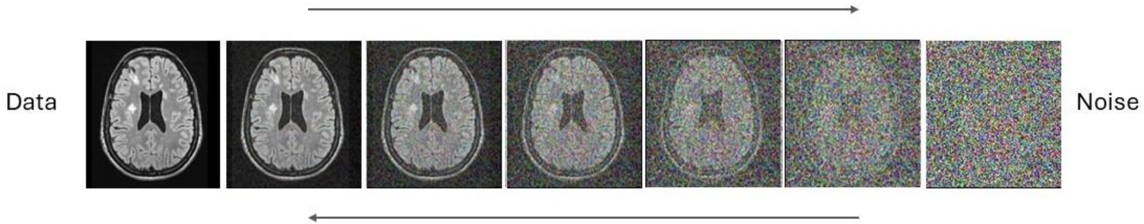
Diffusion-Based Filling and Synthesis of Multiple Sclerosis Lesions

Master Thesis Presentation

Experts: PD Dr. Richard McKinley, Dr. Ivan Diaz

Vinzenz Uhr

12.09.2024

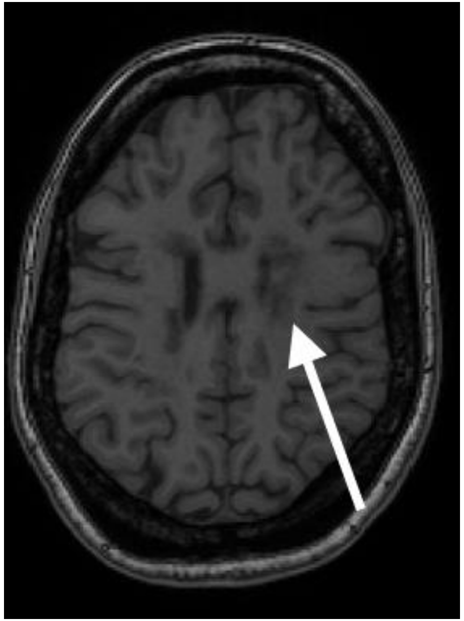


Topics

- Introduction
- Methods
    - Models
    - Dataset
- Results
- Discussion
- Conclusions

Multiple Sclerosis and Lesion Filling

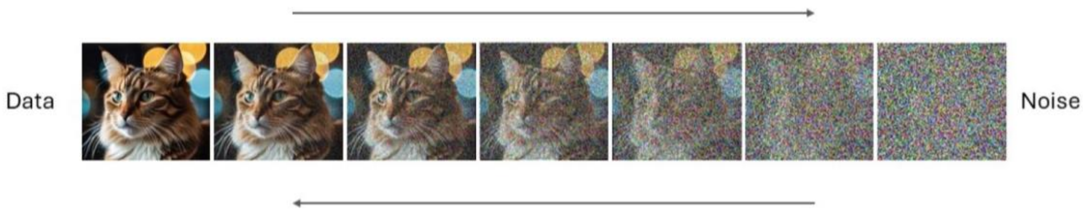
- Multiple sclerosis (MS) is a neurological disease
- Damages the coating of the nerves (myelin), which becomes visible as lesions on an MRI scan
- Computational methods on MRI scans (e.g. cortical thickness measurements) can be affected by the presence of white matter (WM) lesions [1]
- Countermeasure is lesion filling



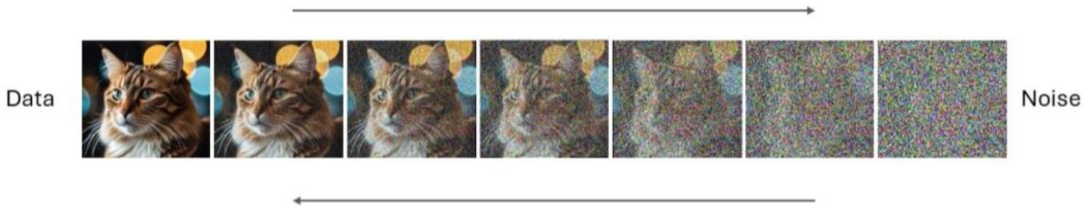
T1w brain MRI with lesions

Noise Diffusion Models [2]

- SOTA for image synthesis
- Technology of famous text-to-image models like Dall-E or Stable Diffusion.
- Transforming a sample x_0 from original distribution to a sample x_T from a normal distribution by adding Gaussian noise over 1000 steps.
- The model learns to revert this process.



Noise Diffusion Models



- Model predicts the integrated noise for each timestep t and calculates the next timestep
- To generate a new image, we start with pure noise $x_T = N(0,1)$ and go through every timestep to obtain final prediction x_0.

Noise Diffusion Models

- Noise diffusion models are data hungry
- Data scarcity for annotated MS Lesions
- Synthetic data generation is promising alternative

Goals

- Develop a method to fill WM lesions in MR-images using noise diffusion models.
- Create a noise diffusion model capable of generating new, synthetic WM lesions within MR-images.
- Assess the impact of lesion filling on cortical thickness measurements using current tools to evaluate their robustness.

