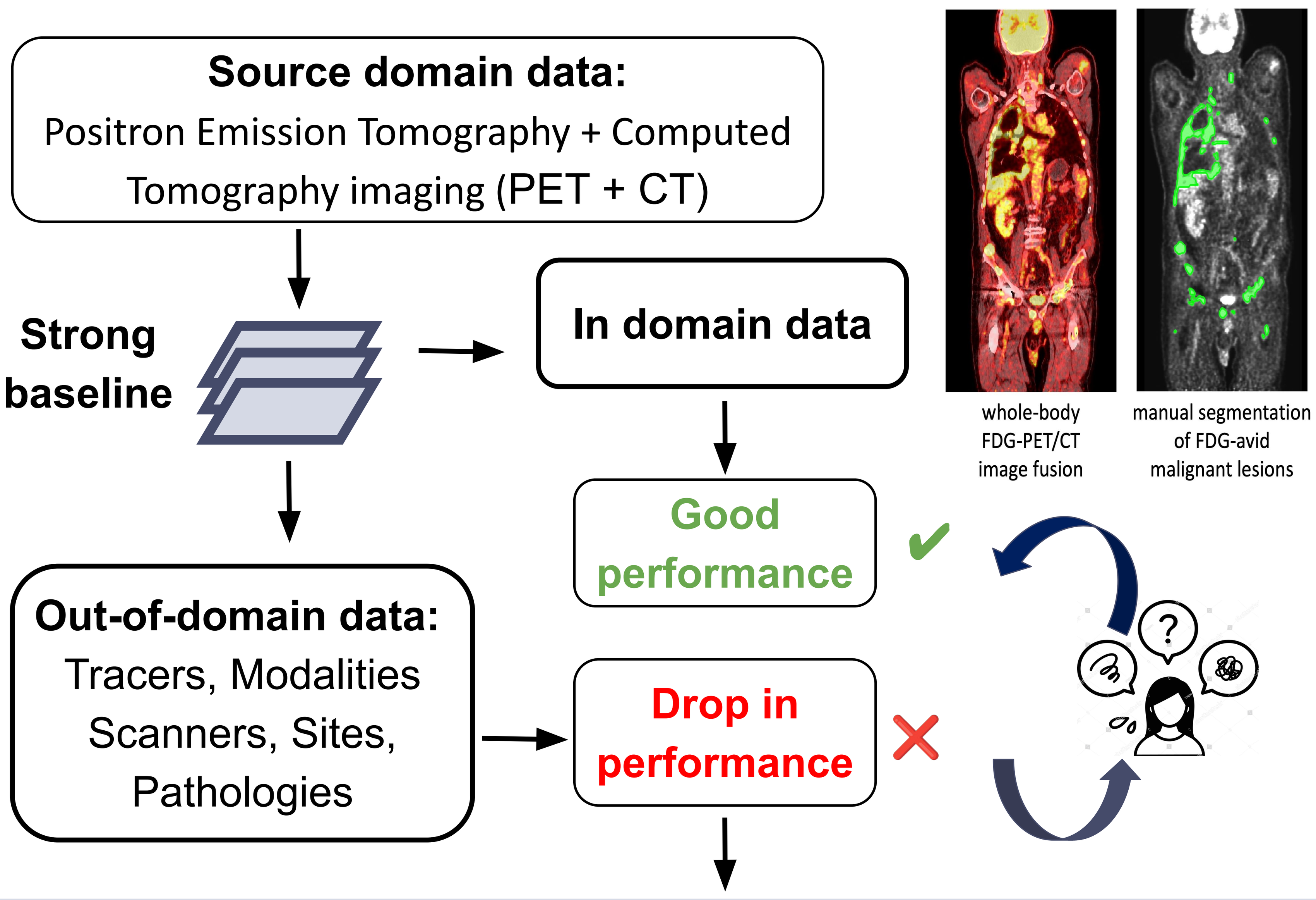
Out-of-domain Analysis in 3D Whole-Body PET/CT Lesion Segmentation

Pauline Ornela Megne Choudja[1], Christian la Fougère[2], Konstantin Nikolaou[3], Clemens Cyran[4], Michael Ingrisch[4], Sergios Gatidis[1], Thomas Kuestner [1]

[1] Medical Image And Data Analysis (MIDAS.lab), Department of Diagnostic and Interventional Radiology, University Hospital Tübingen, Germany.
[2] Department of Nuclear Medicine, University Hospital Tübingen, [3] Department of Radiology, University Hospital Tübingen, [4] Department of Radiology, University Hospital, LMU Munich.

