^{\min} are the maximum and minimum values of the ground truth data in the j -th case. Similar normalized error metrics, which are less sensitive to the absolute scale of the data, have also been employed in other machine learning studies on large-scale CFD datasets [49].

In addition to the above metrics, we also used Mean Squared Error (MSE) and Mean Absolute Error (MAE) for final evaluation. These are commonly used to measure the magnitude and dispersion of prediction errors. Mean Squared Error (MSE) is defined as the average of the squared differences between predicted and ground truth values:

$$MSE = \frac{1}{m} \sum_{j=1}^m \left(\frac{1}{n} \sum_{i=1}^n (y_{i,j} - \hat{y}_{i,j})^2 \right) \quad (7)$$

MSE gives higher weight to larger errors because it squares the error, thus it is more sensitive to outliers.

Mean Absolute Error (MAE) is defined as the average of the absolute differences between predicted and ground truth values:

$$MAE = \frac{1}{m} \sum_{j=1}^m \left(\frac{1}{n} \sum_{i=1}^n |y_{i,j} - \hat{y}_{i,j}| \right) \quad (8)$$

MAE provides a linear error measure, giving the same weight to all errors, and is less sensitive to outliers compared to MSE.

$\overline{L2}$ is influenced by data magnitude; for data with small magnitudes, a more precise model is required to achieve a smaller $\overline{L2}$ value. MNAE is not affected by data magnitude, providing a more stable relative error measure. Considering the translation invariance of fluid equations, when calculating $\overline{L2}$ and MNAE metrics involving velocity, we uniformly translated the u velocity component by a constant (e.g., -0.5) and performed the inverse translation after evaluation. This helps reveal the true sensitivity of $\overline{L2}$ and MNAE to data magnitude, enabling a fairer model evaluation.

B.3 Analysis of Model Training and Performance

B.3.1 Comparison of DeepONet and DeepONet-SwinT Convergence

We conducted a comprehensive evaluation of the performance of DeepONet and DeepONet-SwinT across the training, validation, and testing phases to assess their effectiveness in the aneurysm hemodynamics prediction task. To monitor convergence during training, we employed the $\overline{L2}$ and MNAE as evaluation metrics, computed separately for pressure (p) and velocity components (u, v, w). Figure 9 presents a comparison of the training and validation loss curves for DeepONet and DeepONet-SwinT.

As shown in Figure 9(a), during the training phase, DeepONet-SwinT consistently achieves significantly lower loss values across all metrics ($\overline{L2}$, and MNAE) compared to DeepONet. For instance, within the first 2,000 epochs, the $\overline{L2}_p$ loss of DeepONet-SwinT rapidly decreases to below 0.4, while DeepONet's $\overline{L2}_p$ remains above 0.6, indicating a faster convergence rate for DeepONet-SwinT. Similarly, the MNAE metric reveals that DeepONet-SwinT exhibits smaller prediction errors for the velocity components u, v, w , with notably higher stability in the later stages of training, as evidenced by its reduced fluctuation compared to DeepONet.

During the validation phase (Figure 9(b)), DeepONet-SwinT continues to demonstrate superior performance. Its validation losses across all metrics are consistently lower than those of DeepONet, with particularly pronounced improvements in the \bar{L}_2 and MNAE metrics, where DeepONet-SwinT exhibits a more concentrated error distribution. For example, on the validation set, DeepONet-SwinT's MNAE $_u$ stabilizes at around 0.05 after 5,000 epochs, whereas DeepONet's MNAE $_u$ remains close to 0.10, highlighting DeepONet-SwinT's stronger generalization to unseen geometries.

Overall, DeepONet-SwinT consistently outperforms DeepONet in terms of loss metrics during both training and validation, particularly in capturing complex hemodynamic features. This improvement is closely tied to the introduction of the Swin Transformer-based geometric encoder, which efficiently extracts geometric features of aneurysms, thereby enhancing the model's predictive capabilities. These findings provide strong support for DeepONet-SwinT's superior performance on the test set, as detailed in the main text (see Figures 4 and 5).

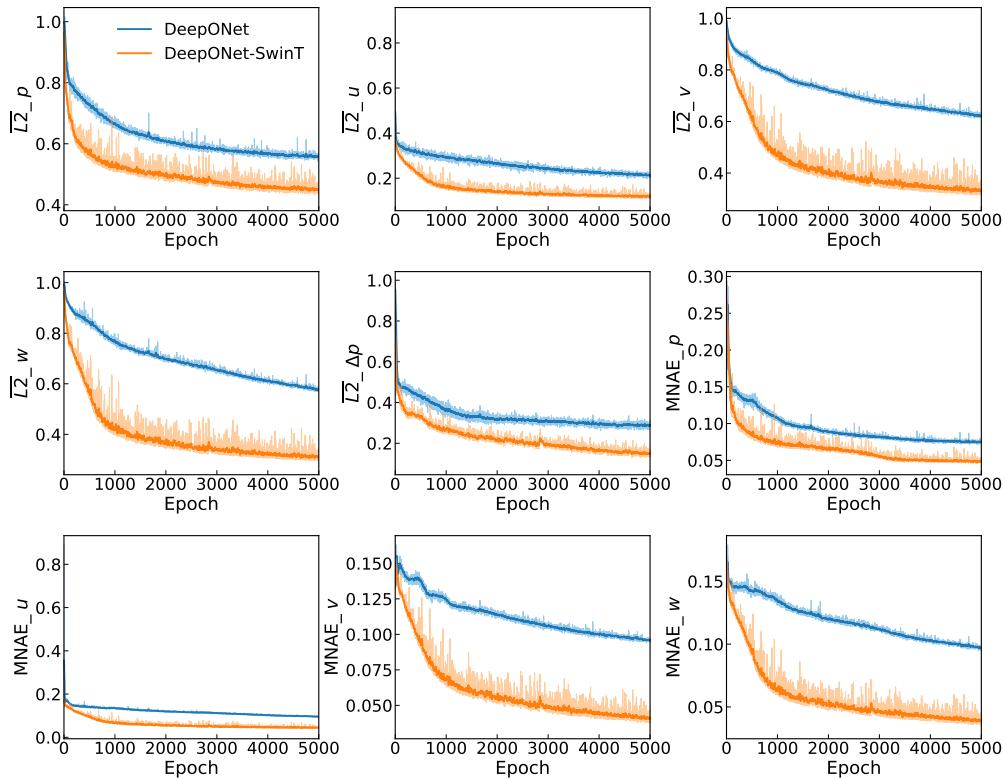
B.3.2 Impact of Training Configuration and Data Characteristics on DeepONet