Methods

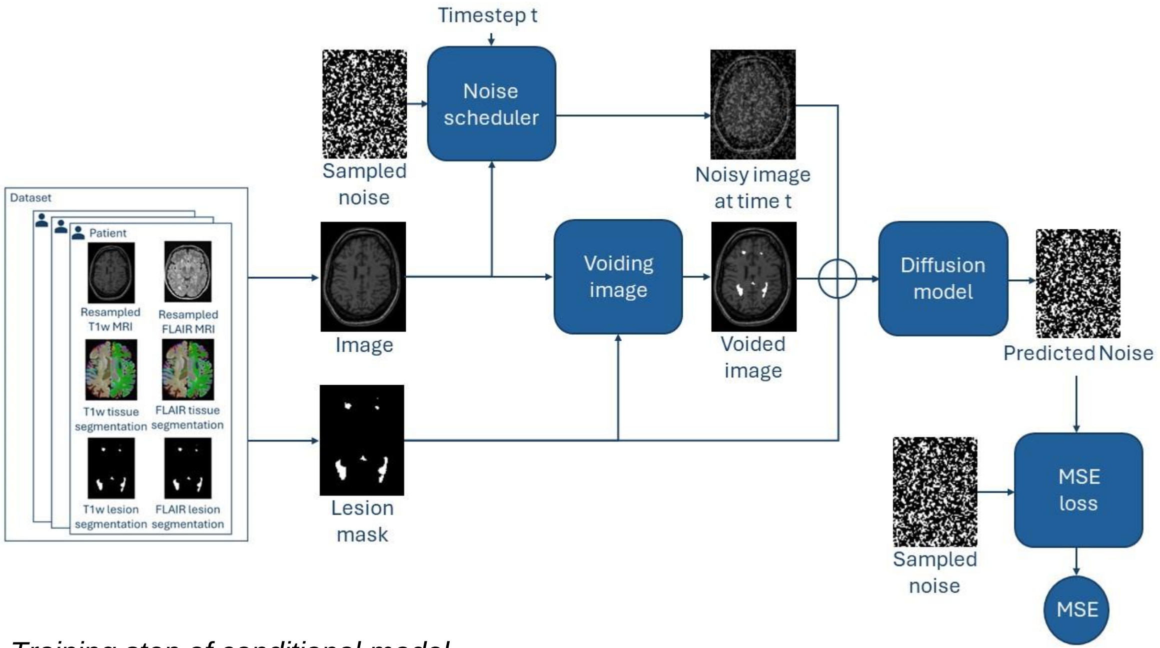
Models

Lesion Filling Conditional

- Model conditions on a binary mask and a voided MR-image
- Incorporated through concatenation at each timestep t

Models

Lesion Filling Conditional

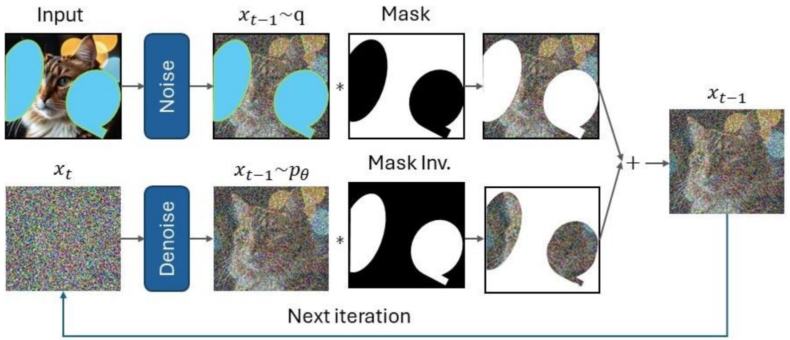


Training step of conditional model

Models

Lesion Filling Unconditional

- Train a noise diffusion model without conditioning
- Use RePaint [3] approach to condition the generation process by replacing pixels outside the mask with the MR-scan



RePaint sampling approach

Models

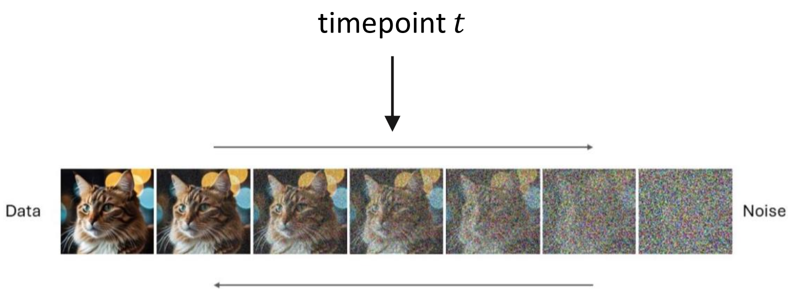
Lesion Synthesis Conditional

- Like conditional lesion filling
- Key distinction: Training target is inpainting WM lesions

Models

Lesion Synthesis Unconditional

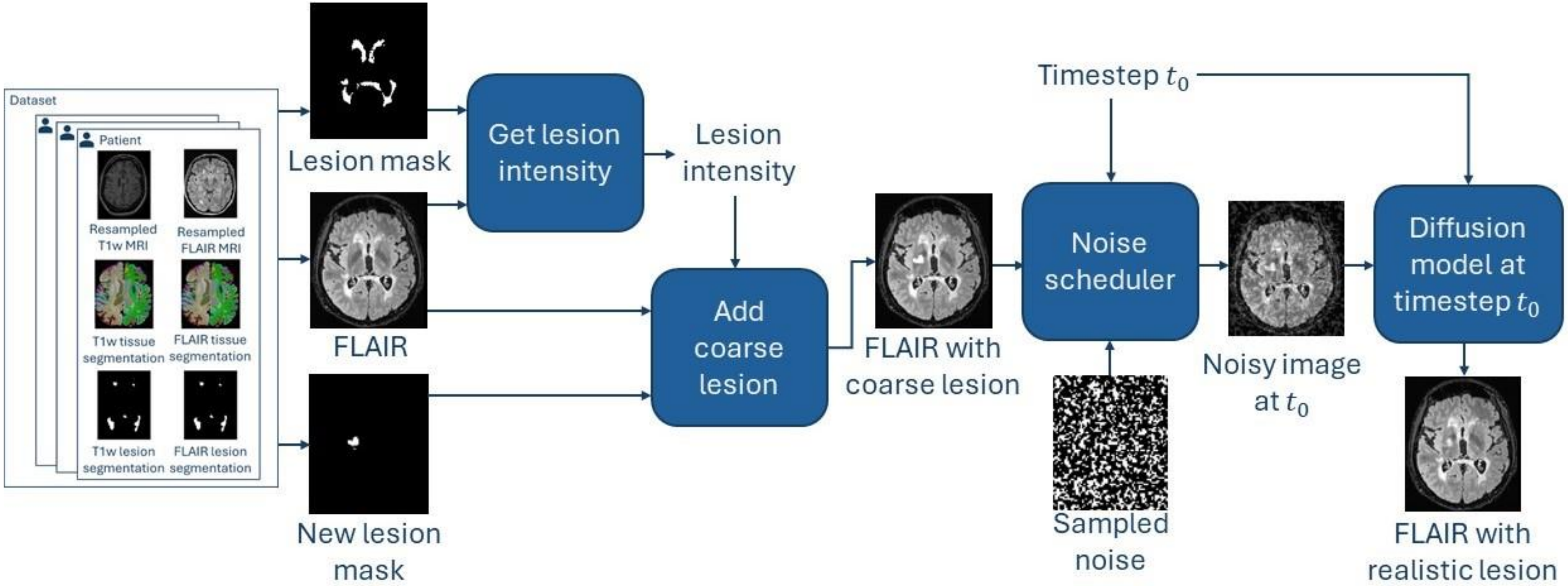
- Train a noise diffusion model without conditioning
- Add coarse lesion to MR-image, add noise and start reverse diffusion process from timepoint t.