MOTIVATION

Image segmentation models have shown remarkable success across various medical imaging applications[1]. However, their transition to oncological Positron Emission Tomography/Computer Tomography (PET/CT) imaging poses significant challenges, particularly when handling out-of-domain data, which impact the models' robustness and generalizability.



Aim

- **Out-Of-Domain (OOD) analysis**
- **Improve models' robustness and generalization**

METHODS

Experiments

Varying image inputs from the autoPET machine learning challenge PET/CT database[2]:

- **PET + CT (baseline)**
- **PET + CT filled with zeros**
- **PET filled with zeros + CT**
- **CT-only**

Models

1. **nnUNet**
   Finetuning (**nnUNet[3]** serves as baseline for segmentation)
2. **TransUNet**
   Training (**TransUNet[4] = nnUNet + Transformer blocks**)

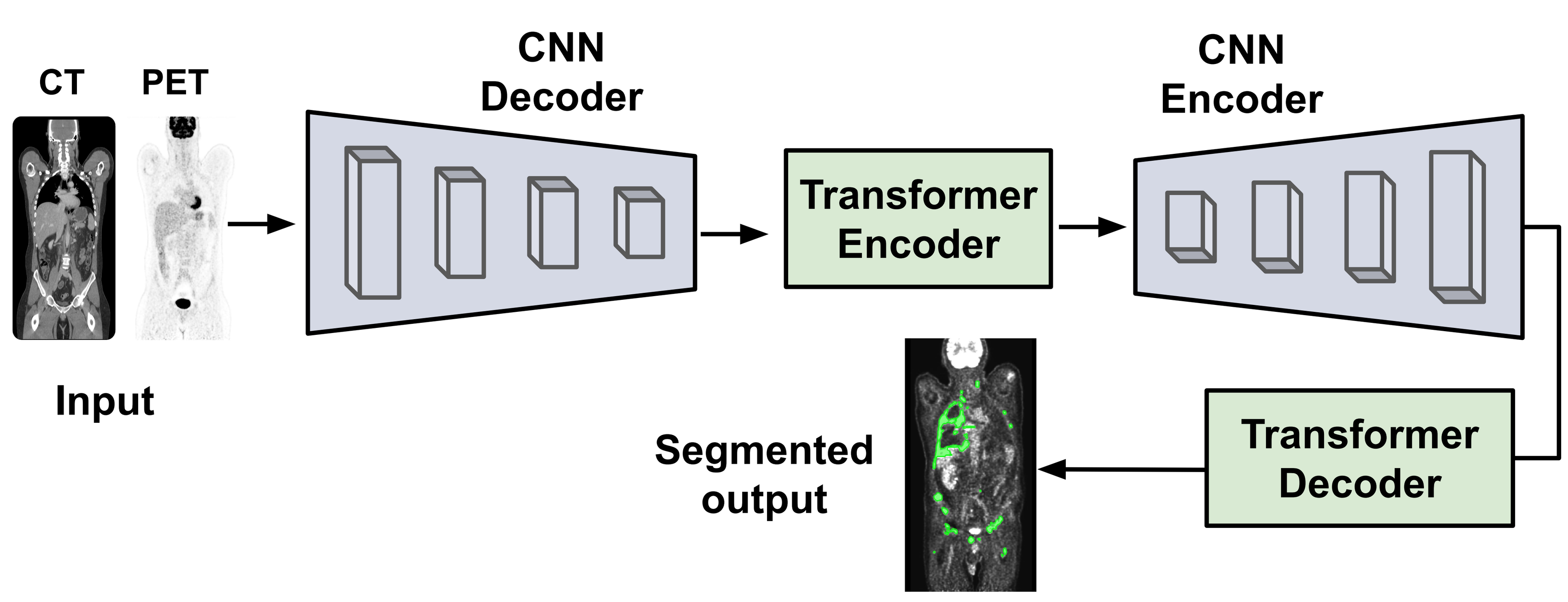


Figure 1: Overview of nnUNet (blue) & TransUNet (blue + green).