Leveraging the defined dataset and evaluation metrics, we conducted detailed experiments using the baseline DeepONet model to assess key factors influencing training and performance, including batch size, training point density, and training set flow condition diversity.

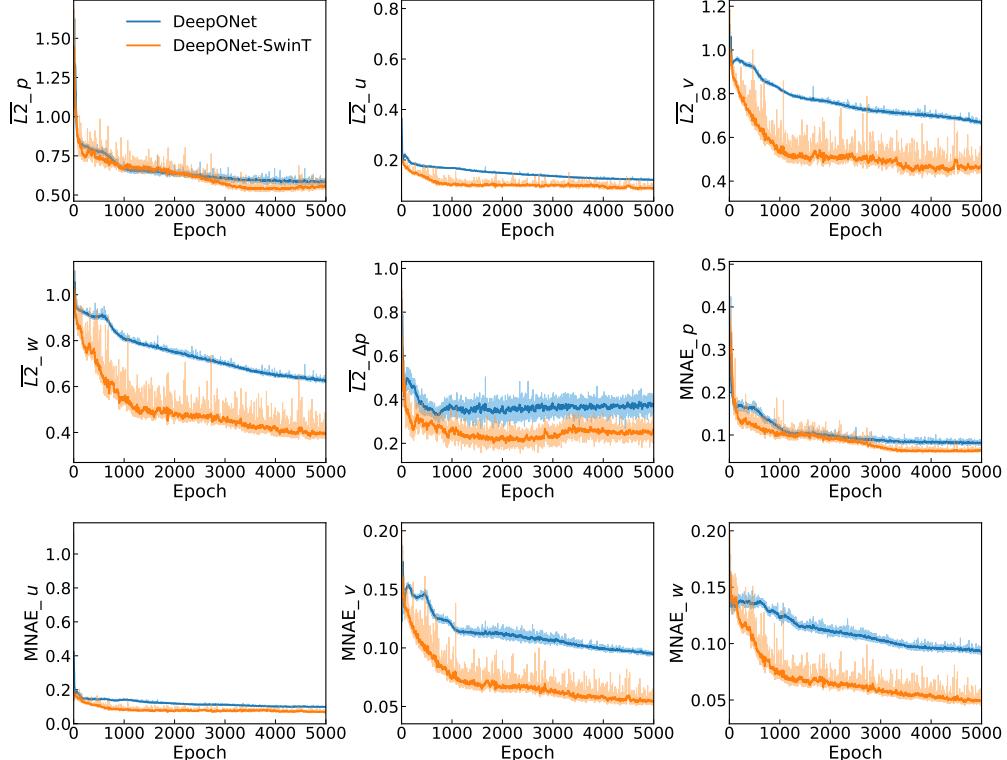
Impact of Batch Size: To investigate the impact of batch size on the training dynamics and final prediction performance of the DeepONet model, we conducted experiments with batch sizes of 4, 8, 16, 32, and 64. Figure 10 illustrates the training and validation curves for the \bar{L}_2 and MNAE metrics across these batch sizes over 5000 epochs. The results show that smaller batch sizes (BS=4 and BS=8) enable faster initial convergence and generally yield lower final \bar{L}_2 and MNAE values for most predicted quantities (pressure, v , and w velocity components) on both the training and validation sets. In contrast, larger batch sizes (BS=32 and BS=64) exhibit slower convergence and higher error values. A batch size of 16 provides a reasonable trade-off, converging faster than larger batch sizes while achieving better performance, though slightly inferior to the smallest batch sizes tested. The MNAE for the u velocity component demonstrates lower sensitivity to batch size variations compared to other metrics and components. Based on these findings, smaller batch sizes appear advantageous for training the DeepONet model on our dataset, improving convergence and overall performance. Considering both computational cost and performance, we selected a batch size of 16 for all subsequent experiments in this paper.

Impact of Training Point Density: We further investigated the sensitivity of the DeepONet model's performance to the number of spatial query points sampled per case during training. This experiment aimed to determine the minimum density of training points required to achieve satisfactory prediction accuracy. We trained the baseline DeepONet model using varying percentages of the available points from each training case: 1%, 5%, 10%, and 20%. The training and validation performance, as measured by \bar{L}_2 and MNAE over 5000 epochs, is presented in Figure 11. The results indicate that while increasing the number of training points generally leads to marginal improvements in performance, the model is remarkably efficient in terms of the required point density. The training and validation curves for 5%, 10%, and 20% of the training points are relatively close, suggesting that using a larger fraction beyond a certain threshold (e.g., 5%) provides diminishing returns. The 1% case shows slightly higher error, particularly in the initial training phase, but the gap narrows as training progresses. This trend is observed across most metrics and velocity components. The MNAE metrics, particularly for pressure and the v , w components, reveal a clearer distinction between the point densities, with higher percentages yielding slightly lower errors, but the overall performance remains comparable for 5% and above. The MNAE for the u component again shows less variance. These findings demonstrate that the DeepONet architecture can learn the operator mapping effectively even when trained on a relatively sparse set of points per case. This efficiency in terms of spatial data requirements per instance is a valuable attribute for applying SciML methods to large-scale simulations where generating dense data points can be computationally expensive.