Diffusion process

Models

Lesion Synthesis Unconditional



Unconditional lesion synthesis pipeline

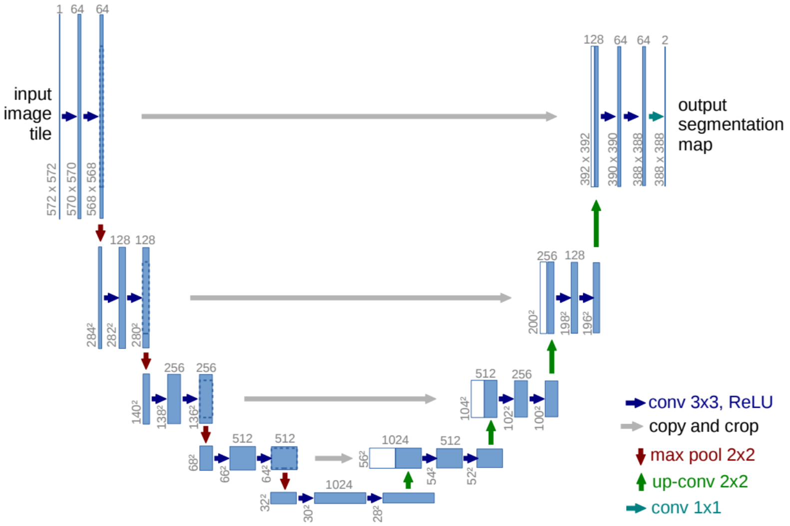
Datasets

- Different datasets with healthy and MS patients with lesions
- T1w and FLAIR images
- Binary lesion masks
- Synthetic random circle masks for lesion filling

U-Net Architectures

- 2D U-Net
- Pseudo 3D U-Net

Pseudo 3D U-Net layer



U-Net architecture [4]

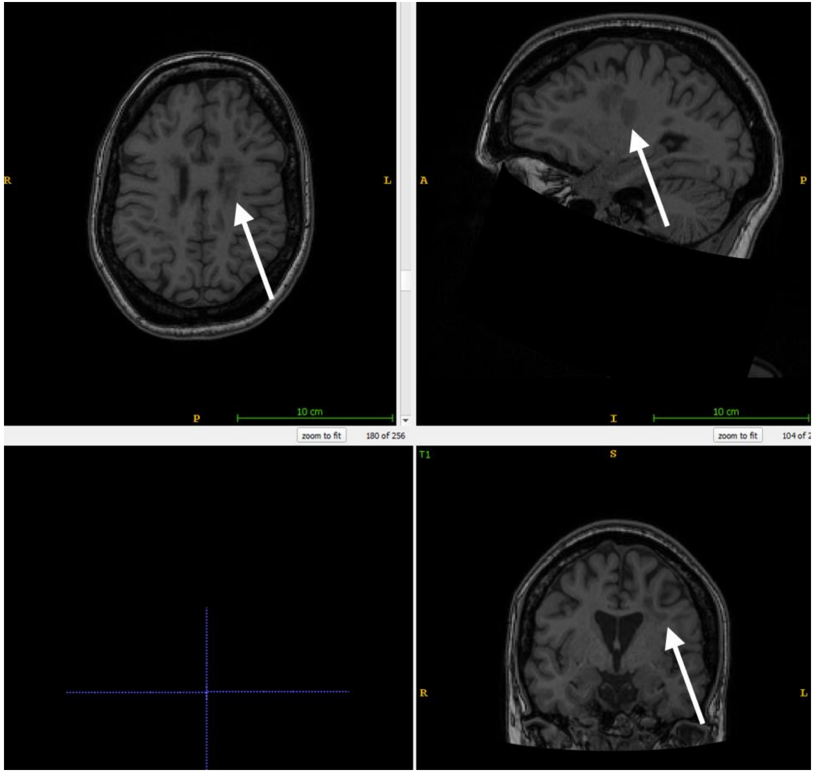
Results

Lesion Filling: Evaluation

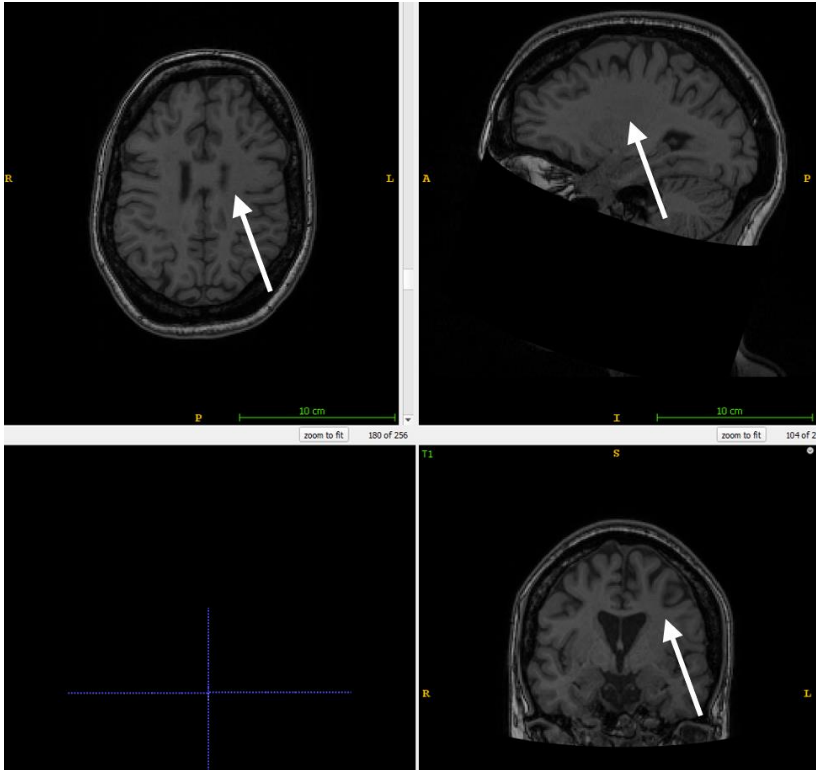
- Winner: Conditional model with pseudo 3D-Unet and 50% synthetic random circle masks
- Pseudo 3D: +5-8%
- Synthetic masks: +3-5%

	SSIM	PSNR	MSE	LPIPS
2D unconditional RePaint	0.83	28	8.2e-3	2.0e-3
2D conditional circles	0.9	32	4e-3	2e-3
2D conditional lesions	0.85	28	0.01	5e-3
2D conditional mixture	0.9	33	4e-3	1e-3
3D unconditional RePaint	0.90	32	3e-3	9e-4
3D conditional circles	0.95	38	1e-3	3e-4
3D conditional lesions	0.93	34	3e-3	4e-4
3D conditional mixture	**0.96**	**39**	**8e-4**	**2e-4**

