U2-Net:
A Bayesian U-Net Model with Epistemic Uncertainty Feedback for Photoreceptor Layer Segmentation in Pathological OCT Scans

José Ignacio Orlando, Philipp Seeböck, Hrvoje Bogunović, Sophie Klimscha, Christoph Grechenig, Sebastian Waldstein, Bianca S. Gerendas, Ursula Schmidt-Erfurth

OPTIMA Ophthalmic Image Analysis

Christian Doppler Forschungsgesellschaft

MEDICAL UNIVERSITY OF VIENNA

World Health Organization

1.3 billion people

suffering some form of visual impairment

Age-Related Macular Degeneration (AMD)

Main cause of visual deficiency in industrialized countries

Global prevalence of 8.7% within 45-85 years old population

Diabetic Macular Edema (DME)

In 2017, 425 million people worldwide were suffering from diabetes

~10% developed vision-threatening DME

Retinal Vein Occlusion (RVO)

14-19 million people affected worldwide

AMD

DME

RVO

Photoreceptor
cell death

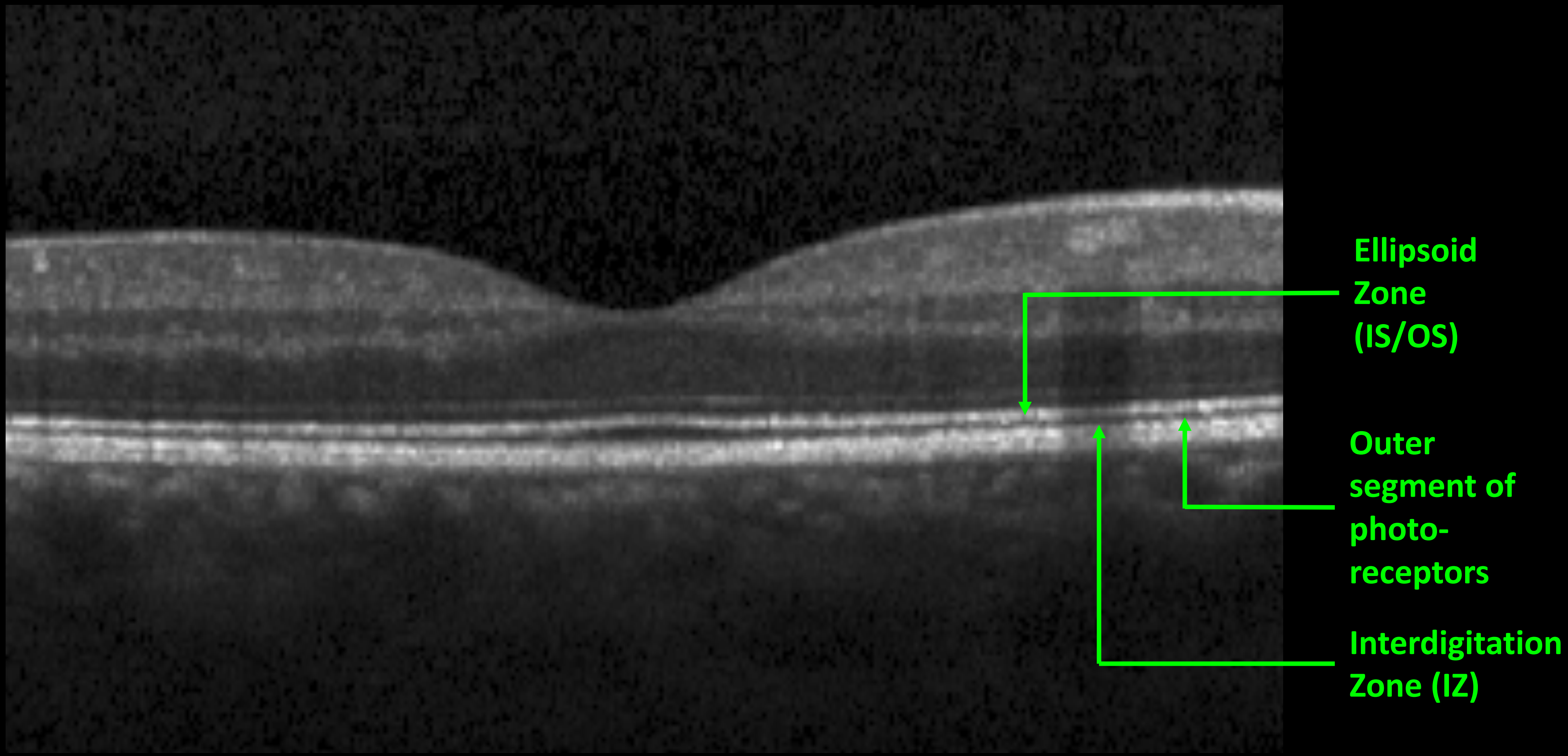
Visual acuity loss

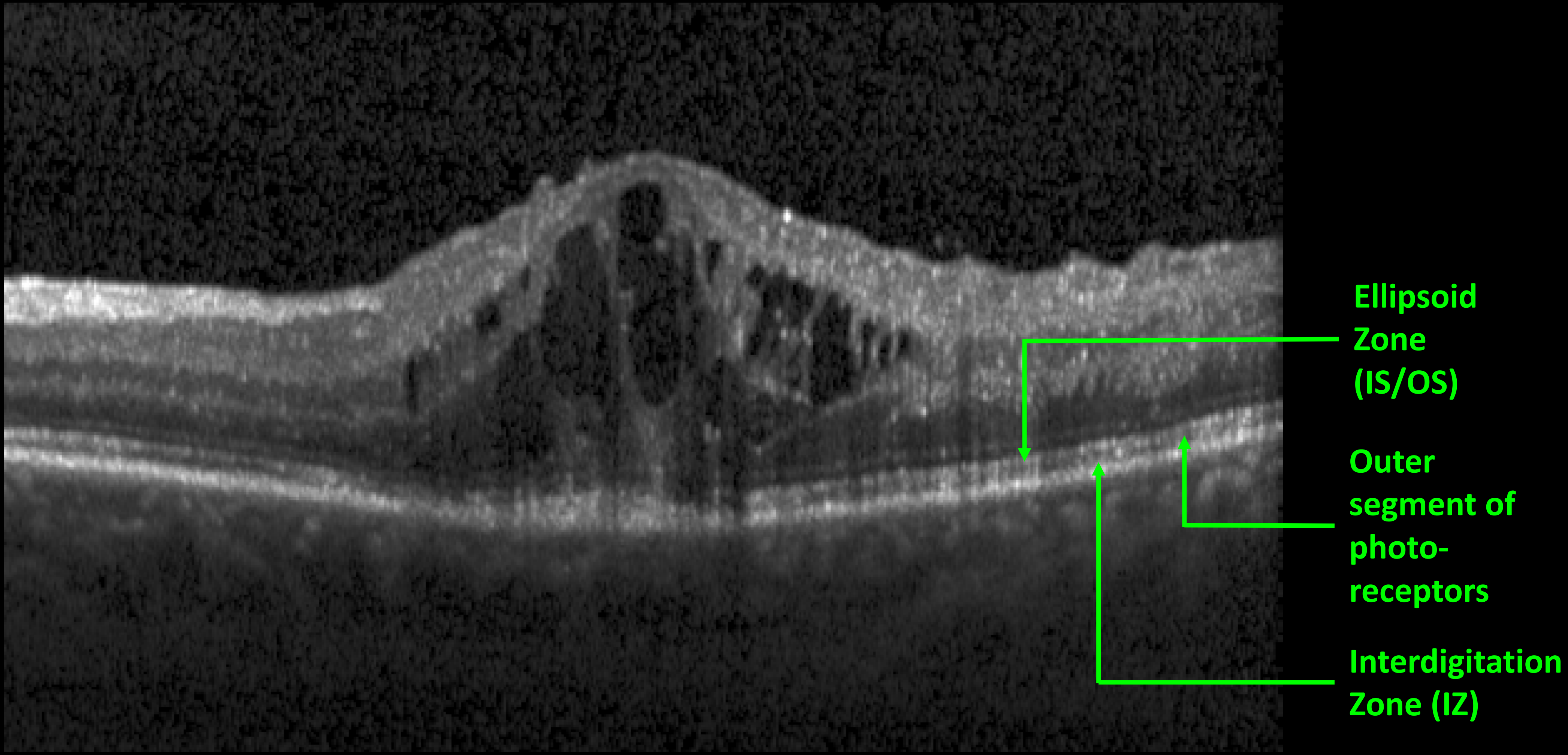
Optical Coherence Tomography (OCT)