Metric

Volume rate analysis

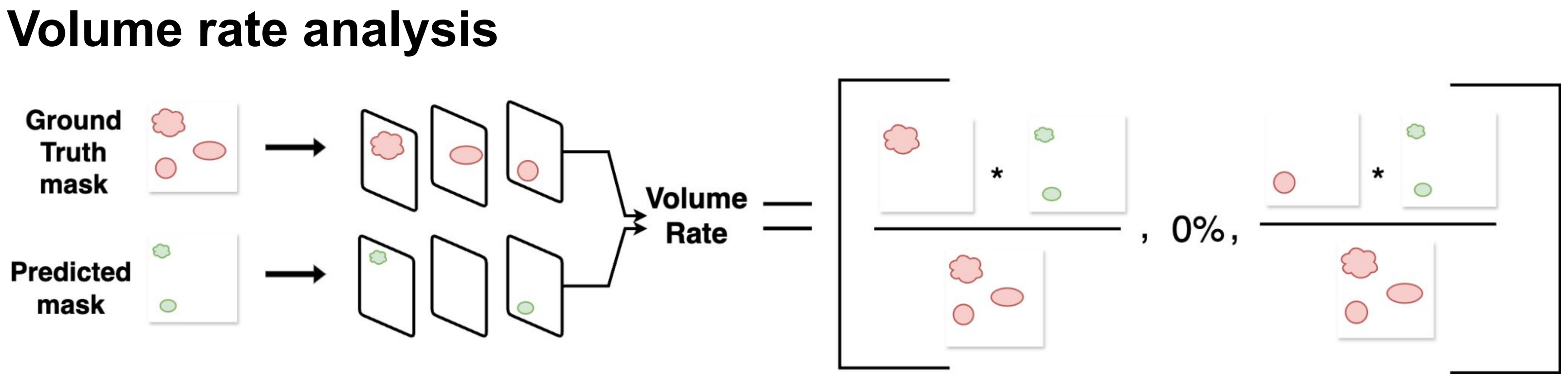
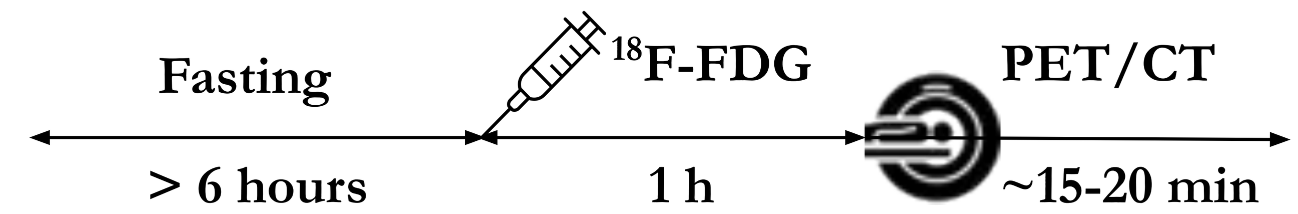


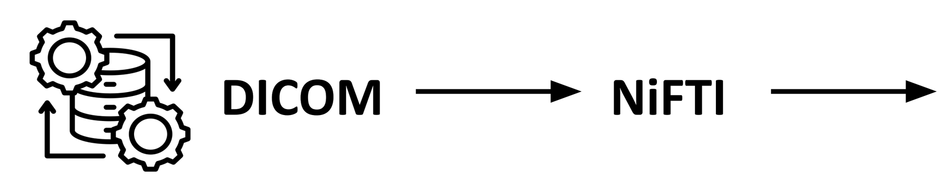
Figure 2: Volume Rate illustration.