Lesion Filling: Examples



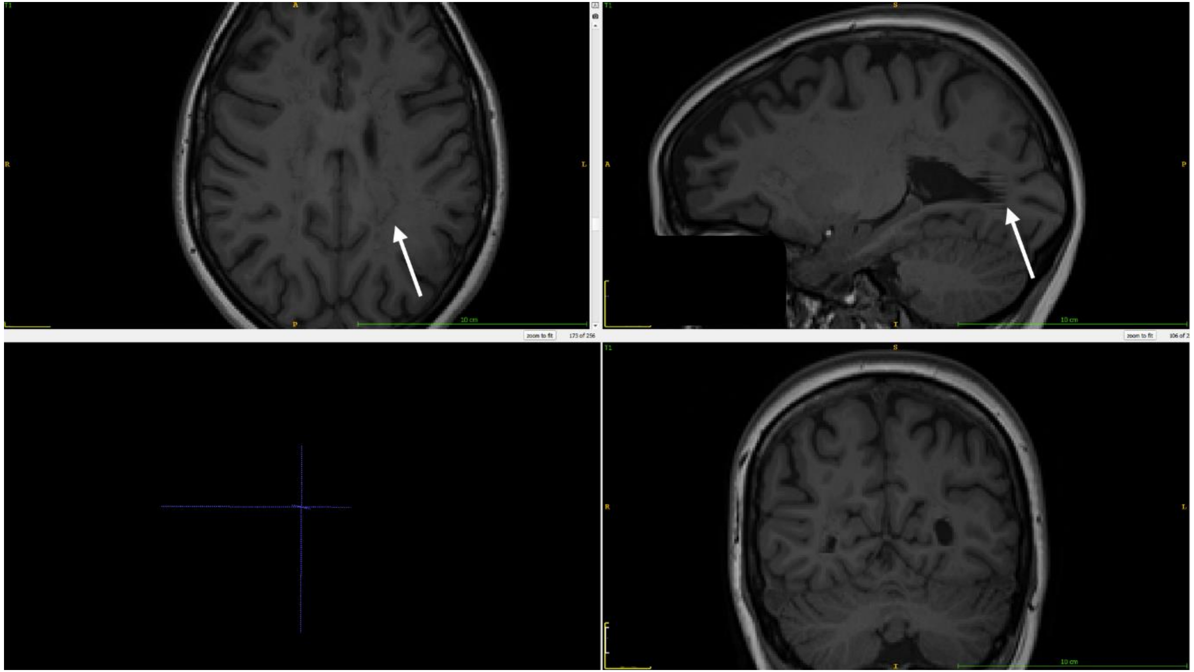
T1w before lesion filling



T1w after lesion filling

Lesion Filling: Solutions against Artifacts

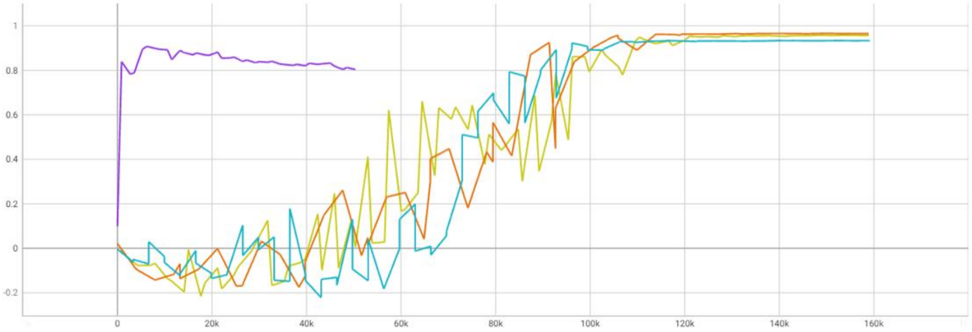
- Dilation to prevent border artifacts
- Pseudo 3D layer to prevent stripe artifacts



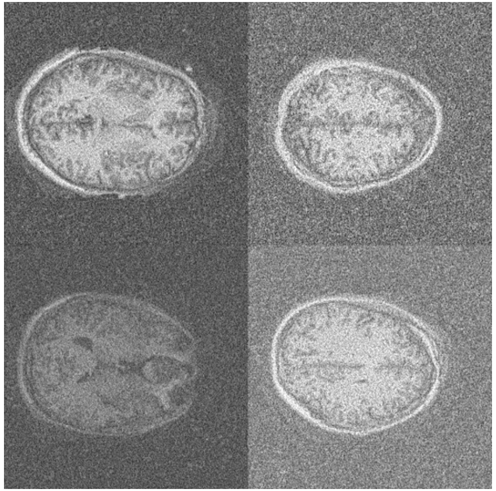
Border (left arrow) and stripe artifacts (right arrow)

Lesion Filling: Training Duration

- RePaint achieves peak SSIM of 0.9 at 6000 training steps
- Conditional models require 90’000 steps for comparable performance
- Unconditional model is still underdeveloped



SSIM score of the 4 3D models unconditional RePaint (violet), conditional mixture (red), conditional circles (yellow) and conditional lesions (blue).



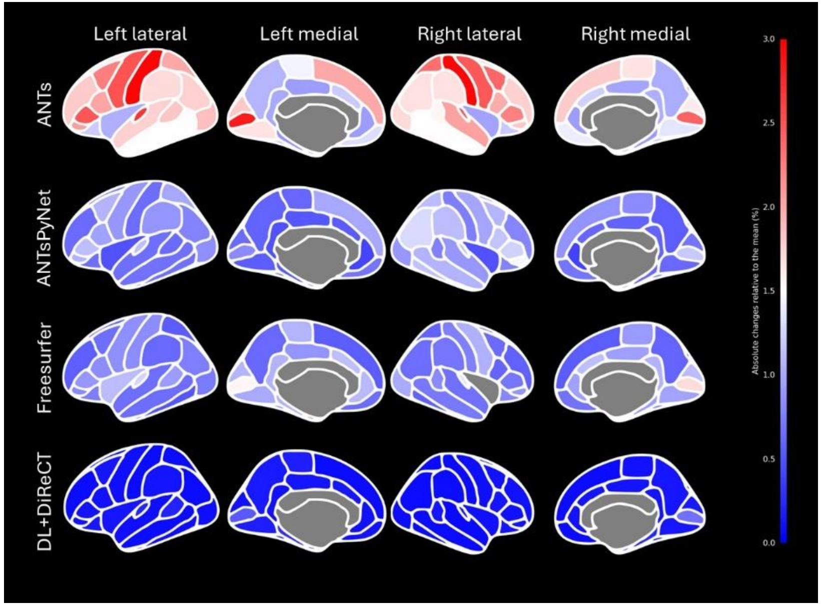
Samples of unconditional model without RePaint

Lesion Filling: Robustness measurements

- Comparing cortical thickness measurements before and after lesion filling

	Global mean thickness (%)	ROI-average (%)
ANTs	1.31	1.68
ANTsPyNet	0.52	0.84
Freesurfer	0.51	0.92
DL+DiReCT	**0.05**	**0.14**

Mean reproducibility errors



Color-coded reproducibility errors of the ROI-wise average cortical thicknesses.