Impact of Training Set Flow Condition Diversity: To quantitatively assess the influence of the diversity of flow conditions present in the training data on the model's ability to generalize, we designed an experiment where the number of unique flow conditions in the training set was varied while the validation set remained fixed to two unseen flow conditions. We trained the

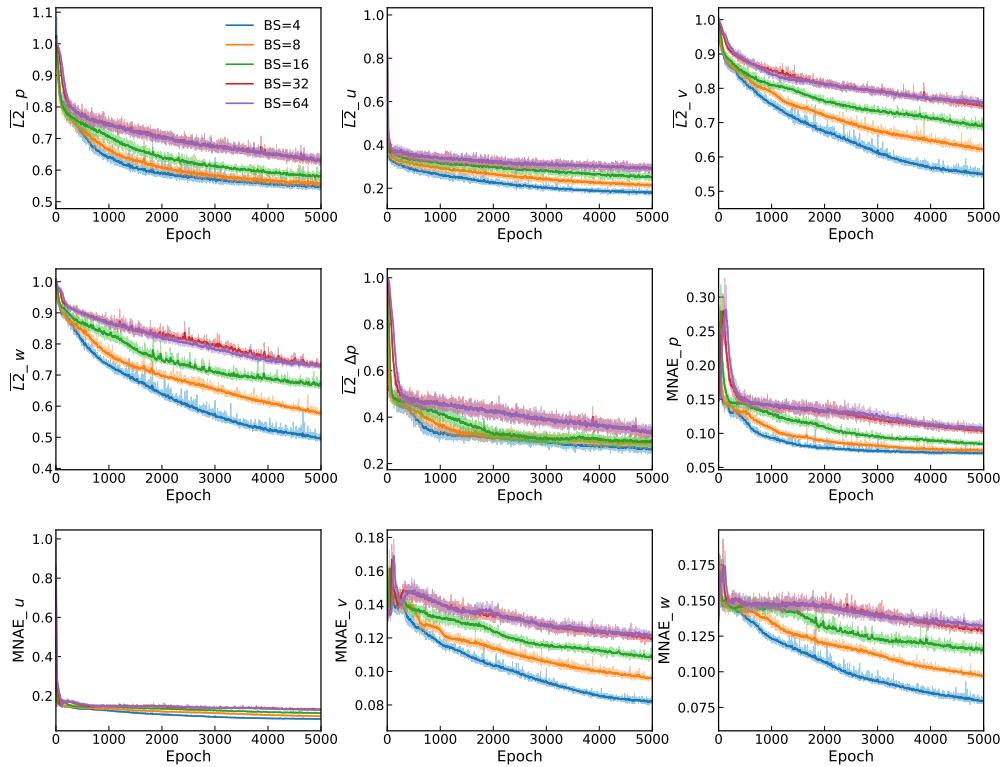


(a) Training loss

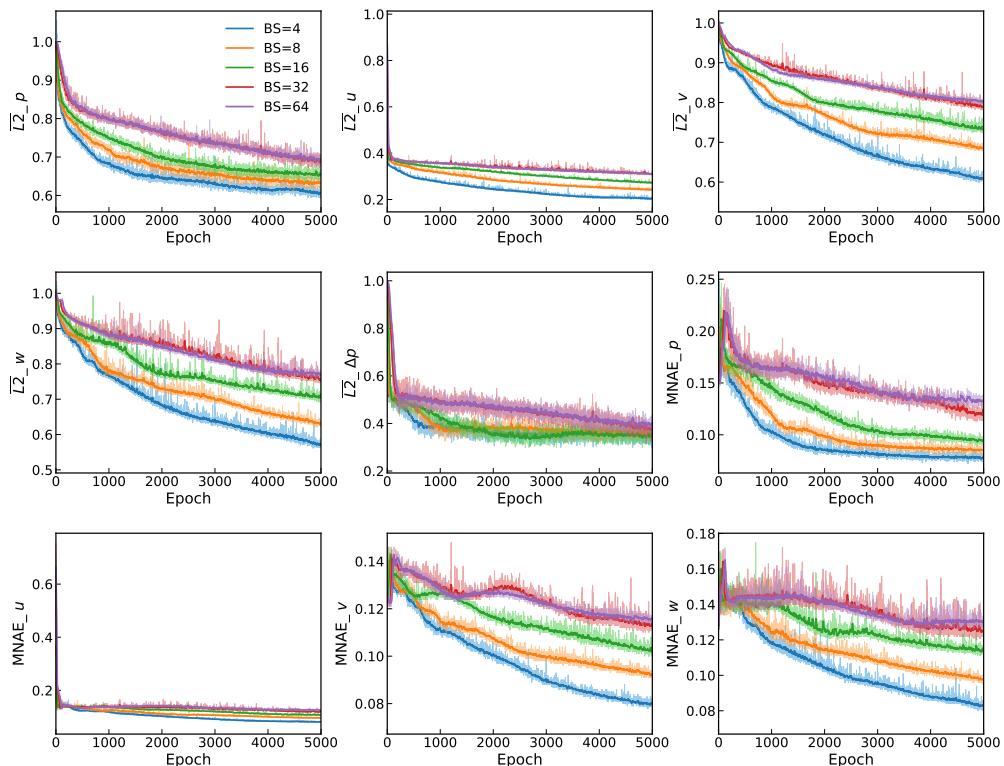


(b) Validation loss

Figure 9: Comparison of DeepONet and DeepONet-SwinT Training and Verification Convergence.