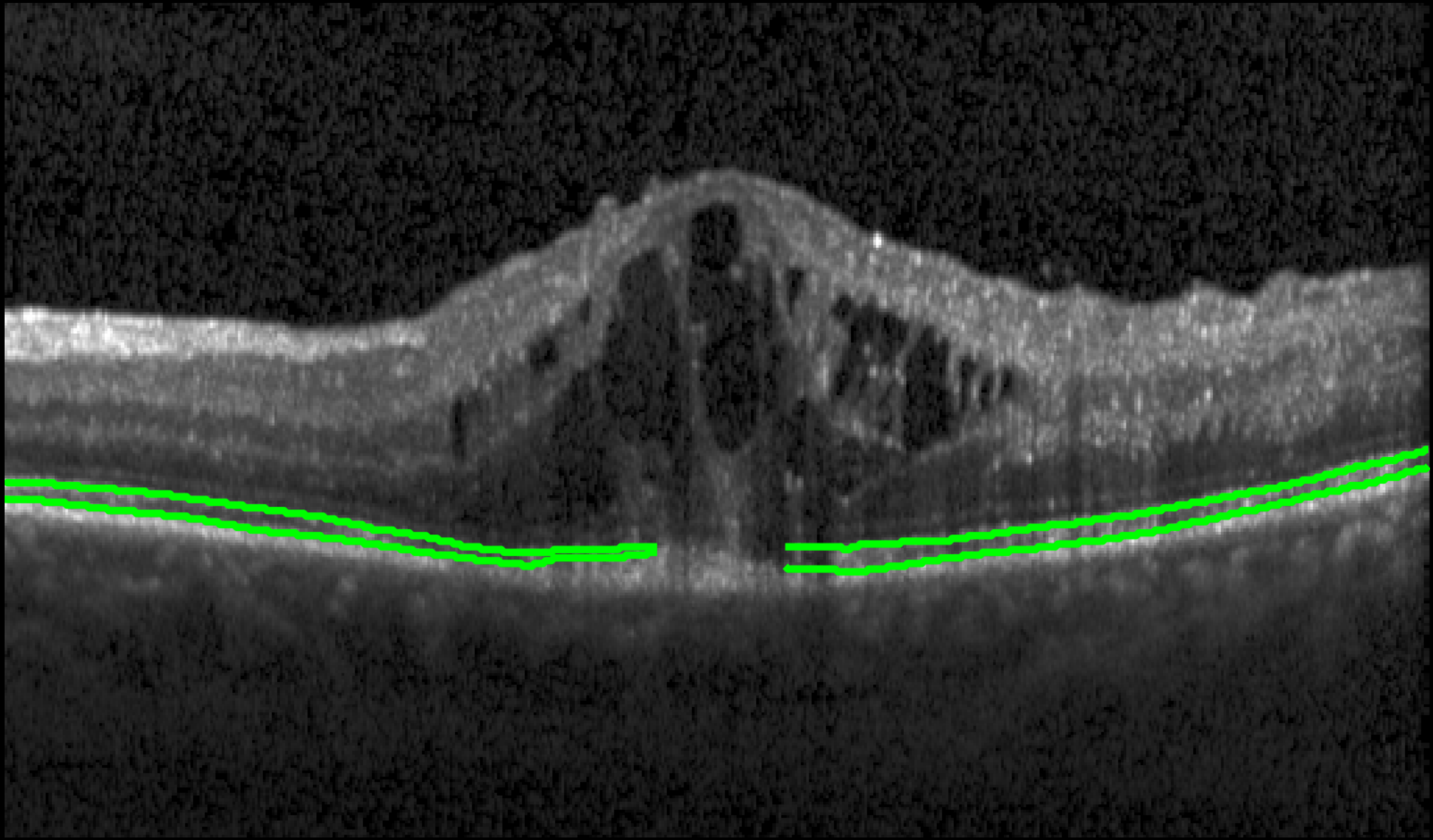
State-of-the-art imaging modality in AMD, RVO and DME

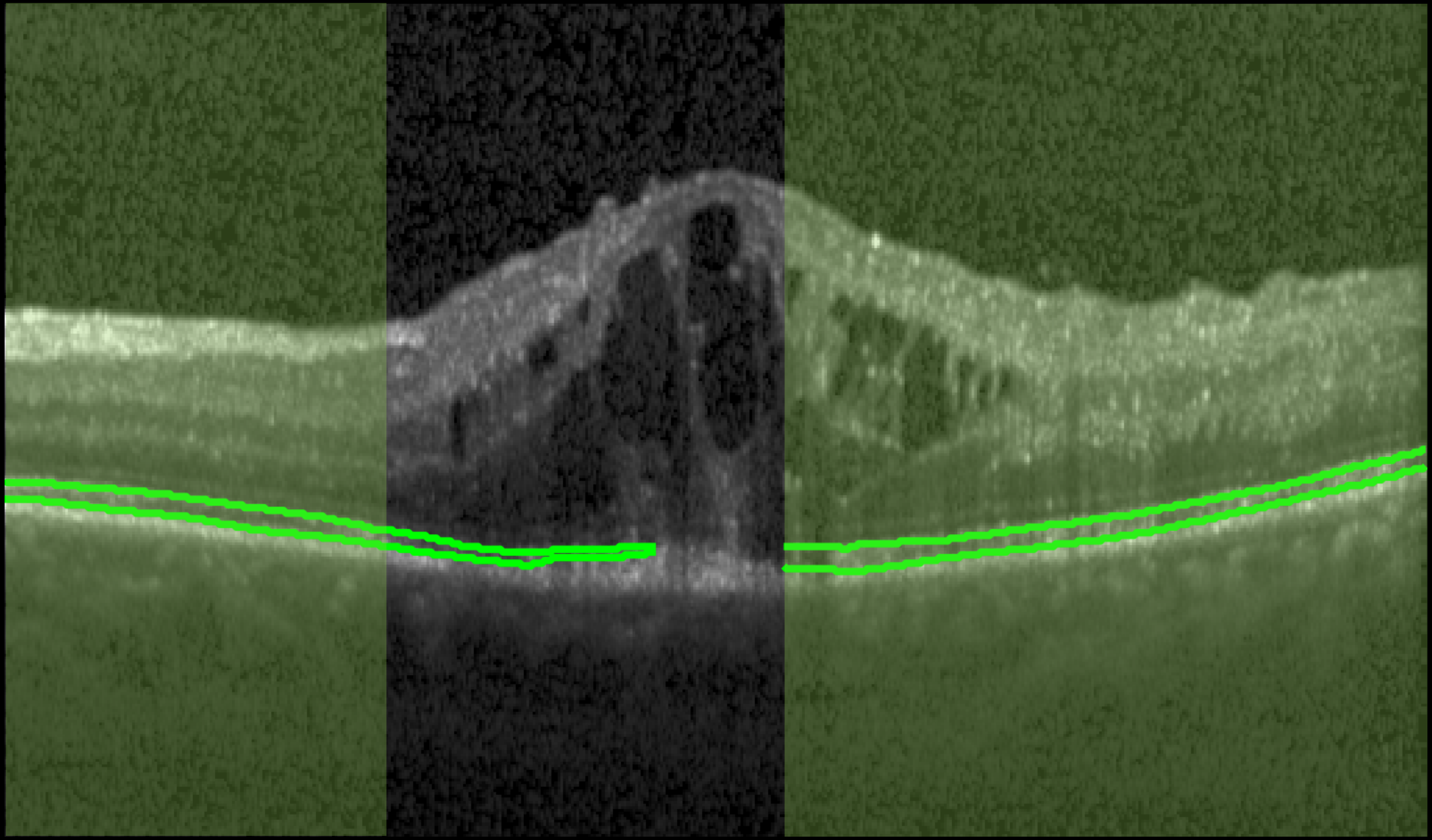
↓

Allows to assess photoreceptor integrity


Ellipsoid Zone (IS/OS)
Outer segment of photo-receptors
Interdigitation Zone (IZ)