SCAN Support Center for Advanced Neuroimaging

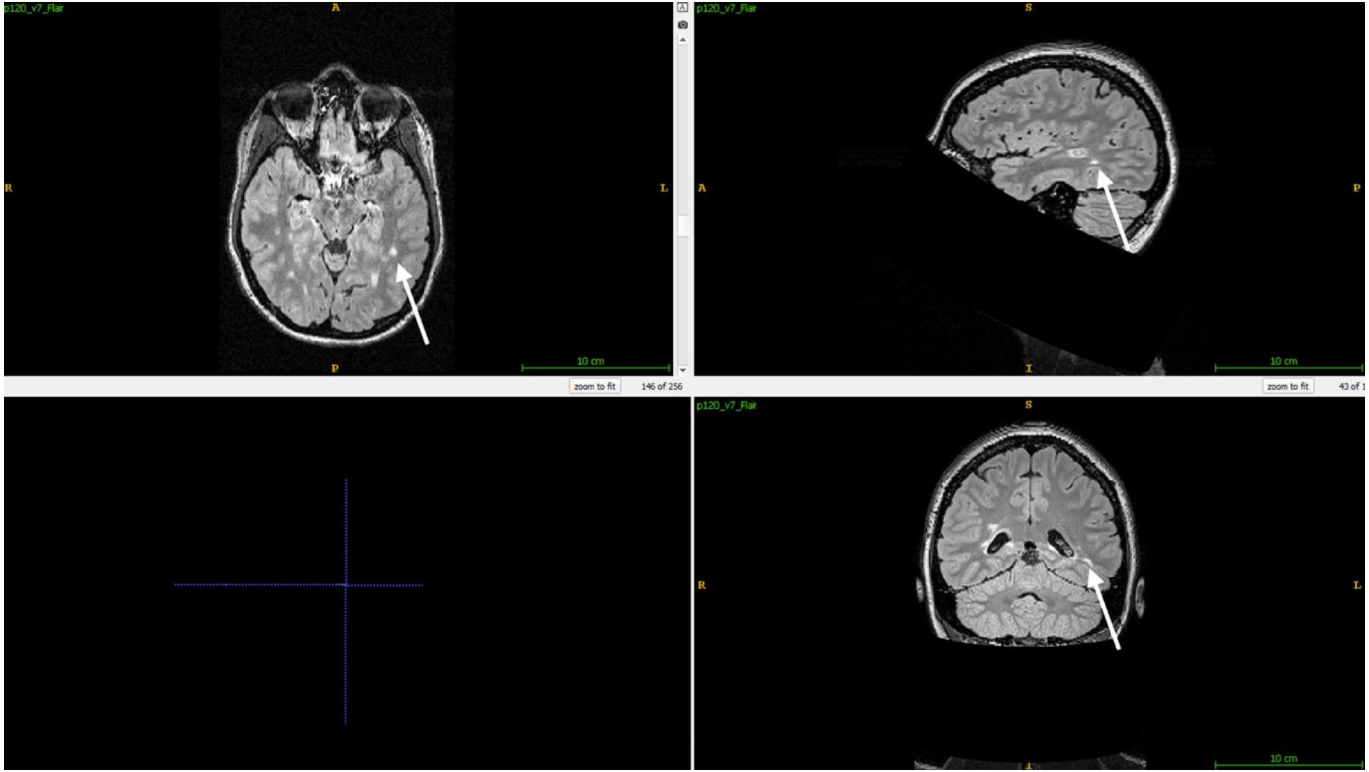
b UNIVERSITÄT BERN

Lesion Synthesis: Evaluation

- Winner: Conditional model with pseudo 3D-Unet

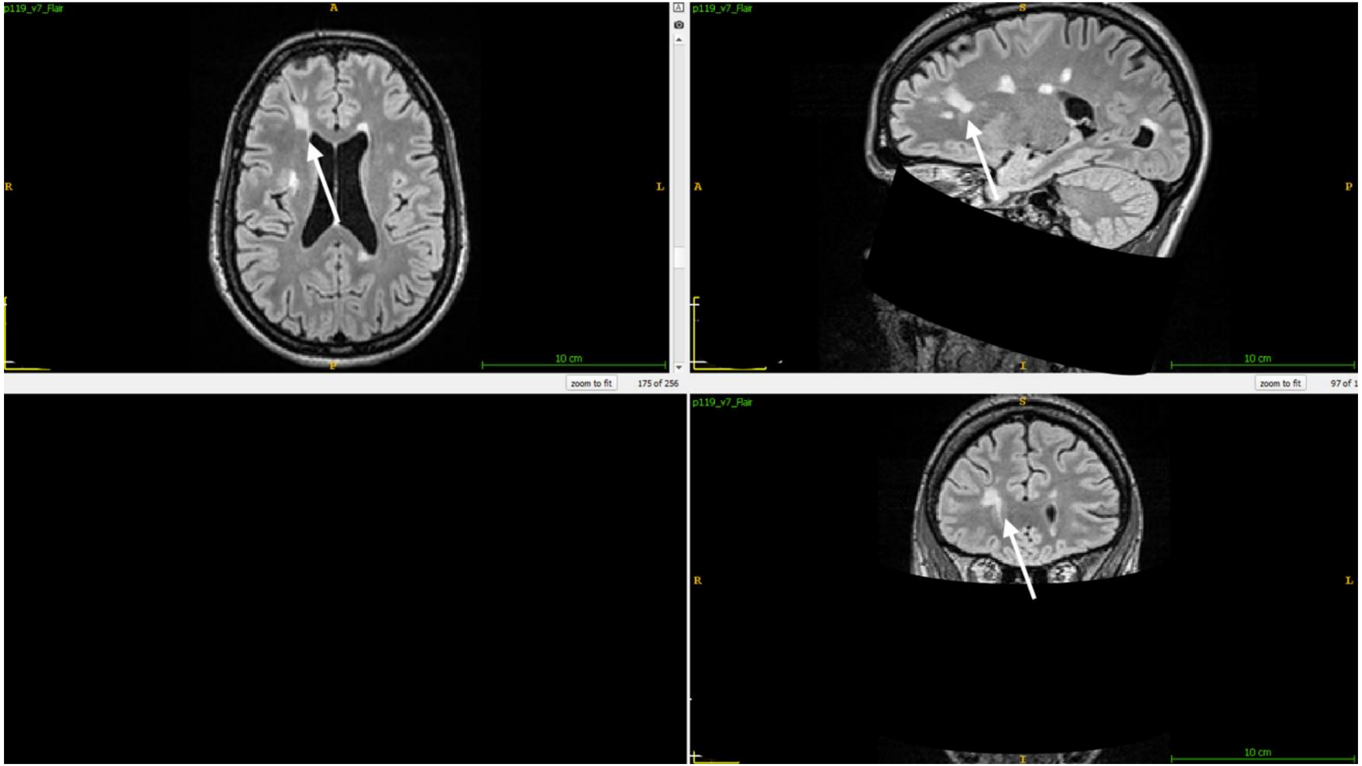
	SSIM	PSNR	MSE	LPIPS
2D Unconditional (with $t_0 = 1$ and median lesion intensity)	0.69	23.16	0.023	2.3e-3
3D Unconditional (with $t_0 = 3$ and median lesion intensity)	0.69	23.92	0.019	1.6e-4
3D Conditional	**0.79**	**27.13**	**0.009**	**1.3e-4**

