Self-supervised learning via inter-modal reconstruction and feature projection networks for label-efficient 3D-to-2D segmentation

José Morano[1,2,✉], Guilherme Aresta[1,2], Dmitrii Lachinov[1,2], Julia Mai[2], Ursula Schmidt-Erfurth[2], Hrvoje Bogunović[1,2]

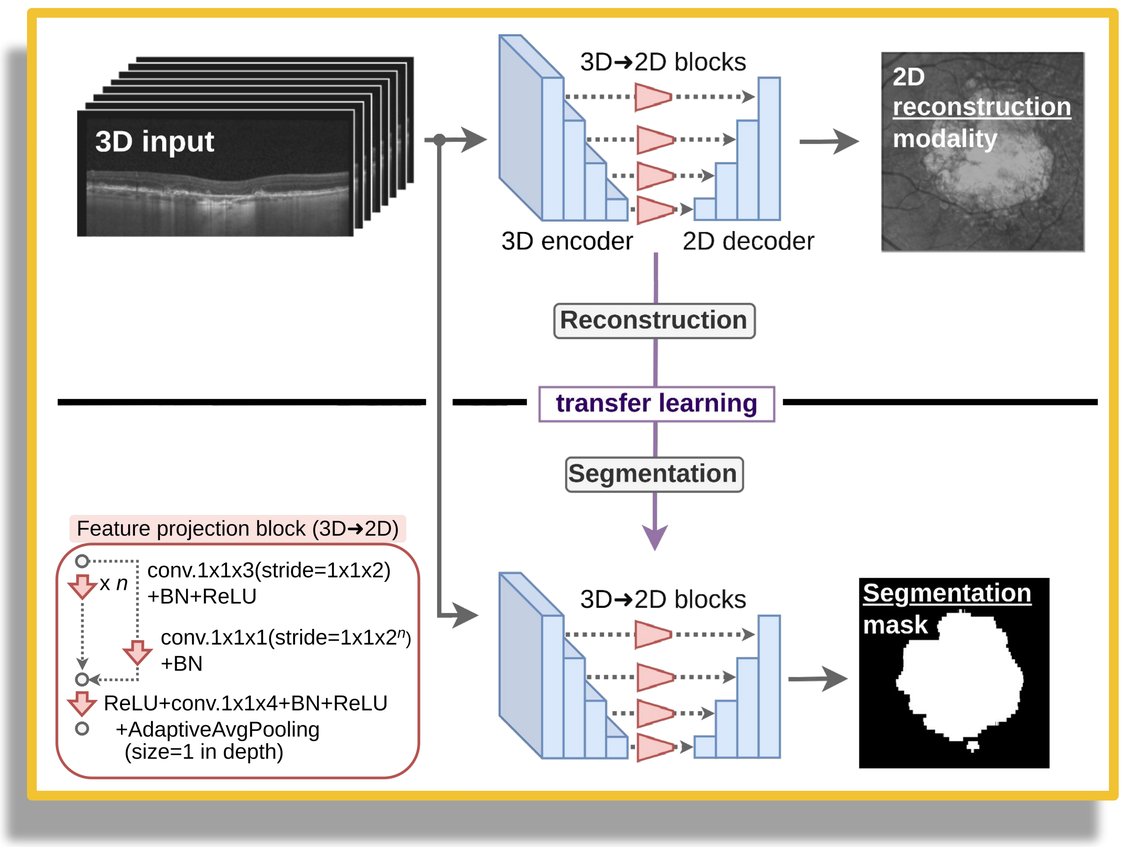
[1] Christian Doppler Laboratory for Artificial Intelligence in Retina, Dept. of Ophthalmology and Optometry, Medical University of Vienna, Austria
[2] Lab for Ophthalmic Image Analysis, Dept of Ophthalmology and Optometry, Medical University of Vienna, Austria

Introduction

Deep learning is a valuable tool for automating medical image segmentation tasks. Some tasks require segmentation to be performed on a subset of the data dimensions, the most common case being **3D-to-2D** [1]. The performance of existing methods is strongly conditioned by the amount of labeled data available, as no label efficiency method has been validated for these tasks. In this work, we propose a novel convolutional neural network (CNN) and self-supervised learning (SSL) method for label-efficient 3D-to-2D segmentation.