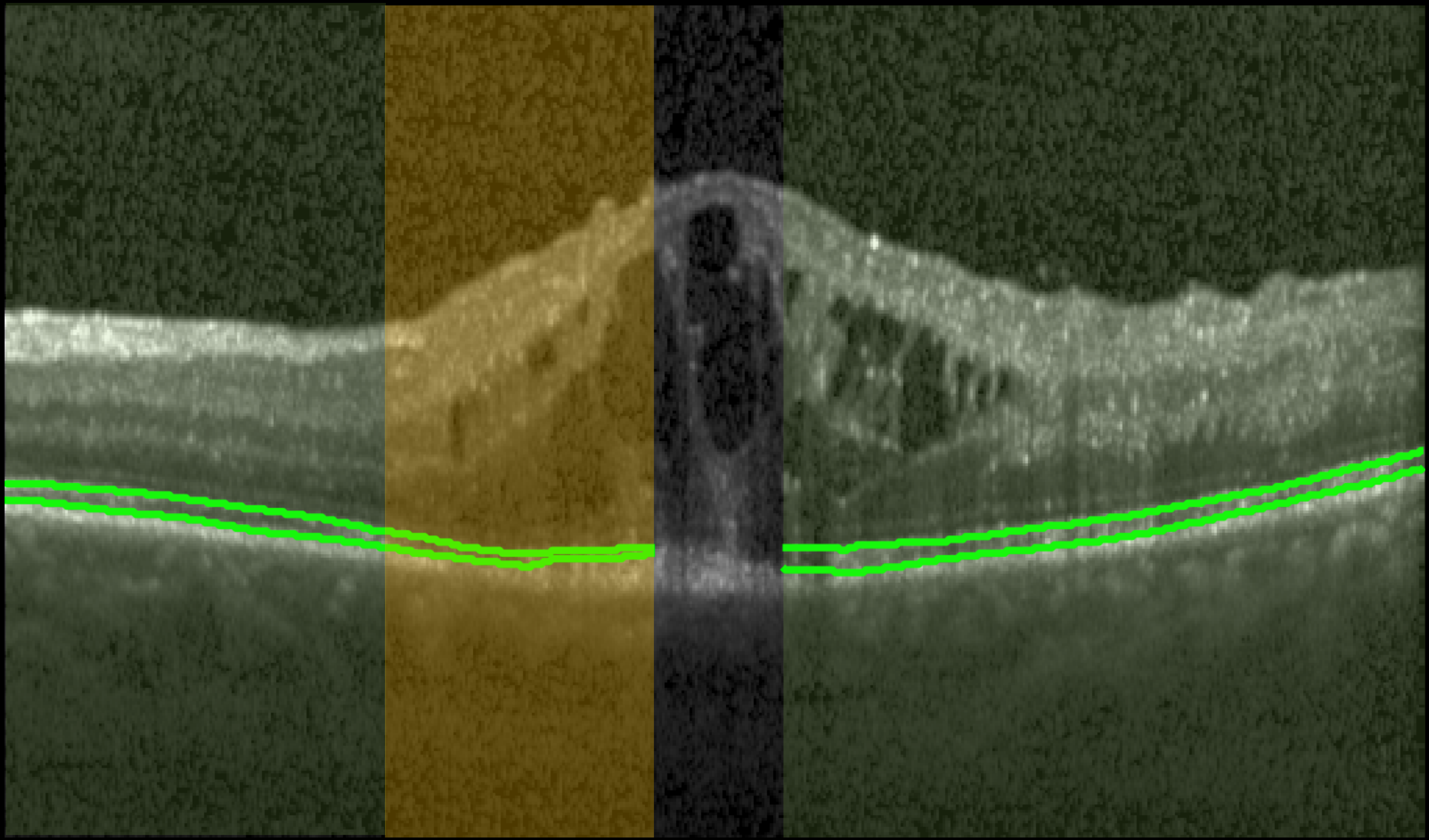

Ellipsoid Zone (IS/OS)
Outer segment of photo-receptors
Interdigitation Zone (IZ)



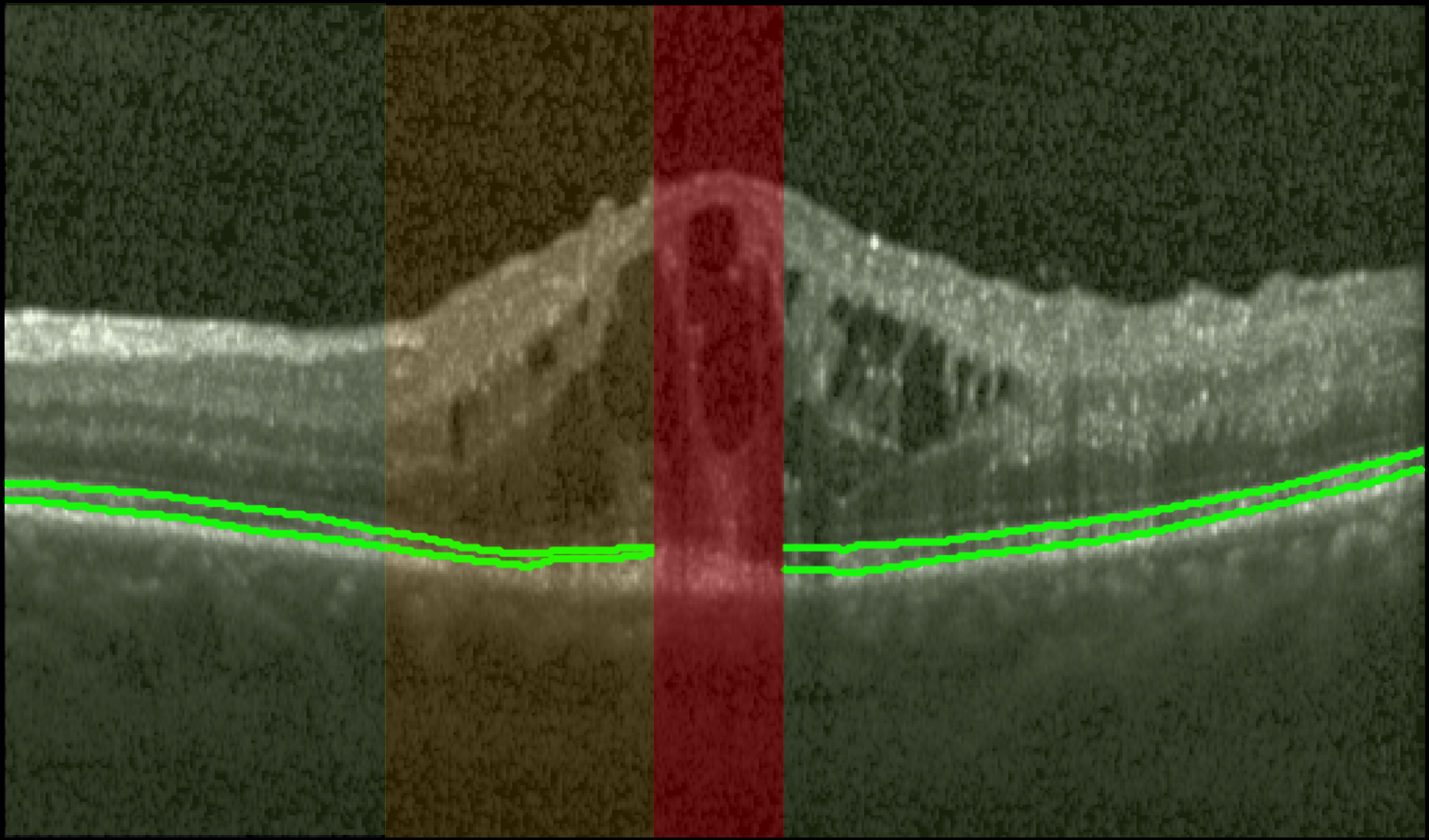


Normal photoreceptors

Normal photoreceptors



Abnormal thinning



Pathological disruption

Our mid-term goal

Understand the pathophysiological processes that cause damage in photoreceptor integrity

(i) Accurate segmentation

(ii) Interpretable feedback to correct the results

Key challenge

Pathological alterations

Ambiguous appearances turn difficult to produce reliable segmentations

Unfeasible to capture every possible pathological feature on a training set

Bayesian deep learning

Bayesian SegNet: Model Uncertainty in Deep Convolutional Encoder-Decoder Architectures for Scene Understanding

Alex Kendall, Vijay Badrinarayanan, Roberto Cipolla

(Submitted on 9 Nov 2015 (v1), last revised 10 Oct 2016 (this version, v2))

What Uncertainties Do We Need in Bayesian Deep Learning for Computer Vision?