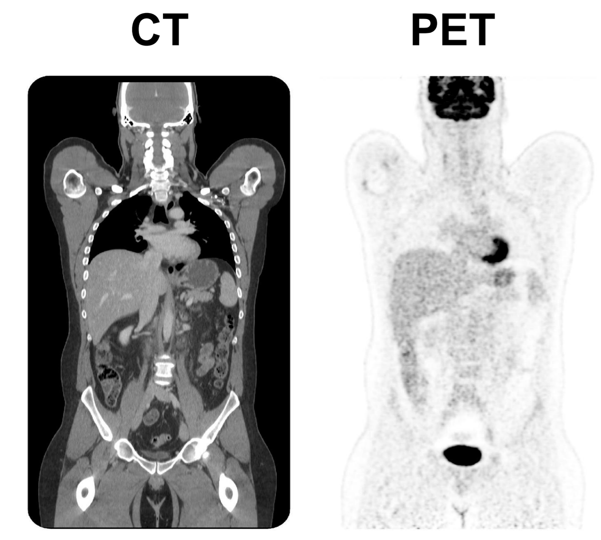
$$\text{Volume Rate(VR)}_i = \frac{\text{GT}_{lv_i} \times \text{Pred\_volume}}{\text{GT\_volume} + \alpha \cdot \max(0, \text{Pred\_volume} - \text{GT\_volume})}$$

DATA



- **Training cohort**: 1014 studies from 900 patients acquired in Tübingen.
- **Test cohort**: 150 studies from 2 imaging centers (Tübingen, Munich).





RESULTS

Table 1: OOD performance in AutoPET I for fine-tuned nnUNet and trained TransUNet. (FN: False Negative Volume; FP: False Positive Volume; SQ: Segmentation Quality; RQ: Recognition Quality; PQ: Panoptic Quality.)

Test cases	Metrics (mean values)						
	FN (ml)	FP (ml)	Dice score (%)	IoU (%)	SQ (%)	RQ (F1 score)	PQ (%)
	nnUNet (epochs=1000)						
(a) PET/CT	1.24	9.5	72.2	65.0	66.6	54.7	47.4
(b) PET/CT(zeroed)	1.43	8.65	62.4	52.8	58.6	39.1	31.6
(c) PET(zeroed)/CT	30.43	126.35	23.40	15.69	24.03	3.10	2.0
(d) CT	27.95	112.85	24.5	16.5	25.0	3.10	2.0
	TransUNet (epochs=125)						
(c) PET(zeroed)/CT	47.26	5.20	46.0	38.0	33.6	30.0	24.6
(d) CT	69.28	4.22	35.0	30.3	25.1	24.0	22.0