Lesion Synthesis: Examples



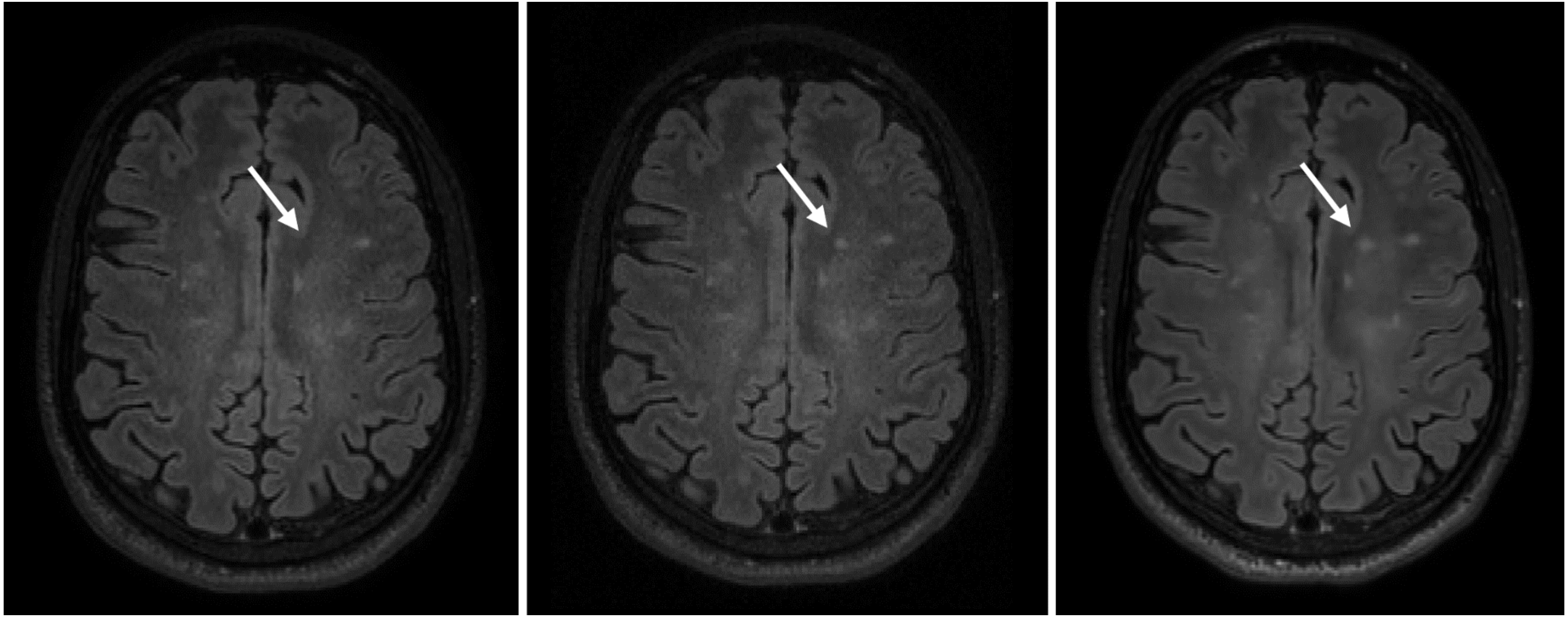
Example of synthetic lesion

Lesion Synthesis: Examples



Example of synthetic lesion

Lesion Synthesis: Examples



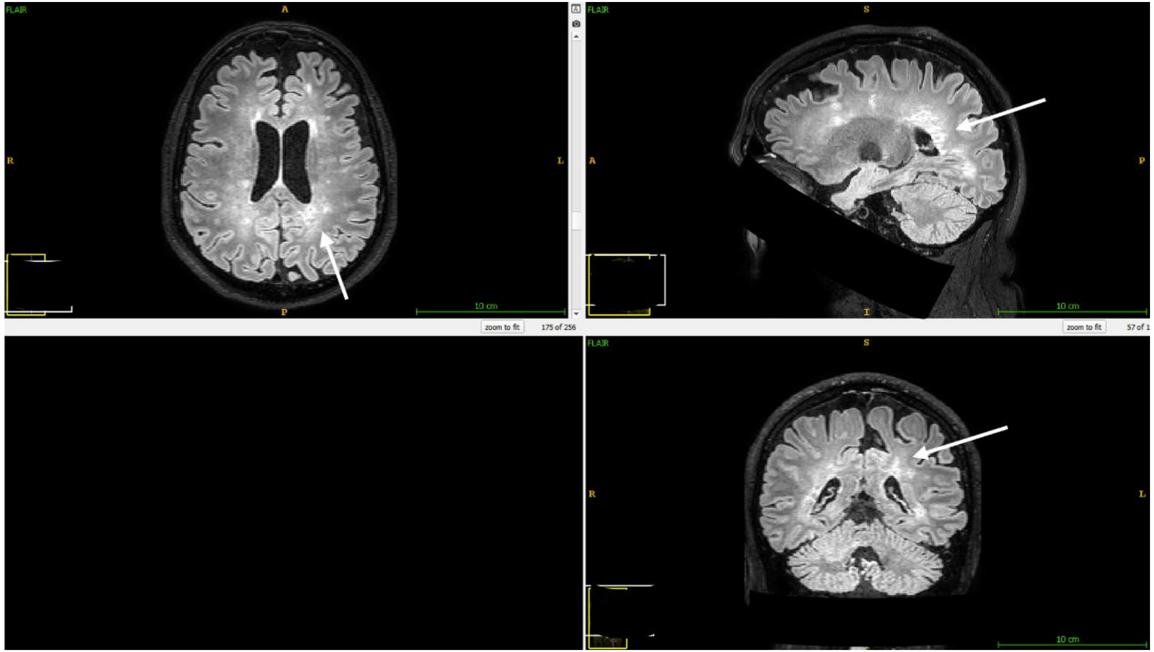
(1) Before and (2) after adding a synthetic lesion to healthy tissue and (3) at a later timepoint with a natural new lesion at the same location.