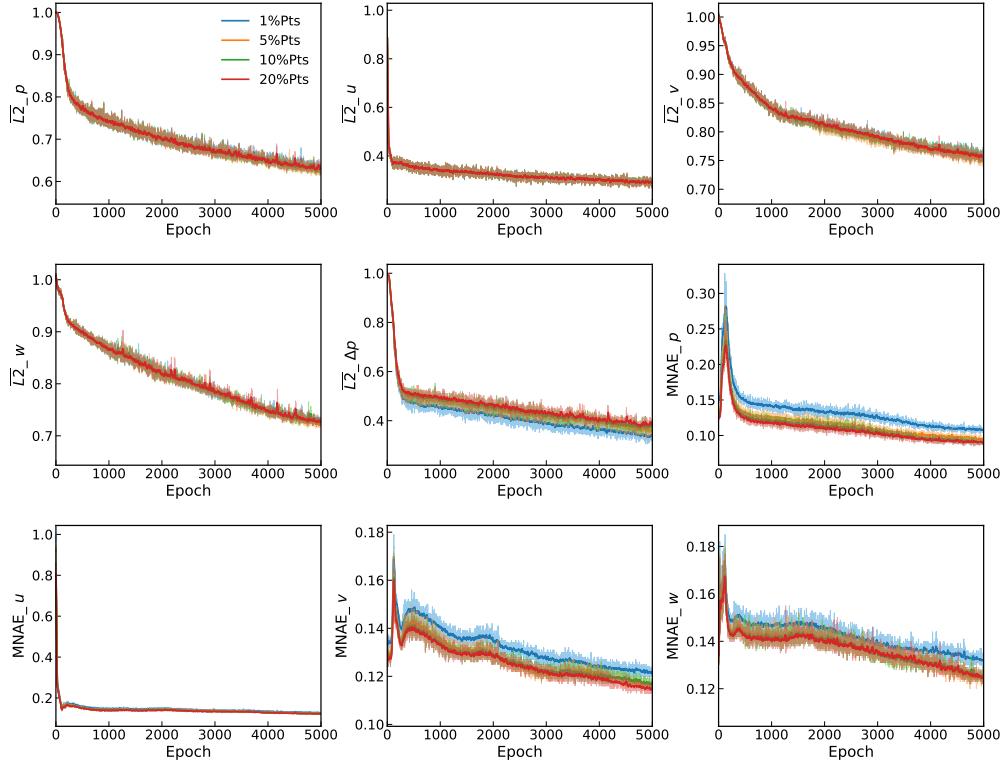


(a) Training loss under different batch sizes

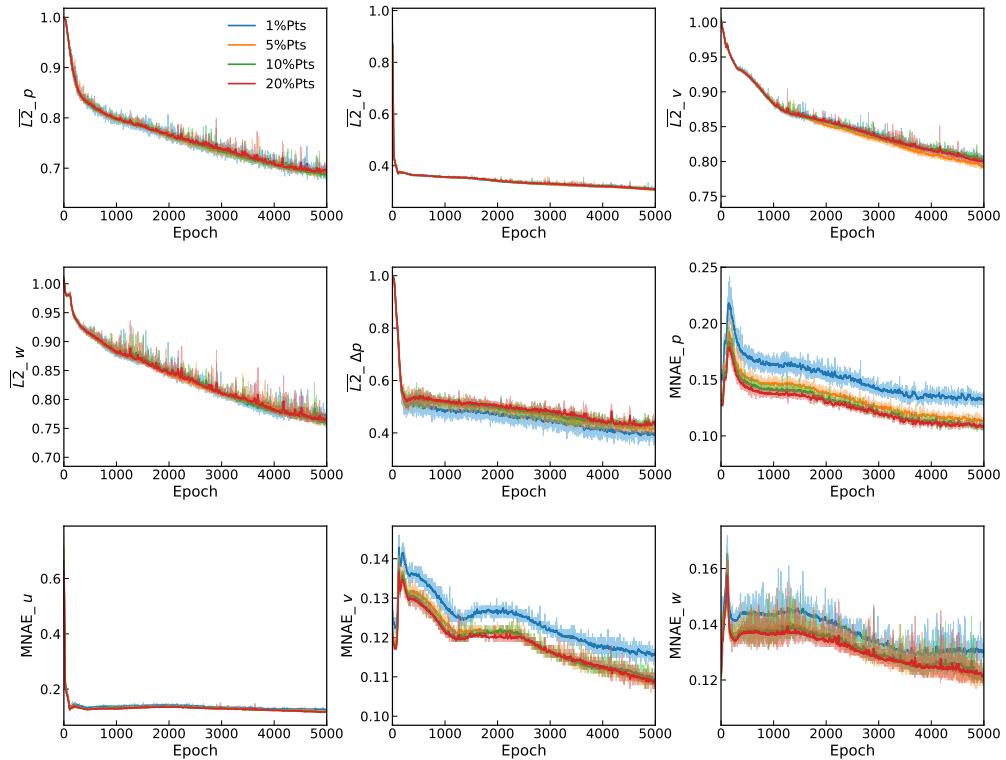


(b) Validation loss under different batch sizes

Figure 10: Effect of Batch Size on Convergence of DeepONet Training and Verification.



(a) Training loss under different training point densities



(b) Validation loss under different training point densities

Figure 11: Effect of training point density on DeepONet training and validation loss.

baseline DeepONet model using training sets containing 1, 2, 4, and 8 distinct flow conditions. The performance on both the training and fixed validation sets is presented in Figure 12, showing the evolution of \bar{L}^2 and MNAE over 5000 epochs. The results clearly indicate that increasing the number of distinct flow conditions in the training set significantly improves the model’s generalization performance on unseen flow conditions. On the validation set (Figure 12), models trained with a greater variety of flow conditions consistently achieve lower \bar{L}^2 and MNAE values for all predicted quantities. Notably, there is a substantial improvement when moving from 1 to 2 and from 2 to 4 training flow conditions. While performance continues to improve with 8 training conditions, the gains appear to diminish compared to the increase from 2 to 4. The case with only 1 training flow condition demonstrates the poorest generalization, highlighting the necessity of training data diversity for capturing the operator mapping across different boundary conditions. These findings underscore the critical role of varied flow conditions in the training data for enabling the DeepONet model to effectively predict hemodynamics under unseen operational parameters.