Volume rate and dice distributions

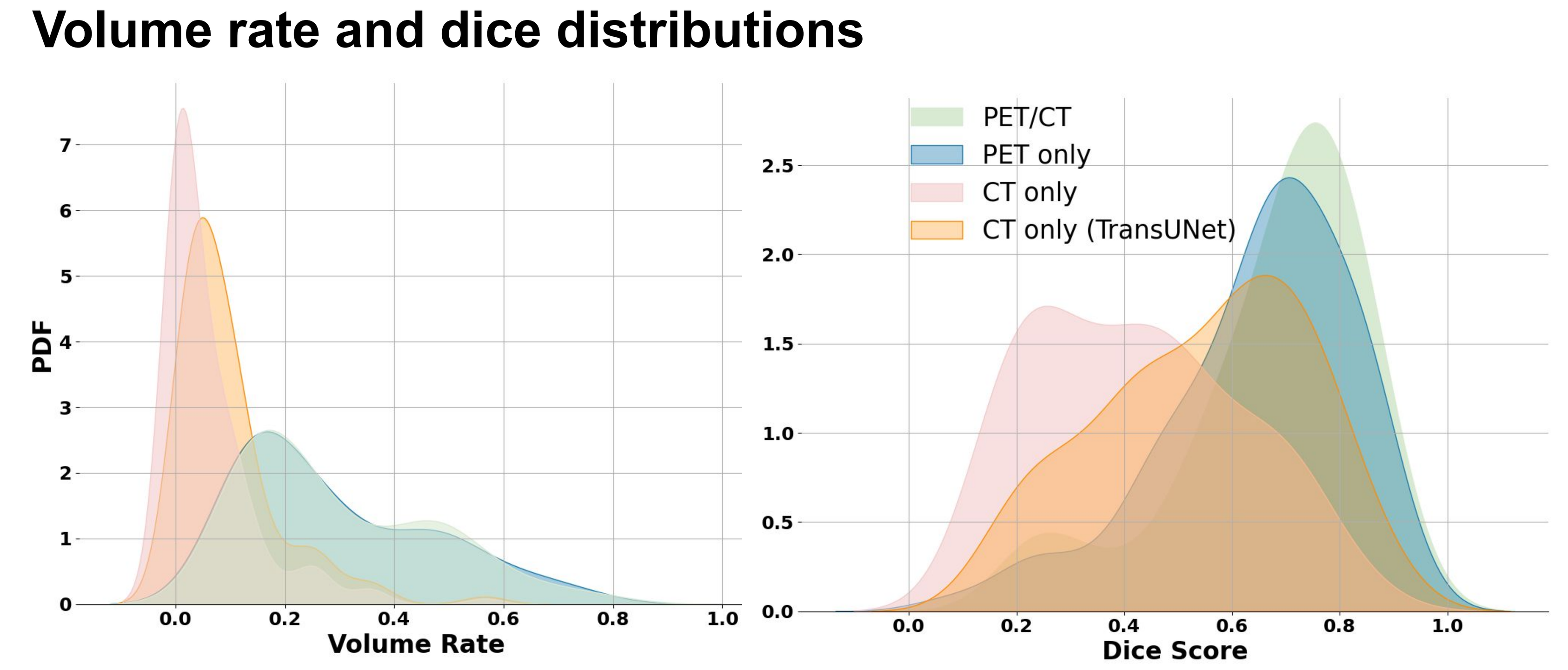


Figure 3: Out-of-domain behavior for lesion segmentation of nnUNet and TransUNet when trained and inferred with varying imaging modality inputs.