Alex Kendall, Yarin Gal

(Submitted on 15 Mar 2017 (v1), last revised 5 Oct 2017 (this version, v2))

Bayesian deep learning

Model uncertainty

- **Aleatoric** — Task uncertainty, what we don't know and we will never learn
- **Epistemic** — Model uncertainty, what we don't know but we can learn given more training data

Bayesian deep learning

Model uncertainty

Aleatoric — Task uncertainty, what we don't know and we will never learn

Epistemic — Model uncertainty, what we don't know but we can learn given more training data

Epistemic uncertainty

BDL is used to compute a posterior distribution

$$p(\mathbf{W}|\mathbf{X}, \mathbf{Y})$$

Approximate distribution learned
by variational inference

$$q(\mathbf{W})$$

Bernoulli distribution to the weights of the i-th
convolutional layer using Dropout

$$q(\mathbf{W}_i) \longrightarrow p_i$$

(Gal et al., 2015)

Epistemic uncertainty

Monte Carlo sampling with dropout in test time

Averaging the outcomes results in better performance

Sampling multiple slightly different outputs

Monte Carlo sampling with dropout in test time

Standard deviation allows to retrieve an epistemic uncertainty estimate

Our approach

Uncertainty U-shaped Network

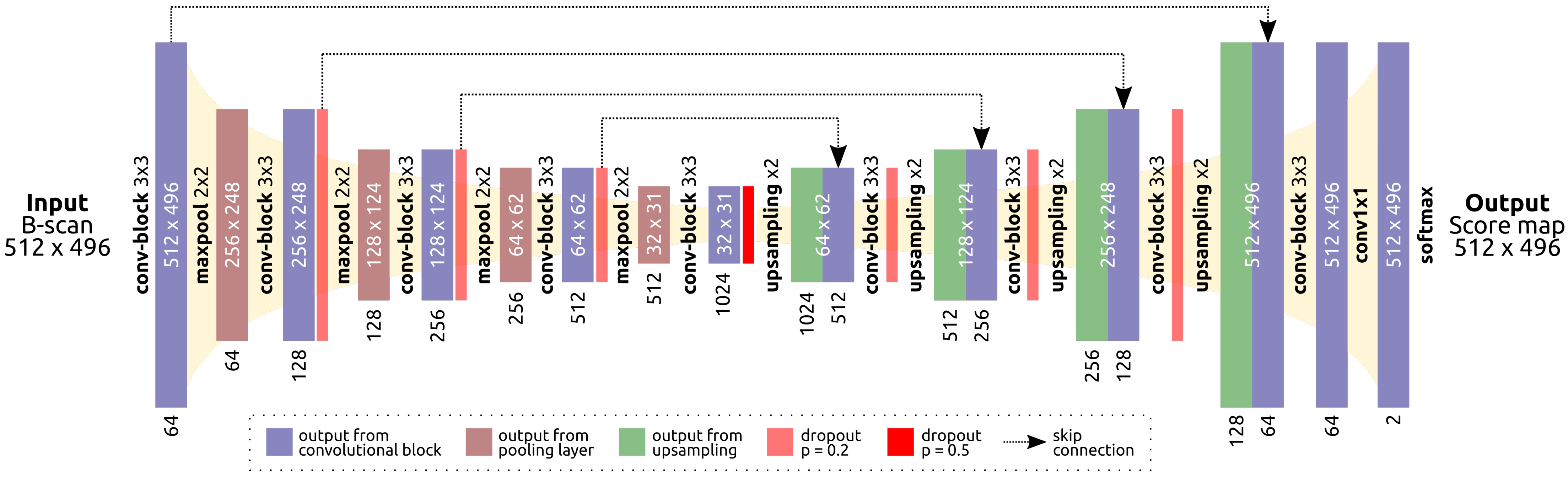
Our approach

Uncertainty U-shaped Network

Our approach

U2-Net

Standard U-Net + Nearest neighbor upsampling + Leaky ReLUs + Batch norm + Dropout



MC sampling with dropout in test time to predict average score map & epistemic uncertainty map

Materials

Data set A

AMD (early, CNV)	10 volumes	490 B-scans
DME	16 volumes	784 B-scans
RVO	24 volumes	1176 B-scans
Total	**50 volumes**	**2450 B-scans**

Split at a patient-basis preserving disease proportion

Training set	Validation	Test
31 volumes (1519 B-scans)	4 volumes (196 B-scans)	15 volumes (735 B-scans)

Data set B

Late AMD (GA)	**10 volumes**	**490 B-scans**

Separate test set

Test

10 volumes
(496 B-scans)

Evaluation metrics

Photoreceptors

- Area under Precision/Recall curve
- Dice index

Disruptions

- Area under Precision/Recall curve (at an A-scan level)

Baselines

Standard U-Net

(Ronneberger et al., MICCAI 2015)

Batch normalization, NN upsampling, dropout in bottleneck

BRU-Net

(Apostolopoulos et al., MICCAI 2017)

Branch residual U-Net with dilated convolutions and residual connections

BU-Net

Bayesian U2-Net with aleatoric uncertainty estimates

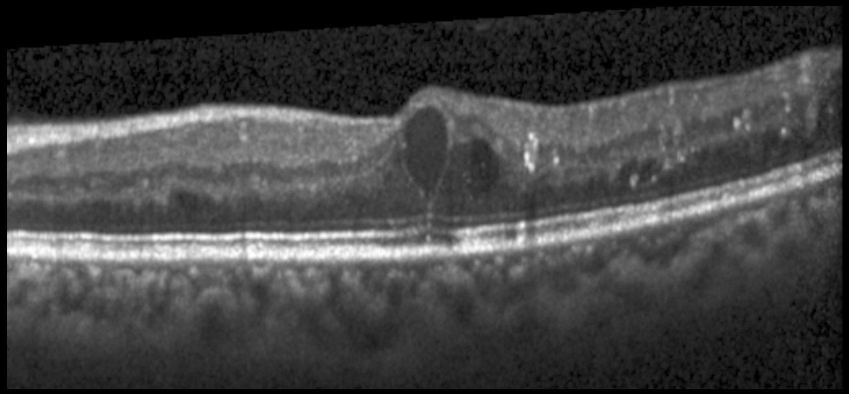
(Inspired in Nair et al., MICCAI 2018)

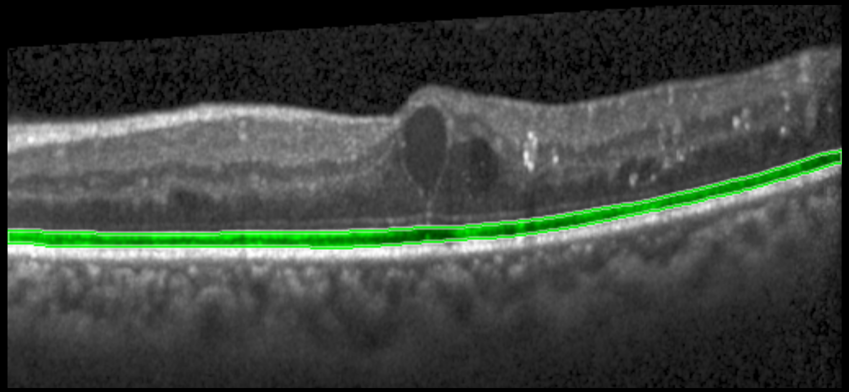
Results

Quantitative evaluation

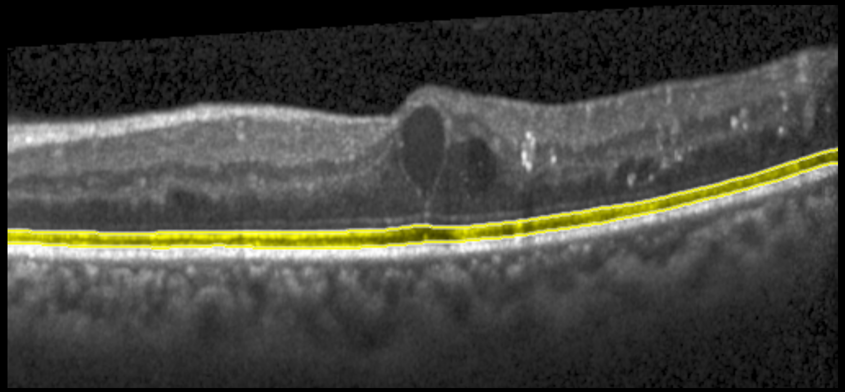
Model	Test set A AMD (early, CNV), DME, RVO			Test set B Late AMD (GA)		
	Photoreceptors		Disruptions	Photoreceptors		Disruptions
	AUC	Dice	AUC	AUC	Dice	AUC
U-Net [10]	0.9566	0.8815 $\pm$0.06	0.5077	0.9390	0.8375 $\pm$0.07	0.8795
BRU-Net [16]	0.9593	0.8767 $\pm$0.08	0.2621	0.9295	0.7890 $\pm$0.13	0.8333
BU-Net $T=1$	0.9466	0.8647 $\pm$0.08	0.2222	0.8969	0.7311 $\pm$0.14	0.8065
BU-Net $T=10$	0.9505	0.8678 $\pm$0.08	0.2405	0.8998	0.7428 $\pm$0.14	0.8129
U2-Net $T=1$	0.9653	0.8932 $\pm$0.04	**0.6712**	**0.9500**	**0.8546 $\pm$0.06**	0.9085
U2-Net $T=10$	**0.9669**	**0.8943 $\pm$0.04**	0.6417	0.9472	0.8457 $\pm$0.08	**0.9101**

