vesselFM: A Foundation Model for Universal 3D Blood Vessel Segmentation

Bastian Wittmann
University of Zurich

Yannick Wattenberg
ETH Zurich

Tamaz Amiranashvili
Technical University of Munich

Suprosanna Shit
University of Zurich

Bjoern Menze
University of Zurich

Motivation:

3D blood vessel segmentation is a critical task in (bio)medical image analysis, with high relevance to numerous downstream applications.
However, substantial domain gaps - both across and within imaging modalities - as well as the continuous emergence of novel imaging techniques, often require researchers and clinicians to annotate datasets from scratch to segment and analyze their data at hand. This presents a significant bottleneck to efficient and scalable vascular analysis.
To overcome this, we propose vesselFM, a foundation model precisely tailored to universal 3D blood vessel segmentation that is resilient to domain shifts. Trained on three heterogeneous, large-scale data sources, vesselFM achieves state-of-the-art performance in *zero*-, *one*-, and *few*-shot segmentation across diverse vascular images.

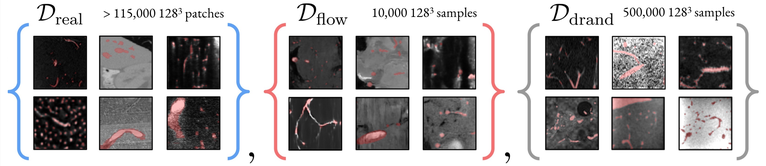
Contributions:

- We propose a foundation model for universal 3D blood vessel segmentation and demonstrate *zero*-shot generalization to unseen domains.
- We curate the largest dataset for 3D blood vessel segmentation to date, consisting of three heterogeneous, carefully curated data sources.
- We propose a domain randomization strategy tailored to 3D blood vessel segmentation.
- We introduce mask- and class-conditioned flow matching to 3D medical image synthesis.
- We release code and model weights.

Code (GitHub)

Connect (LinkedIn)

Data Curation: We generate 3 heterogeneous, large-scale data sources



real, carefully pre-processed data containing several imaging modalities and anatomical regions from 17 annotated datasets

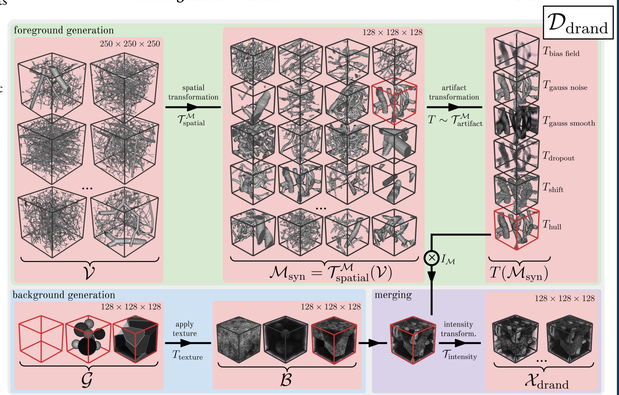
synthetic data sampled from a mask- and class-conditioned **flow matching-based generative model**

synthetic data of semi-randomized style generated by our proposed **domain randomization pipeline**

***Foreground generation*:** Synthesis of vascular patches of varied characteristics with subsequent emulation of realistic foreground artifacts.

***Background generation*:** Synthesis of highly diverse background intensity patterns and shapes.

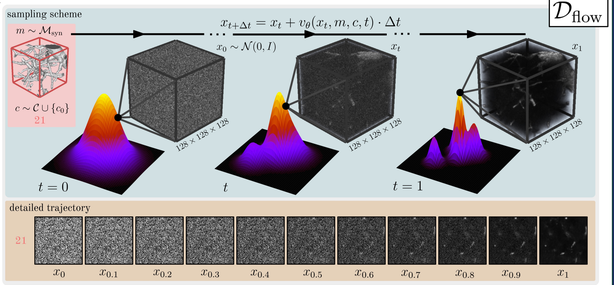
***Merging*:** Combining fore- and background images followed by applying an ample, carefully selected range of intensity transformations to generate highly diverse image-mask pairs of semi-randomized style.



Schematic overview of our domain randomization pipeline for data synthesis.

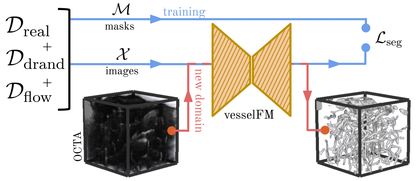
***Flow matching*:** We train a mask- and class-conditioned flow matching-based generative model using the data from our other two data sources.
Mask conditioning is implemented by concatenating the mask channel-wise with the input image.

***Sampling*:** Subsequently, we sample additional synthetic images to broaden the distributions of our data sources in a data-driven manner.

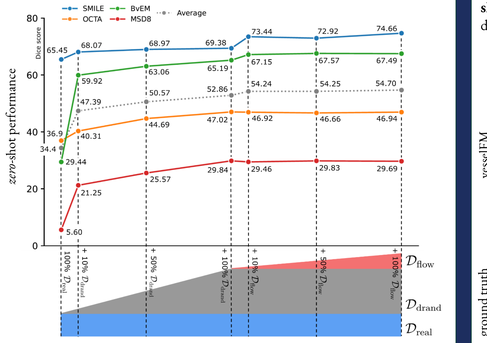


Sampling from our flow matching-based generative model for data synthesis.