Lesion Synthesis: Qualitative Evaluation

- Set of 20 examples
- Neuroradiologist nr. one identified three synthetically added lesions
- Neuroradiologist nr. two detected only one.

Lesion Synthesis: Artifacts with unconditional model

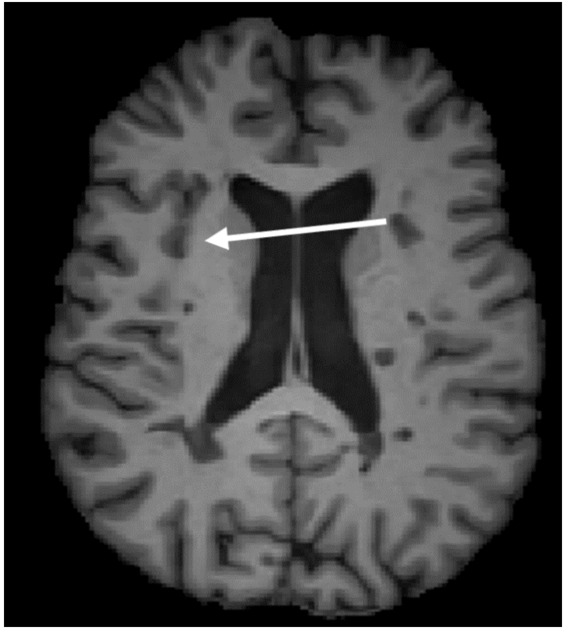
- Removal of coarse lesion
- High sensitivity to initial timestep and lesion intensity



Synthetic lesion: Partial lesion removal resulted in diffuse patterns

Lesion Synthesis: Artifacts with unconditional model

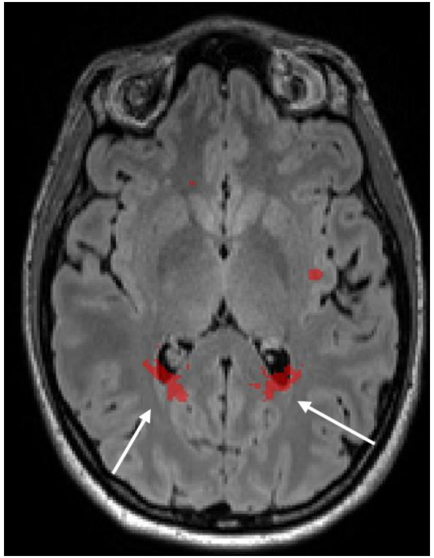
- Lesions got inpainted as gray matter with T1w images.



Synthetic lesion inpainted as GM

Lesion Synthesis: Mask Registration

- Inpainting needs masks
- Registration of lesion masks across patients has often poor results



Poorly registered lesion mask

Discussion

Lesion Filling

- Good results
- Conditional model was better
- RePaint converges faster, but overfits
- Pseudo 3D U-Net and synthetic random circle masks improved performance

- Robustness varies between different tools
- Deep learning tools are more robust to WM lesions
- Lesion filling might become obsolete in the future

Lesion Synthesis

- Good results with conditional model
- Pseudo 3D U-Net improved performance
- Unconditional model never delivered satisfactory results

- Next step: Train model with synthetic dataset
- Better mask registration process is needed

Conclusions

- We could train good models for lesion filling and synthesis
- Lesion filling might become obsolete
- Next step for lesion synthesis is train a model with synthetic dataset

Thank you for your attention!

Vinzenz Uhr
vinzenz.uhr@gmail.com

Thesis repository:
https://github.com/vinzenzuhr/Thesis_Diffusion_Lesions

Bibliography

1. M. Battaglini, M. Jenkinson, and N. De Stefano. Evaluating and reducing the impact of white matter lesions on brain volume measurements. Human Brain Mapping, 33(9), 2012.
2. J. Ho, A. Jain, and P. Abbeel. Denoising diffusion probabilistic models. In Advances in Neural Information Processing Systems, volume 2020-December, 2020.
3. A. Lugmayr, M. Danelljan, A. Romero, F. Yu, R. Timofte, and L. Van Gool. RePaint: Inpainting using Denoising Diffusion Probabilistic Models. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, volume 2022-June, 2022.
4. O. Ronneberger, P. Fischer, and T. Brox. U-Net: Convolutional Networks for Biomedical Image Segmentation. 5 2015.

Thank you for your attention!