Impact of Validation Set Diversity on Evaluation: Complementary to the preceding experiments, we investigated the influence of the diversity of flow conditions within the validation set on the evaluated model performance. For this analysis, the training set was fixed to include 8 distinct flow conditions, while the validation sets were constructed to contain varying numbers of flow conditions: 1, 2, 4, and 8. This experiment aimed to understand how the composition of the validation set affects the assessment of the baseline DeepONet model’s generalization capability to unseen flow conditions. Figure 13 presents the results for the \bar{L}^2 and MNAE metrics. As expected, the training curves (Figure 13(a)) are nearly identical across different validation set configurations, as the training data remained constant.

The validation curves (Figure 13(b)), however, reveal a significant dependency of the evaluated performance metrics on the number of flow conditions included in the validation set. Specifically, validating the model on a set with only 1 flow condition (ValCount=1) yields substantially lower error metrics, suggesting an optimistic performance assessment. As the number of distinct flow conditions in the validation set increases (ValCount=2, 4, and 8), the measured \bar{L}^2 and MNAE values generally rise. This trend is consistent across pressure and velocity components, although the magnitude of the effect varies. The \bar{L}^2 for pressure and the MNAE for velocity components v and w show the most pronounced difference between ValCount=1 and higher counts. For most metrics, the performance evaluated with ValCount=2, 4, and 8 are closer to each other than to ValCount=1.

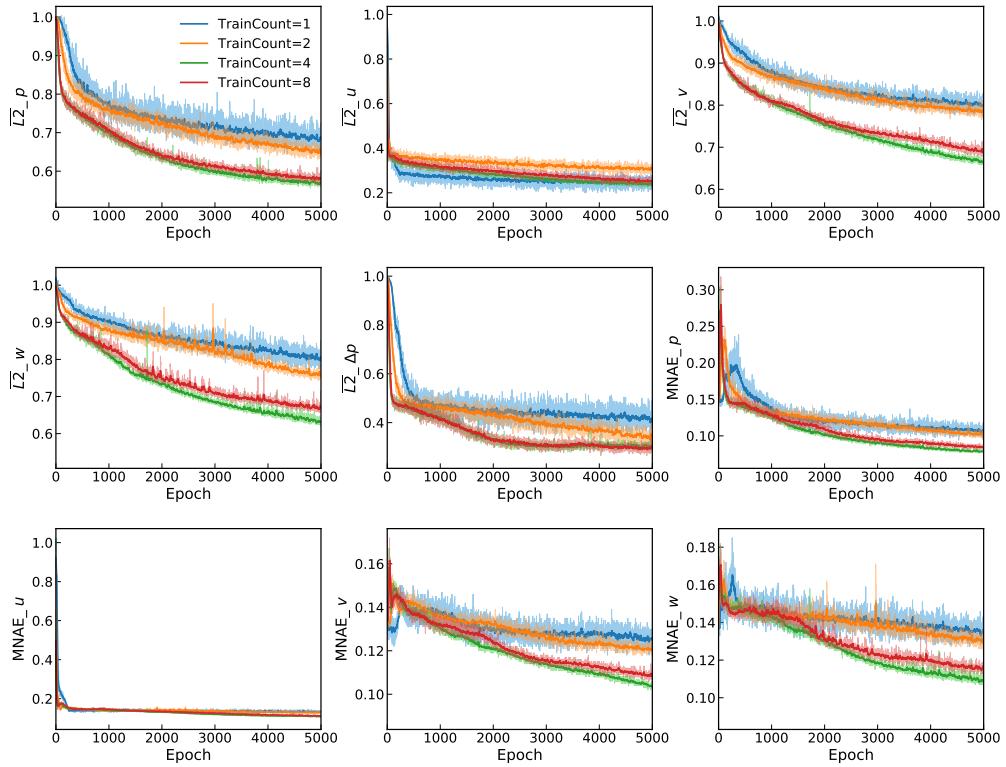
These results underscore the critical importance of using a sufficiently diverse validation set when evaluating SciML models designed to generalize across different boundary conditions. Evaluating on a limited set of unseen conditions can lead to an overestimation of the model’s true generalization performance. A validation set incorporating a greater variety of unseen flow conditions provides a more robust and realistic measure of the model’s ability to predict hemodynamics in novel scenarios. In this study, using at least 2 distinct flow conditions in the validation set appears crucial for obtaining a more reliable generalization assessment.

Scaling Performance Analysis: We ultimately conducted a scaling law study to evaluate the performance and scalability of the models under larger-scale training datasets. This analysis covered both the baseline DeepONet model and our proposed DeepONet-SwinT architecture. In these large-scale experiments, we adopted a data splitting strategy based on geometric deformation: 70% of the samples from each unique geometric deformation were allocated for training, with the remaining 2 samples reserved for validation. This methodology enabled us to systematically expand the overall scale of the training data by incorporating a broader range of geometric deformations and flow conditions. It is important to note that during these large-scale training processes, constrained by the available GPU resources (four NVIDIA A100 80GB GPUs, which allowed a maximum batch size of 8), we set the training batch size to 8. By progressively increasing the scale of the training data, we analyzed the trend of performance metrics such as \bar{L}^2 and MNAE for both models, thereby evaluating their potential for application in large-scale real-world scenarios.