B-scan



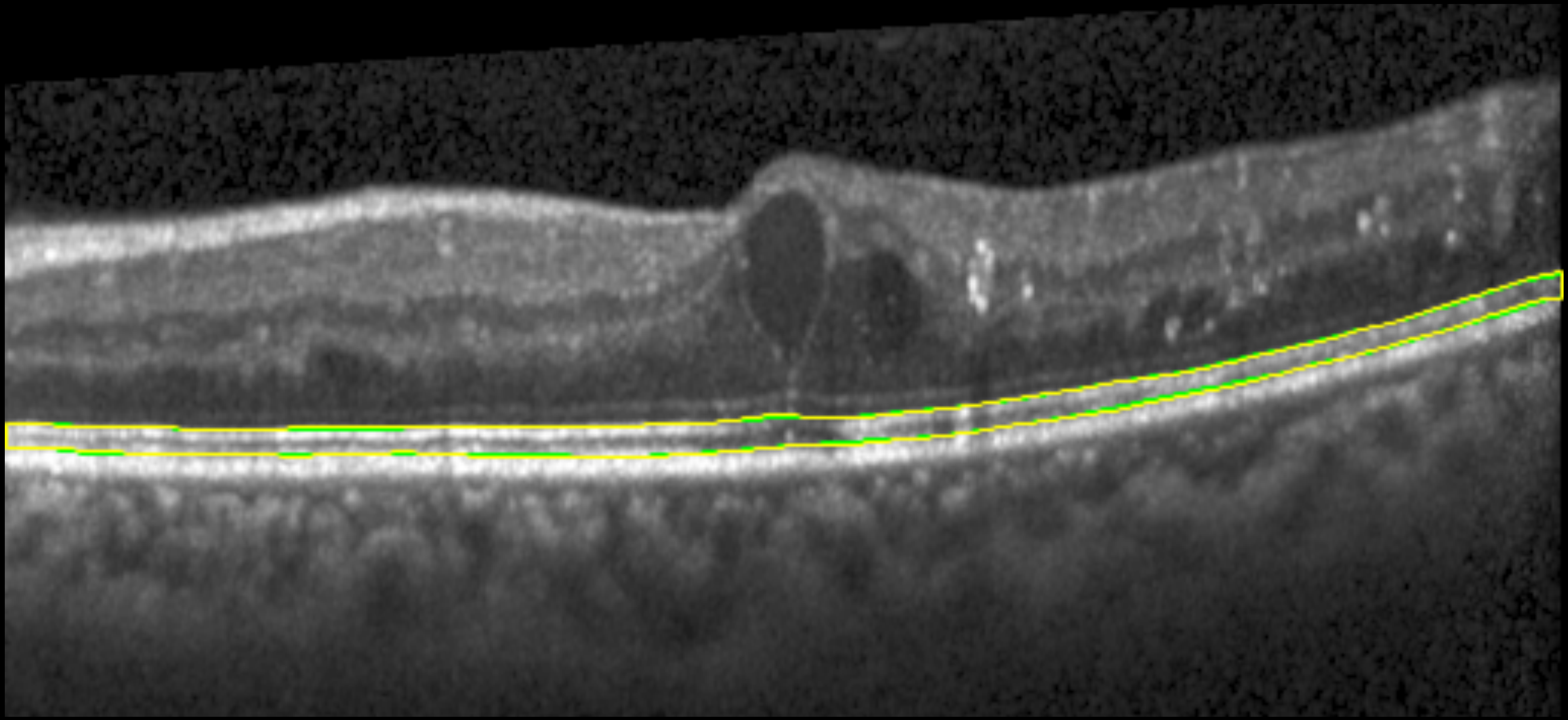


Manual



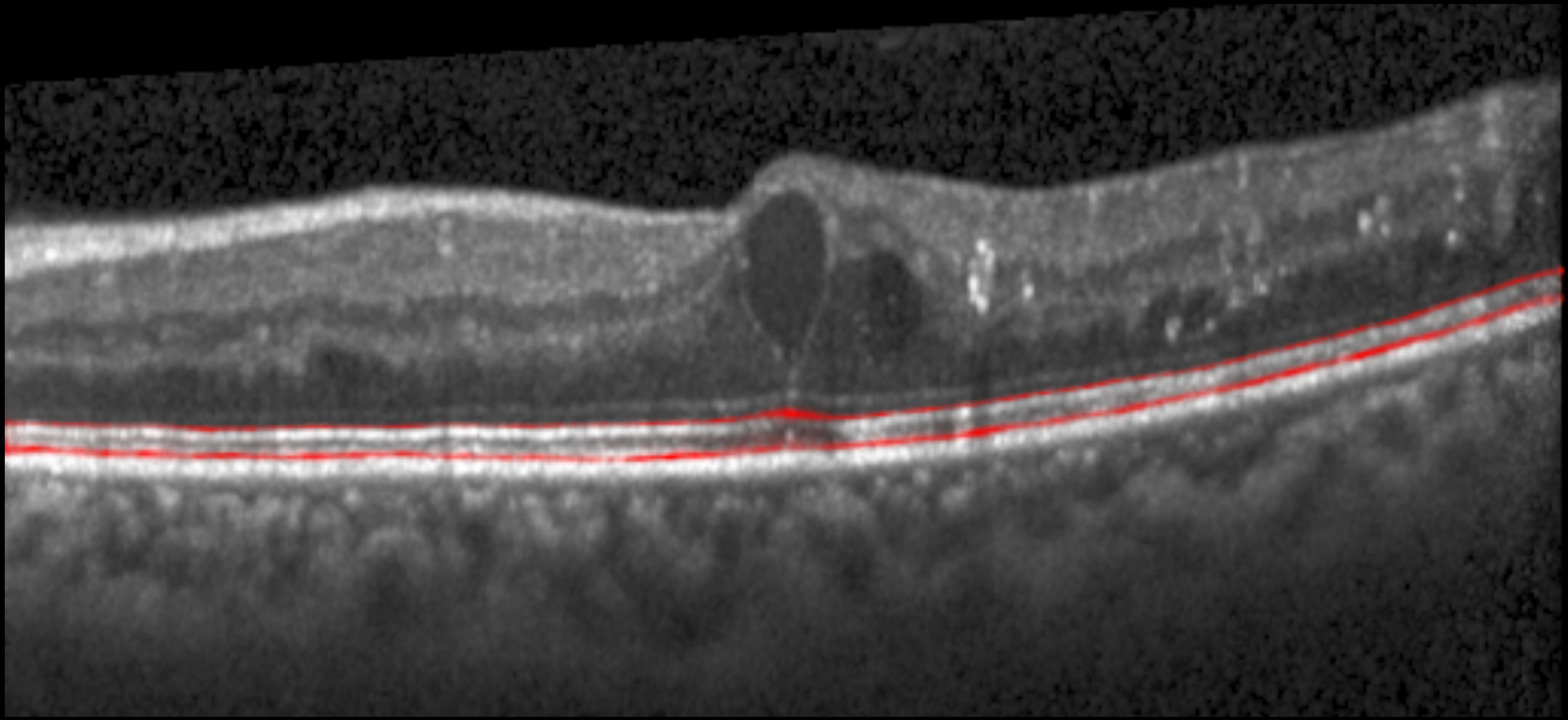
U2-Net

Test set A – Dice= 0.9624 (B-scan level) – Mean uncertainty: 6.004e-4 (B-scan level)