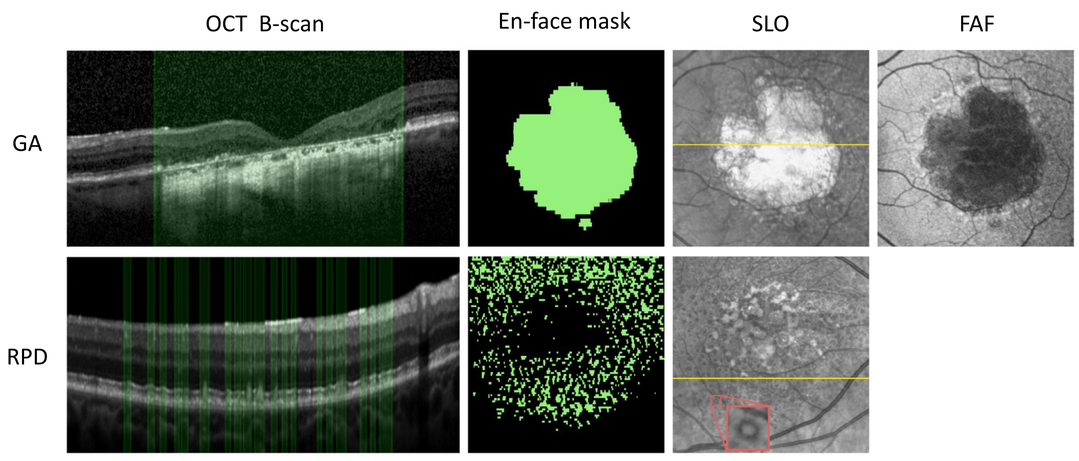
Methods

The SSL approach consists of training a novel CNN to **reconstruct image pairs** [2] from modalities with different dimensionalities and then fine-tuning it for the target segmentation tasks. The proposed CNN is composed of a 3D encoder and a 2D decoder [3] connected by **novel feature projection blocks (3D-to-2D)**.



Experiments and Results

The approach was validated on two clinically relevant tasks: the **segmentation of geographic atrophy** (GA) and **reticular pseudodrusen** (RPD) in optical coherence tomography (OCT), using fundus autofluorescence (FAF) and scanning laser ophthalmoscopy (SLO) as the modalities to reconstruct.