Training: We train on all 3 data sources

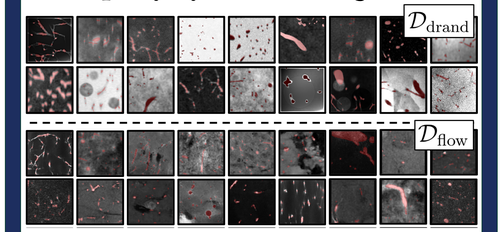


VesselFM is simply **trained in a supervised manner** on image-mask pairs from our three data sources (blue) and subsequently applied to new, unseen domains.



Experiment highlighting the necessity of training vesselFM on all three of our proposed data sources. In general, the ***zero*-shot performance benefits from the additional diversity introduced by the synthetic datasets.**

Exemplary Synthetic Images:



Quantitative & Qualitative Results:

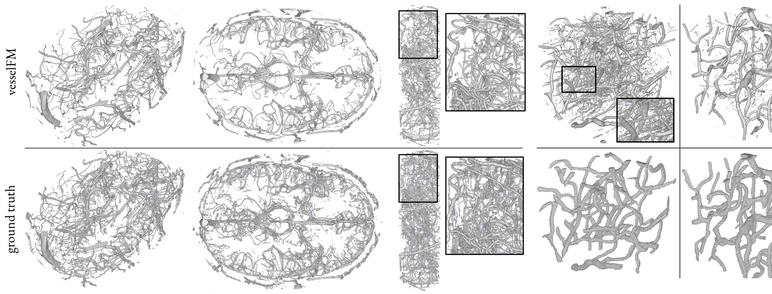
VesselFM is evaluated on **four (pre-)clinically relevant datasets** of unseen domains and is compared to state-of-the-art generalist 3D segmentation (foundation) models.

We define three tasks: ***Zero*-, *one*-, and *few*-shot segmentation**. *One*- and *few*-shotting is defined as fine-tuning models on either one or three patches of shape 128^3.

VesselFM exhibits **exceptional *zero*-shot generalization** on all four datasets, outperforming all baselines.

Task	Model	OCTA Dice ↑	OCTA clDice ↑	BvEM Dice ↑	BvEM clDice ↑	SMILE-UHURA Dice ↑	SMILE-UHURA clDice ↑	MSD8 Dice ↑	MSD8 clDice ↑
zero-shot	tUbeNet	36.01	23.64	10.03	11.17	48.32	36.85	5.13	5.84
	VISTA3D	13.60	3.72	0.94	2.03	5.05	1.62	23.83	20.25
	SAM-Med3D	6.74	6.56	5.98	7.38	2.12	1.66	7.94	7.89
	MedSAM-2	28.56	15.76	10.92	12.27	3.85	5.46	14.53	14.27
	vesselFM (ours)	**46.94**	**67.07**	**67.49**	**62.04**	**74.66**	**75.27**	**29.69**	**36.14**
one-shot	tUbeNet	38.09	59.37	10.75	11.53	57.67	53.25	13.66	15.41
	VISTA3D	51.24	25.69	8.25	15.04	56.53	42.42	31.73	32.94
	SAM-Med3D	38.33	54.90	49.47	52.14	38.57	36.94	29.29	36.78
	MedSAM-2	56.68	50.95	24.07	24.69	19.78	11.87	30.21	23.89
	vesselFM (from scratch)*	65.57	73.79	63.85	39.55	37.99	45.72	27.13	29.48
	vesselFM (ours)	**72.10**	**83.73**	**78.27**	**79.91**	**76.43**	**78.36**	**36.88**	**48.65**
few-shot	tUbeNet	41.61	57.98	5.41	10.22	56.31	49.28	17.67	18.97
	VISTA3D	54.25	32.59	24.04	38.10	61.17	51.05	41.90	46.45
	SAM-Med3D	37.85	56.94	57.86	66.04	46.59	44.63	31.30	35.48
	MedSAM-2	56.96	51.99	18.76	19.66	58.15	42.72	29.24	22.38
	vesselFM (from scratch)*	67.37	75.79	63.03	56.69	50.51	58.77	32.64	36.03
	vesselFM (ours)	**75.70**	**84.03**	**78.11**	**84.54**	**78.77**	**79.37**	**45.04**	**57.25**

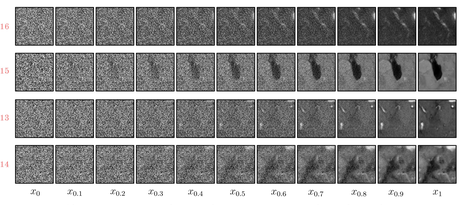
*Model not pre-trained on $\mathcal{D}_{\text{real}}$, $\mathcal{D}_{\text{drand}}$, and $\mathcal{D}_{\text{flow}}$.



Qualitative *zero*-shot results on SMILE-UHURA.

Qualitative *zero*-shot results on BvEM.

Exemplary Flow Matching Traj.:



Class IDs corresponding to classes in $\mathcal{D}_{\text{real}}$ are indicated in red.

Summary & Conclusion:

- VesselFM is the **first foundation model for universal 3D blood vessel segmentation**, demonstrating strong generalization to unseen domains.
- **Trained on three diverse data sources**, vesselFM achieves state-of-the-art *zero*-shot performance, setting a new benchmark for vascular segmentation.
- By **addressing a key (pre-)clinical challenge**, vesselFM paves the way for novel insights into vascular disorders and the development of next-generation diagnostic tools.
- Future work should extend vesselFM to **multi-class or instance segmentation tasks**.

Keywords: 3D Segmentation · Foundation Models · Domain Generalization · Synthetic Data · Deep Generative Models · Medical Image Analysis · Blood Vessels