Manual / U2-Net

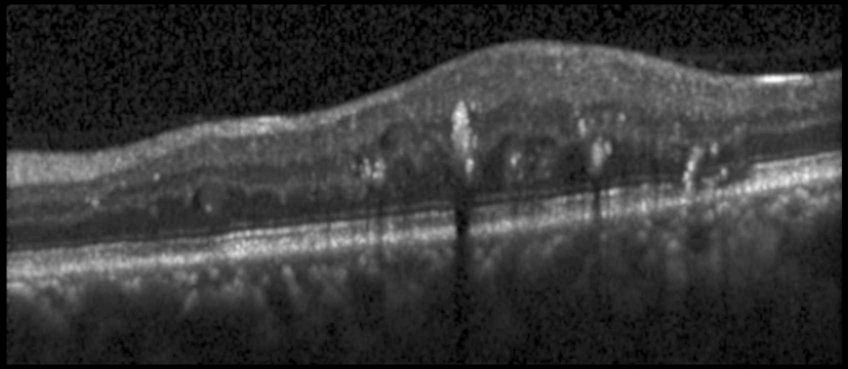
Test set A – Dice= 0.9624 (B-scan level) – Mean uncertainty: 6.004e-4 (B-scan level)

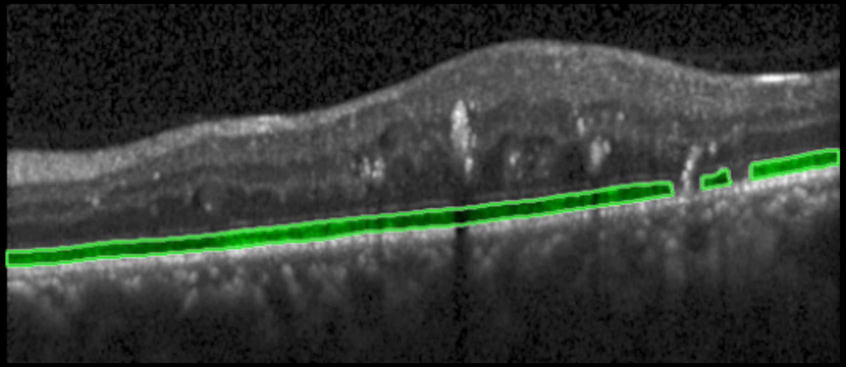


Epistemic uncertainty estimate

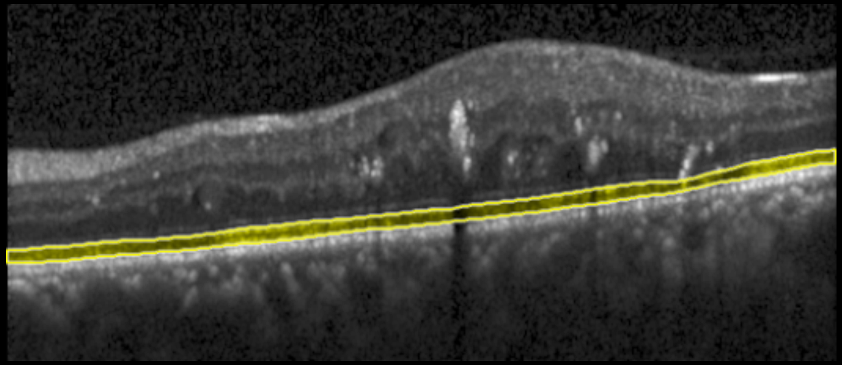
Test set A – Dice= 0.9624 (B-scan level) – Mean uncertainty: 6.004e-4 (B-scan level)