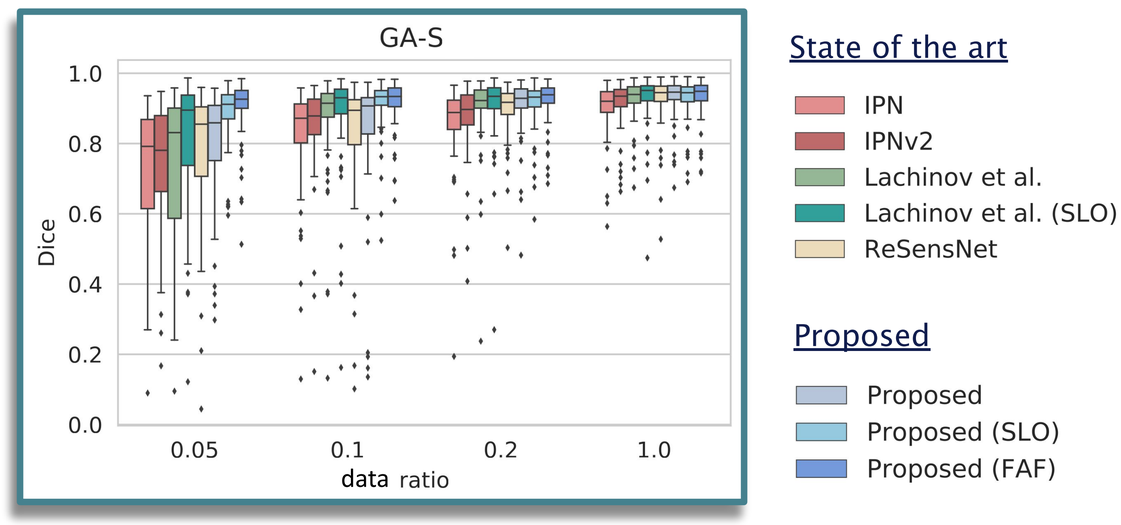


Paper and code available at https://github.com/j-morano/multimodal-ssl-fpn

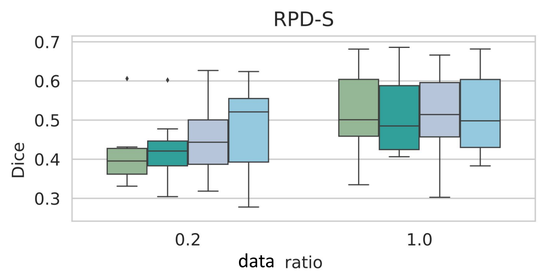
jose.moranosanchez@meduniwien.ac.at

Quantitative & qualitative results

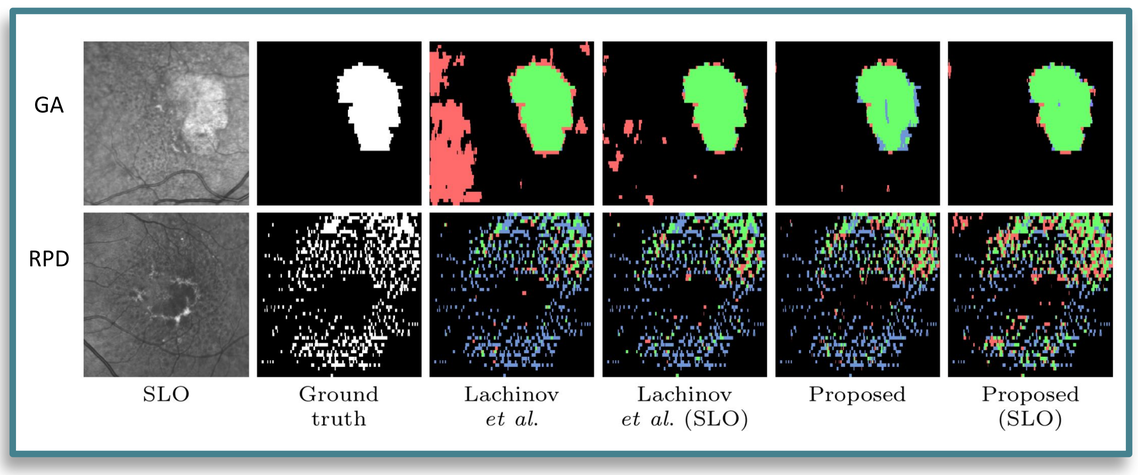
Results on different segmentation datasets (GA-S, RPD-S) and data ratios show that the proposed CNN significantly **improves the state of the art** [1] in label-scarce scenarios (by up to 8% in Dice score). Moreover, the SSL method allows further improvement of this performance (up to 23%). The proposed SSL is beneficial regardless of the network architecture.



GA-S (Dice)				
Method ↓ / Ratio →	0.05	0.1	0.2	1.0
IPN	$71.45 \pm 19.72^{***}$	$82.44 \pm 14.77^{***}$	$86.03 \pm 11.60^{***}$	$90.26 \pm 7.48^{***}$
IPNv2	$74.22 \pm 17.46^{***}$	$84.90 \pm 12.94^{***}$	$87.63 \pm 9.78^{***}$	$91.57 \pm 6.62^{***}$
Lachinov et al.	$73.46 \pm 21.50^{***}$	$86.89 \pm 13.75^{***}$	$90.20 \pm 9.98^{***}$	$92.89 \pm 5.48^{**}$
Lachinov et al. (SLO)	$81.78 \pm 17.68^{***}$	$88.45 \pm 13.78^{***}$	$90.98 \pm 9.58^{**}$	92.82 ± 7.31
ReSensNet	$78.47 \pm 17.44^{***}$	$83.81 \pm 16.02^{***}$	$89.94 \pm 7.31^{***}$	$92.70 \pm 7.22^{*}$
Proposed	$79.90 \pm 15.51^{***}$	$84.51 \pm 17.70^{***}$	$90.90 \pm 7.76^{***}$	93.15 ± 5.67
Proposed (SLO)	$88.36 \pm 8.98^{***}$	91.40 ± 7.49	$91.48 \pm 6.86^{***}$	$92.93 \pm 5.72^{***}$
Proposed (FAF)	$\mathbf{90.47 \pm 8.33}$	$\mathbf{91.61 \pm 6.52}$	$\mathbf{92.21 \pm 6.08}$	$\mathbf{93.15 \pm 5.71}$



RPD-S (Dice)		
Method ↓ / Ratio →	0.2	1.0
Lachinov et al.	41.34 ± 7.98	$\mathbf{51.62 \pm 11.07}$
Lachinov et al. (SLO)	42.24 ± 8.84	51.45 ± 9.96
Proposed	45.47 ± 10.31	50.60 ± 12.57
Proposed (SLO)	$\mathbf{47.55 \pm 12.65}$	48.28 ± 17.24



Conclusions

The proposed approach is an **effective solution for label-efficient 3D-to-2D segmentation**. Moreover, both the network and the SSL strategy have the potential to be used in other 3D-to-2D tasks, and the SSL strategy can be easily extended to other imaging domains.

Acknowledgements. This work was supported in part by the Christian Doppler Research Association, Austrian Federal Ministry for Digital and Economic Affairs, the National Foundation for Research, Technology and Development, and Heidelberg Engineering.

References

[1] Lachinov *et al.*, Projective skip-connections for segmentation along a subset of dimensions in retinal OCT, MICCAI 2021.
[2] Morano *et al.*, Multimodal transfer learning-based approaches for retinal vascular segmentation, ECAI 2020.
[3] Seeböck *et al.*, Linking function and structure with ReSensNet, *Ophthalmology Retina*, 2022.

Repository