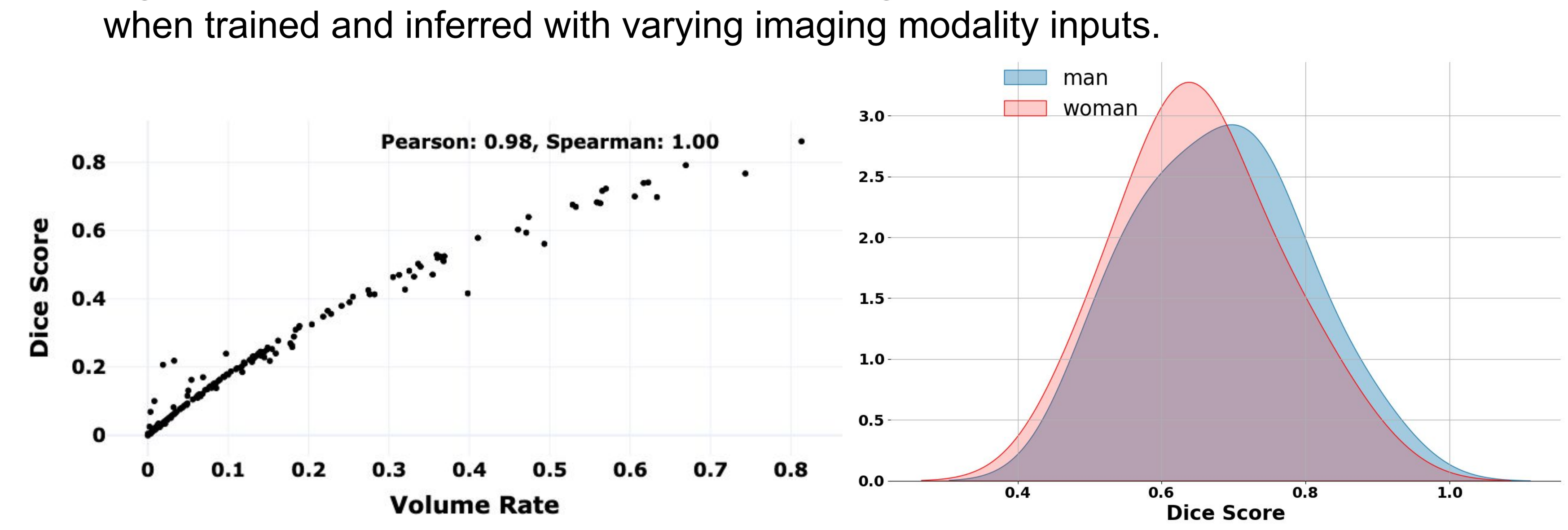


Figure 4: Volume Rate-Dice scatter plot and gender-based dice distributions for TransUNet.

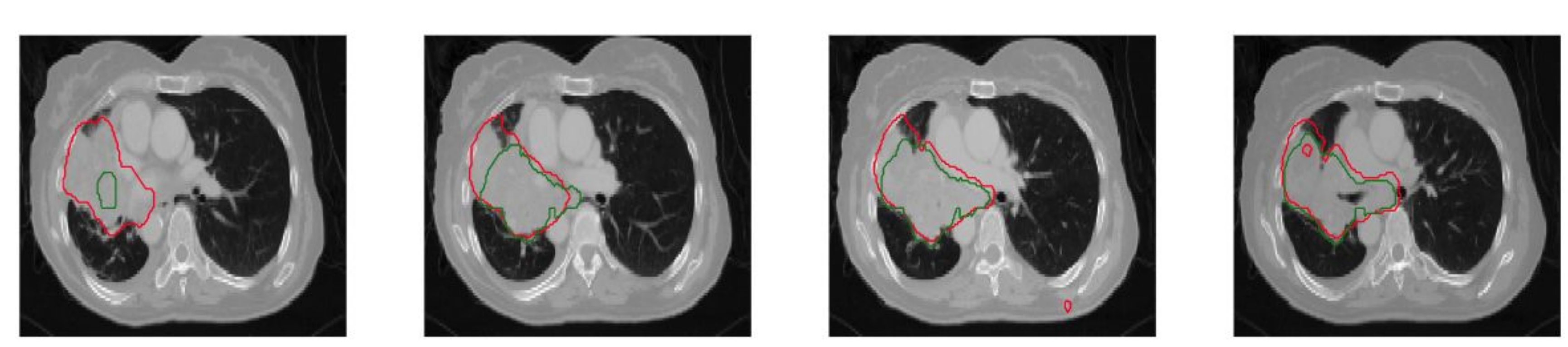


Figure 5: Axial image slice visualization showing the ground truth mask (red) and the predicted mask (green).

CONCLUSION

The proposed approach enhances model evaluation and informs the development of more refined and clinically relevant metrics especially in OOD settings, ultimately guiding improvements in model design. We aim to further improve upon OOD behaviour to realize a clinical translation.

REFERENCES

[1] Wenjian Yao, et al. From CNN to Transformer: A Review of Medical Image Segmentation Models, 2023.
[2] Sergios Gatidis, T Küstner, M Ingrisch, M Fabritius, and C Cyran. Zenodo, 3, 2022.
[3] Fabian Isensee, Paul F Jaeger, Simon AA Kohl, Jens Petersen, and Klaus H Maier-Hein, 2021
[4] Jieneng Che, et al. 3d transunet: Advancing medical image segmentation through vision transformers, 2023.

Presented by Pauline Ornela Megne Choudja (pauline-ornela.megne-choudja@student.uni-tuebingen.de)