B-scan



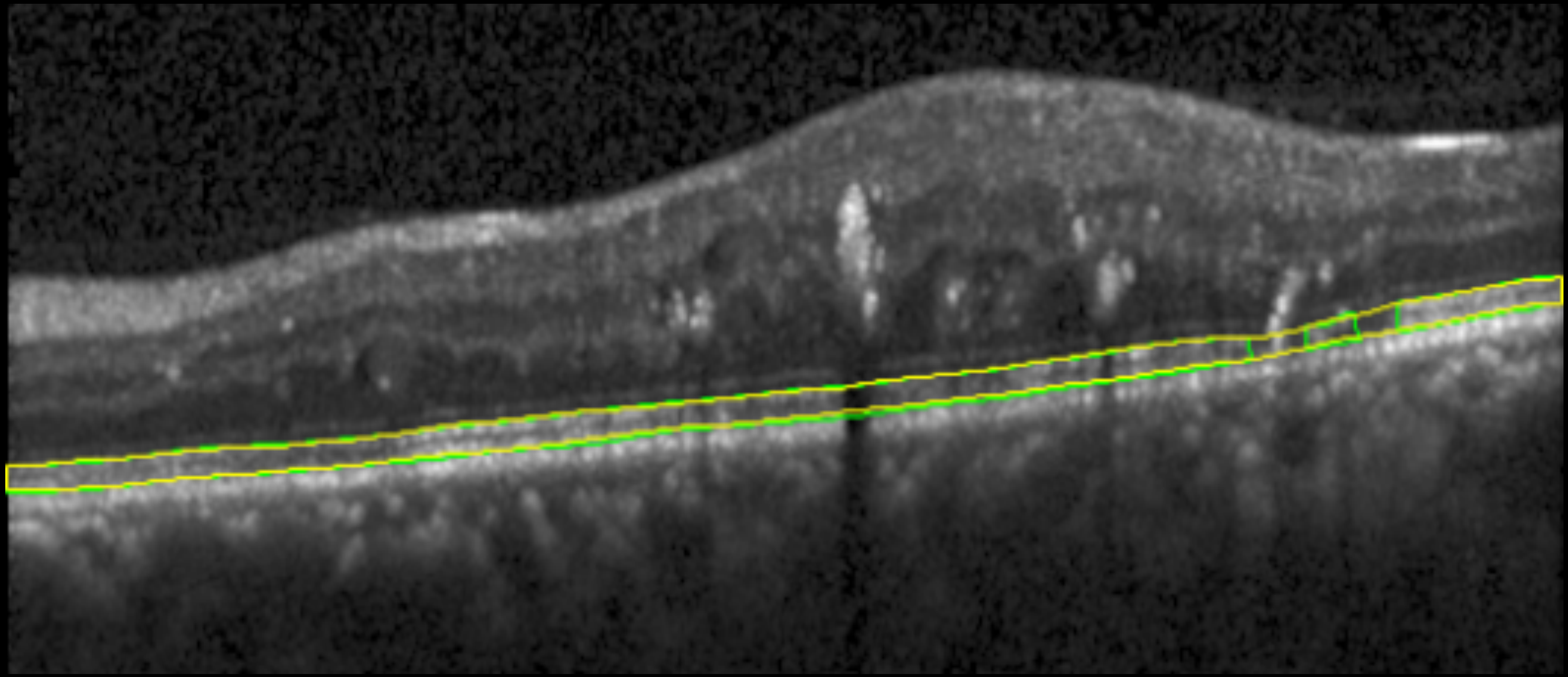


Manual



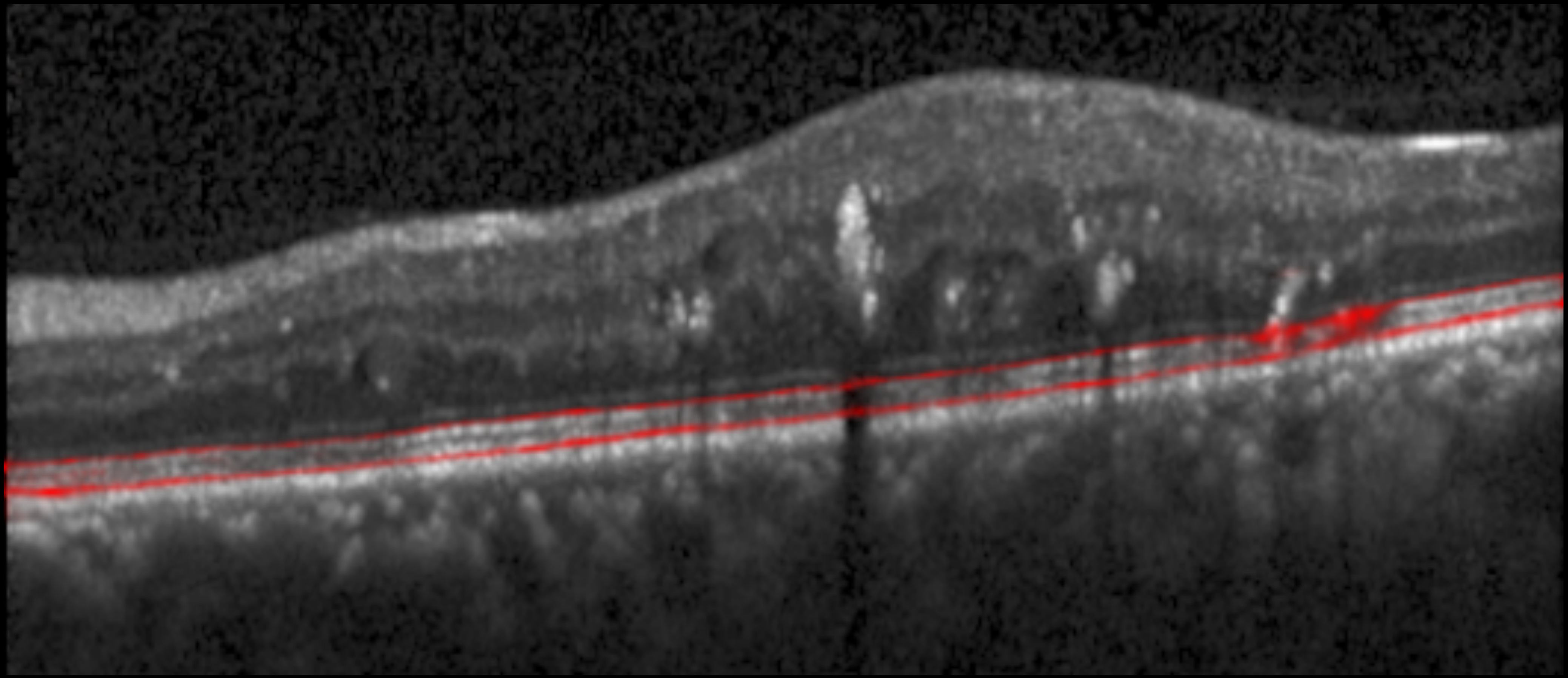
U2-Net

Test set A – Dice= 0.9196 (B-scan level) – Mean uncertainty: 6.720e-4 (B-scan level)



Manual / U2-Net

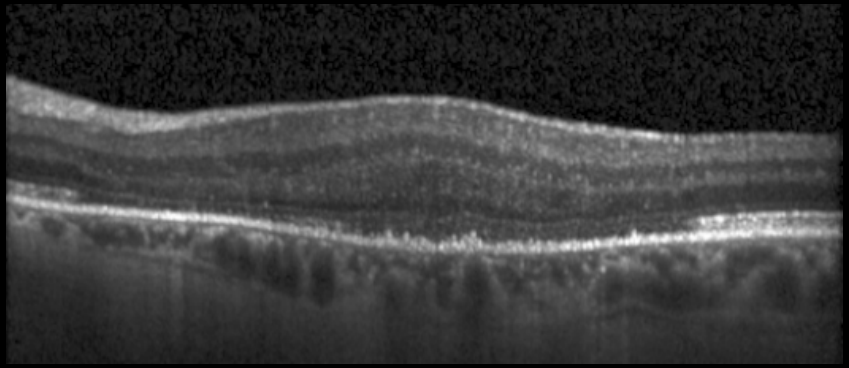
Test set A – Dice= 0.9196 (B-scan level) – Mean uncertainty: 6.720e-4 (B-scan level)

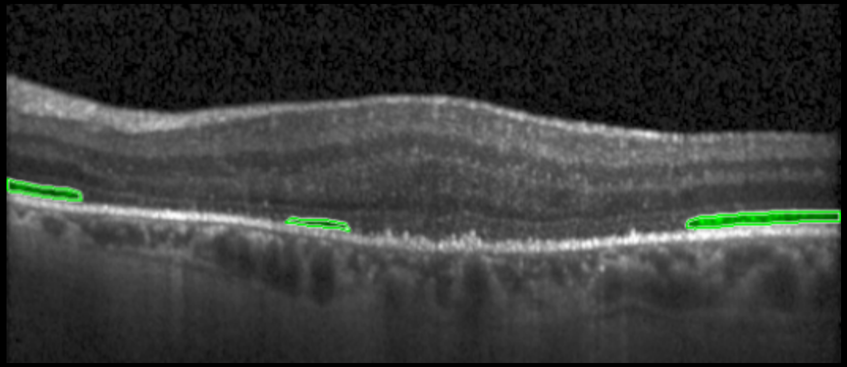


Epistemic uncertainty estimate

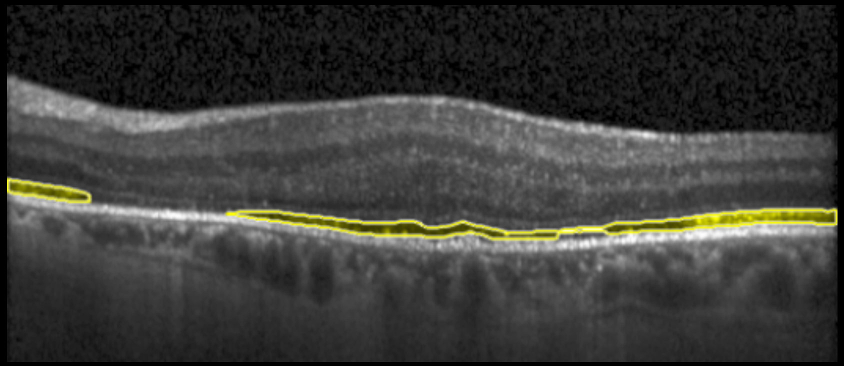
Test set A – Dice= 0.9196 (B-scan level) – Mean uncertainty: 6.720e-4 (B-scan level)