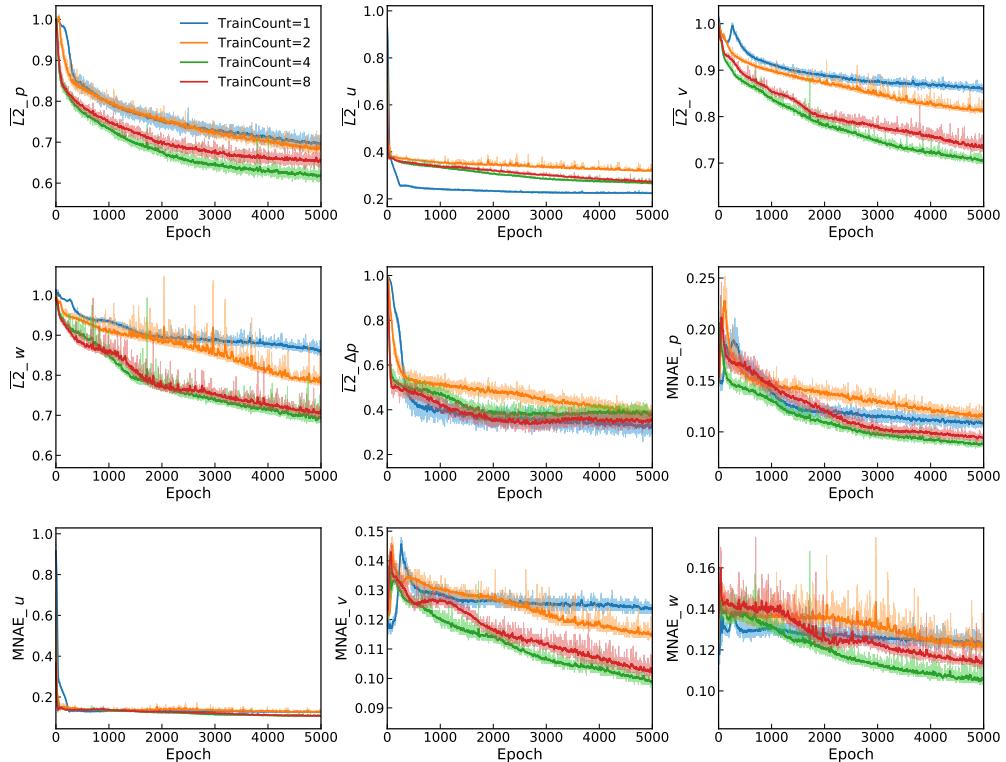
Figure 14 presents the training and validation curves for the baseline DeepONet model trained with two different data scales: TrainData=1280 and TrainData=12800 samples. Figure 15 shows the corresponding curves for the DeepONet-SwinT model under the same conditions. The results clearly indicate that increasing the scale of the training data significantly enhances the performance of both

architectures. For both the training and validation sets, the $\overline{L2}$ and MNAE metrics are substantially lower at TrainData=12800 compared to TrainData=1280.

When comparing the two models at the larger data scale (TrainData=12800), DeepONet-SwinT consistently exhibits lower validation errors across most metrics (see Figure 15(b) in contrast to Figure 14(b)). Notably, DeepONet-SwinT demonstrates a more pronounced benefit from the expansion of data volume, underscoring the advantage of its Swin Transformer-based geometric encoder in effectively utilizing the richer geometric information present in a larger and more diverse dataset. This further validates the effectiveness of our constructed dataset and the proposed architecture. The observed scaling trend provides valuable insights into the potential of applying SciML methods to large-scale computational fluid dynamics problems.