Vinzenz Uhr
vinzenz.uhr@gmail.com

Thesis repository:
https://github.com/vinzenzuhr/Thesis_Diffusion_Lesions

Backup Slides

Noise Diffusion Models

$$q(x_t|x_{t-1}) = \mathcal{N}(x_t; \sqrt{1-\beta_t}x_{t-1}, \beta_t I)$$

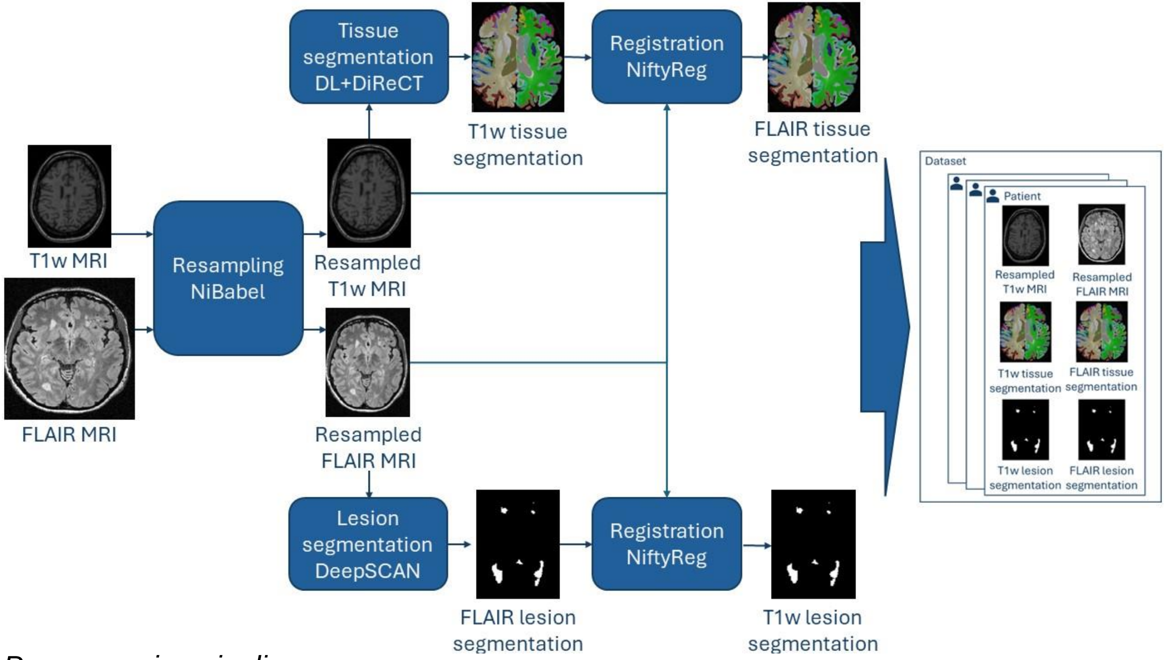
Data

Noise

$$p_\theta(x_{t-1}|x_t) = \mathcal{N}(x_{t-1}; \mu_\theta(x_t;t), \sigma_t^2 I)$$

Datasets

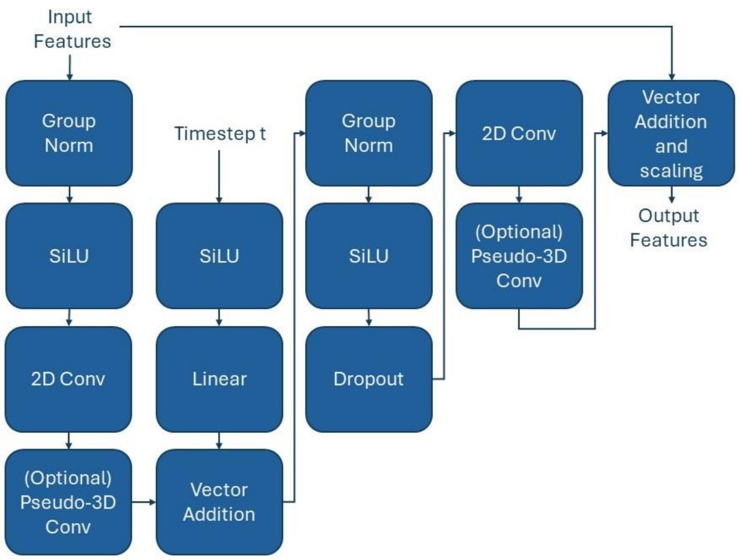
Preprocessing



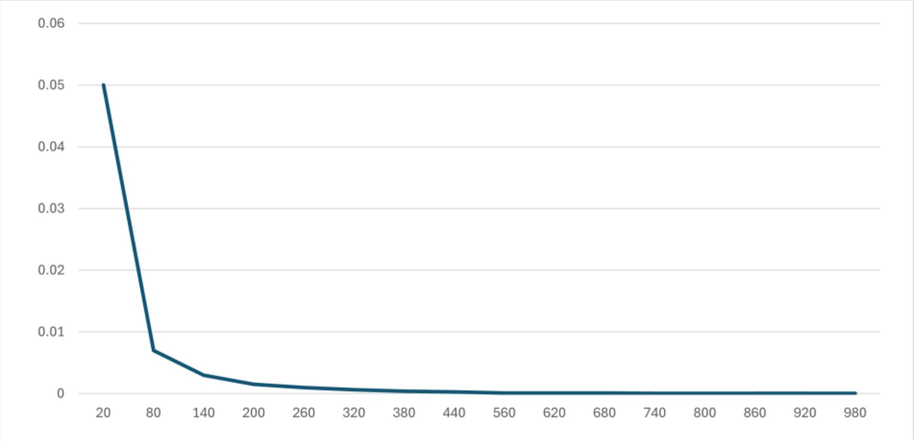
Preprocessing pipeline

Pseudo 3D U-Net

- six feature map resolutions with two convolutional residual blocks per resolution level and one self-attention block.
- From highest to lowest resolution the U-Net stages use (128, 128, 256, 256, 512, 512) channels.

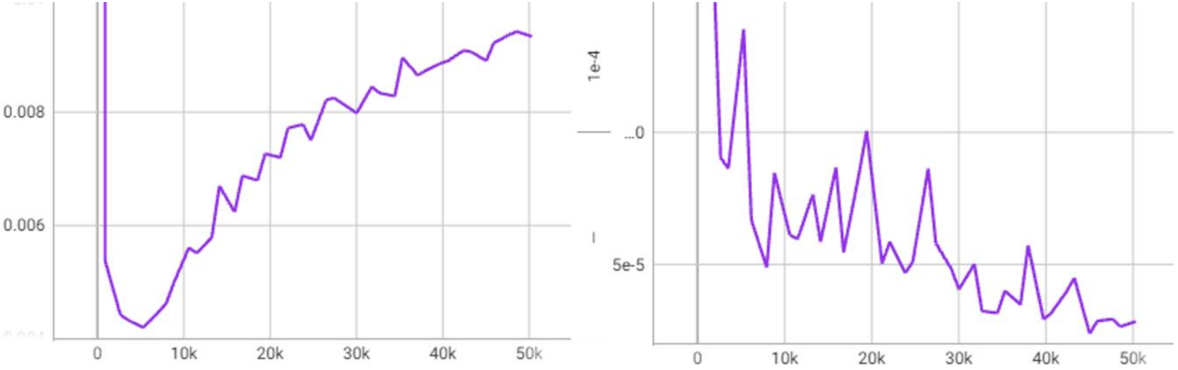


Min-SNR loss

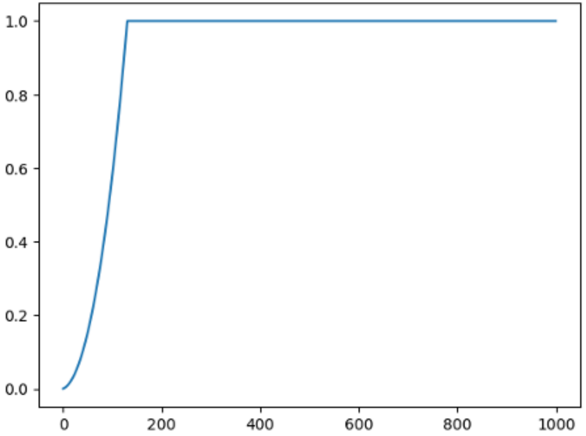


Validation loss of all 1000 timesteps

- Smaller timesteps overfit faster
- Min-SNR reweights loss per timestep



Validation loss of smaller vs bigger timestep



Loss timestep weights