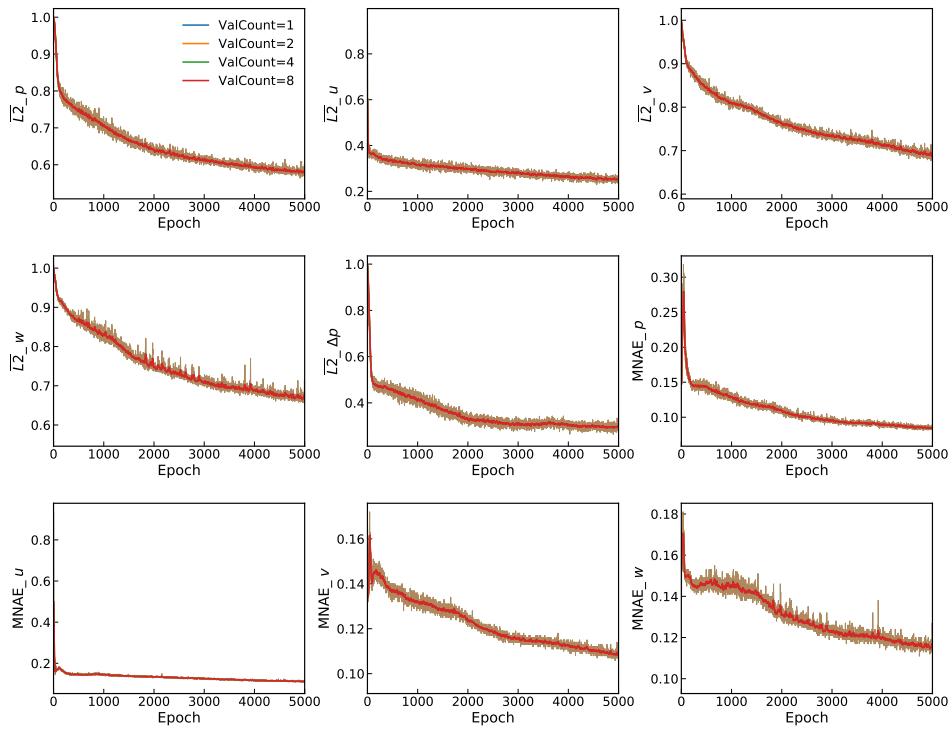


(a) Training loss under different numbers of distinct training flow conditions

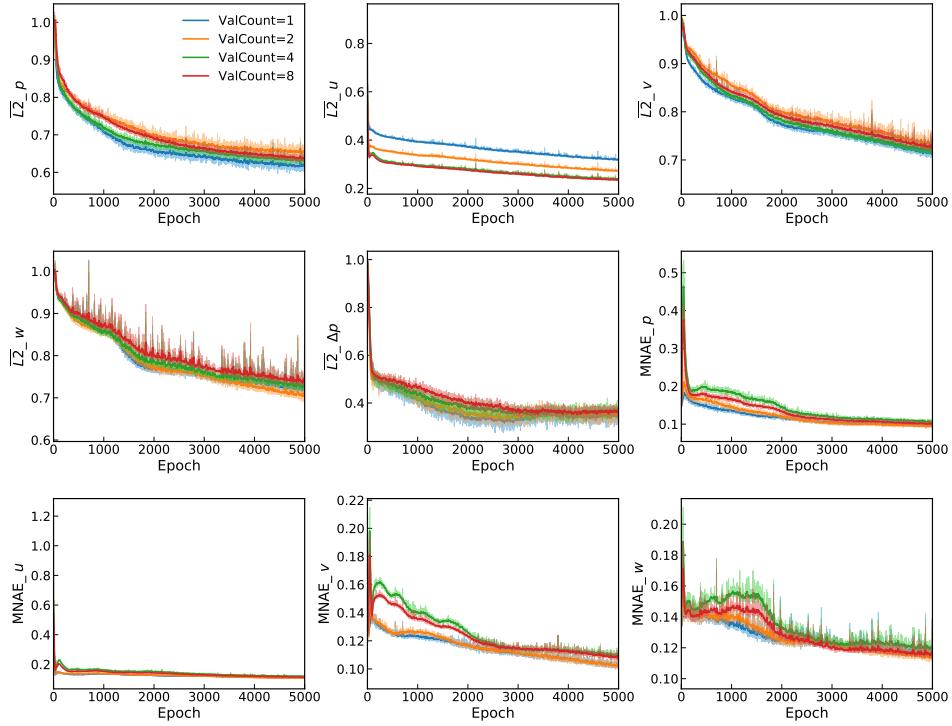


(b) Validation loss under different numbers of distinct training flow conditions

Figure 12: Effect of training set flow condition diversity on DeepONet training and validation loss.



(a) Training loss with different number of validation set flow field conditions (training set fixed)



(b) Validation loss with different number of validation set flow field conditions (training set fixed).

Figure 13: Effect of validation set flow condition diversity on DeepONet training and validation loss.

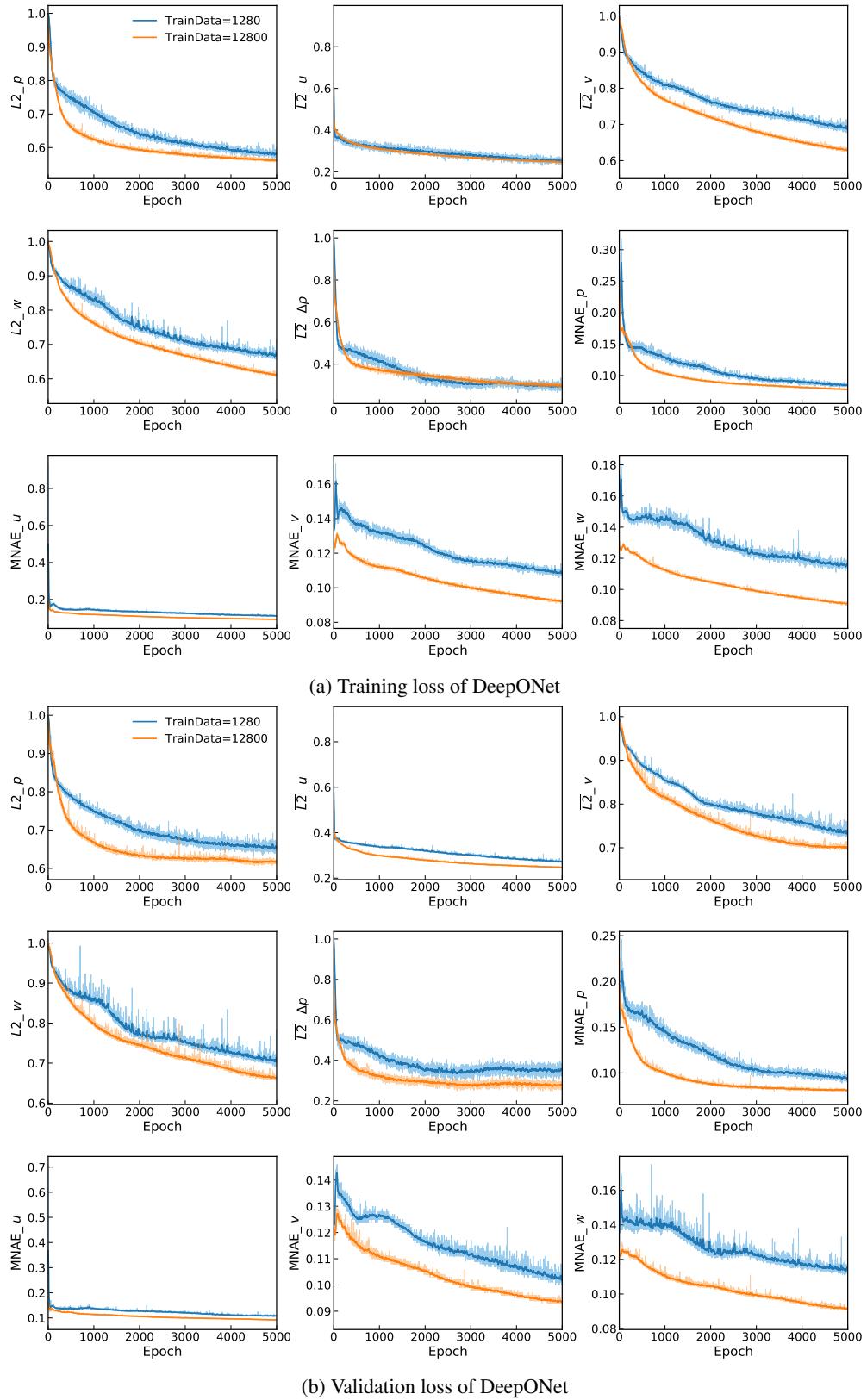
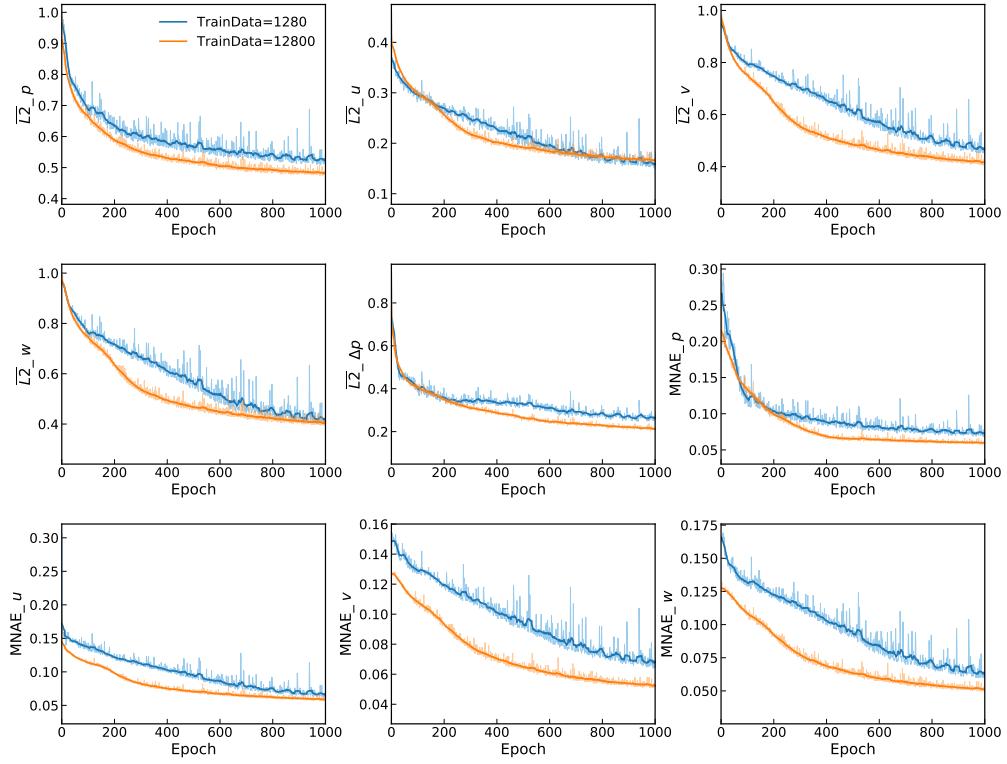
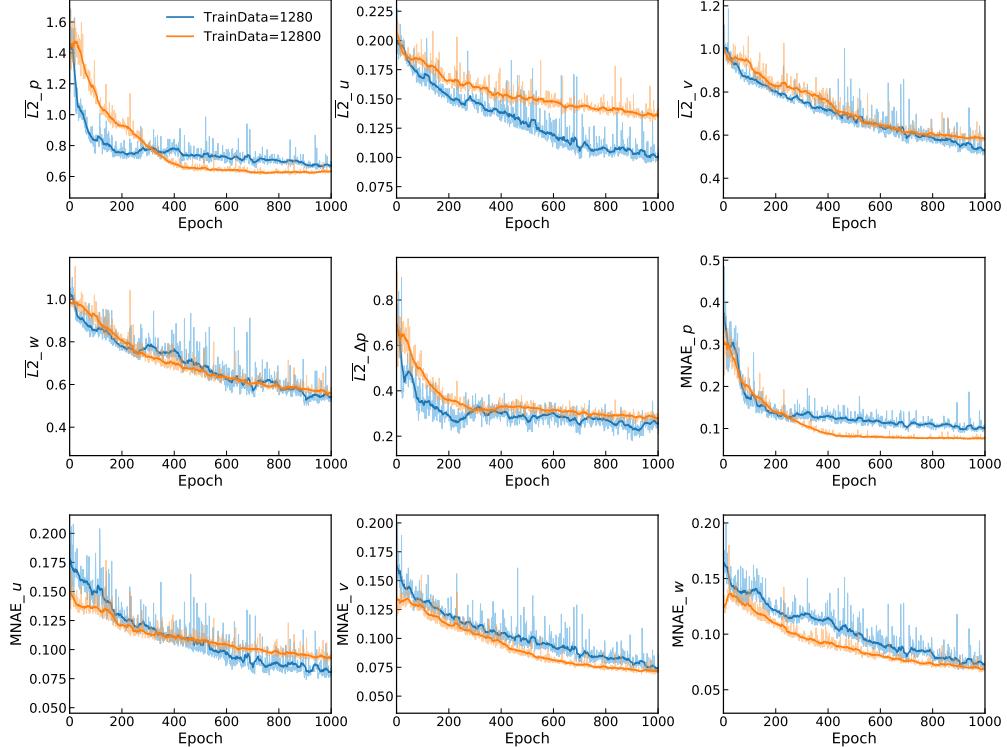


Figure 14: Scaling performance of DeepONet with different training data scales.



(a) Training loss of DeepONet-SwinT



(b) Validation loss of DeepONet-SwinT

Figure 15: Scaling performance of DeepONet-SwinT with different training data scales