B-scan



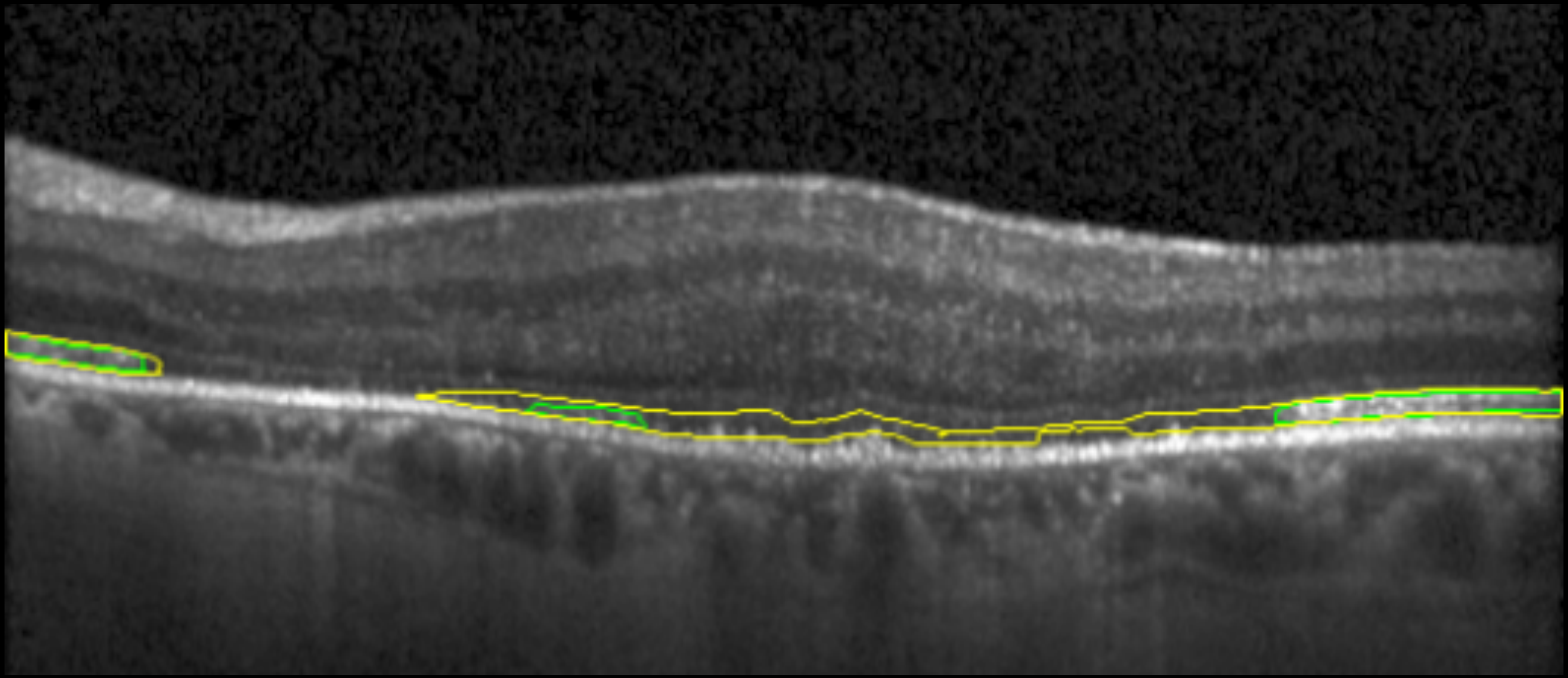


Manual



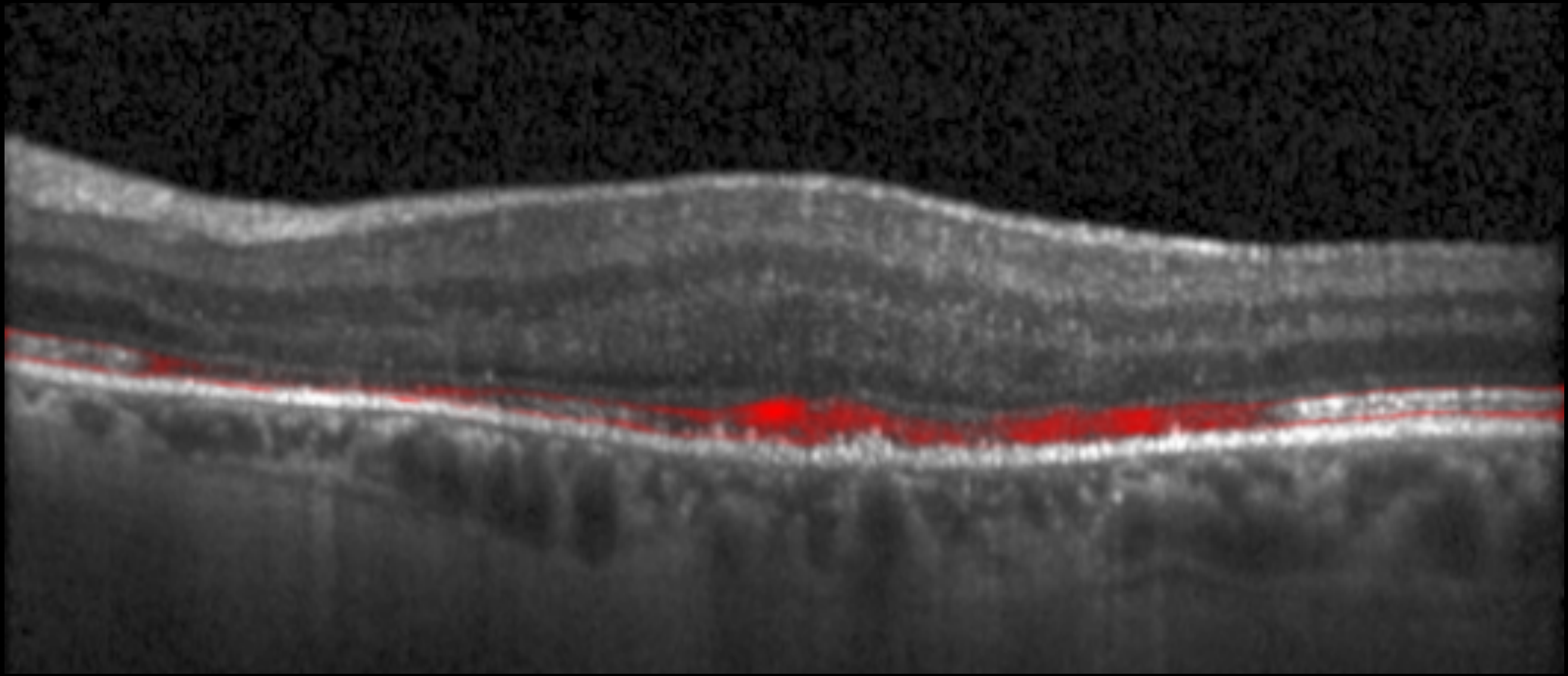
U2-Net

Test set A – Dice= 0.5400 (B-scan level) – Mean uncertainty: 0.0014 (B-scan level)



Manual / U2-Net

Test set A – Dice= 0.5400 (B-scan level) – Mean uncertainty: 0.0014 (B-scan level)

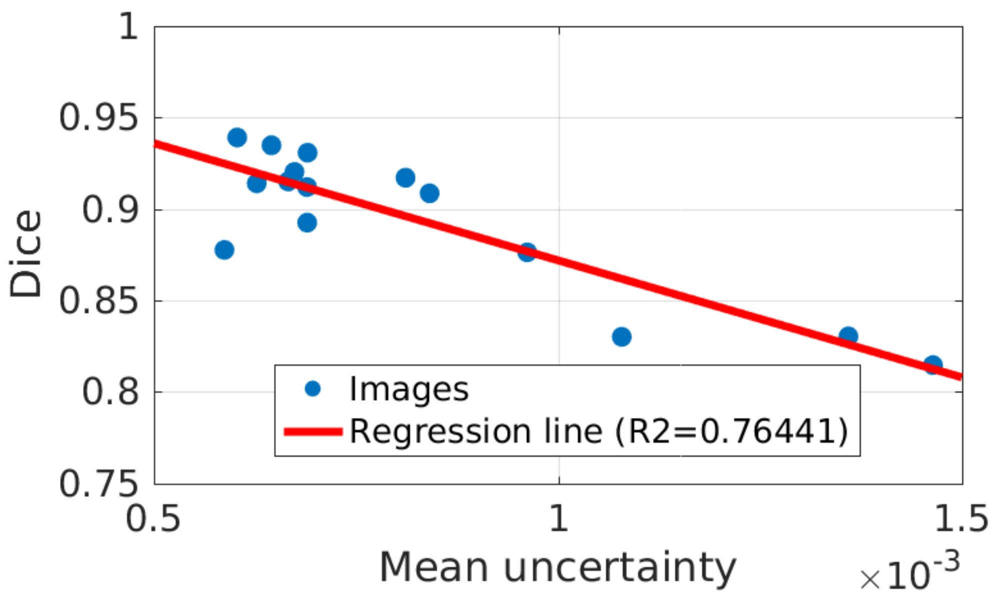


Epistemic uncertainty estimate

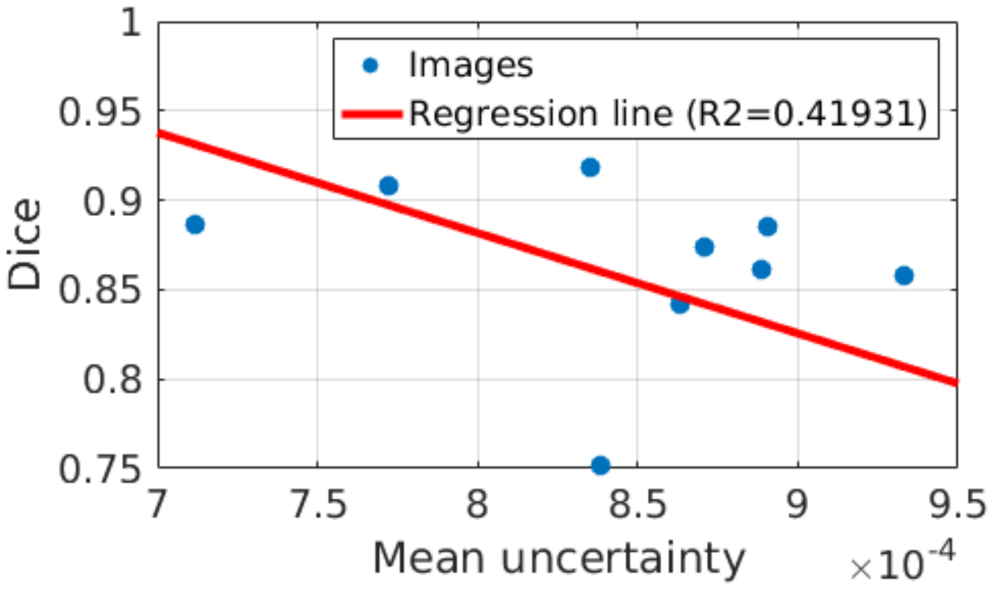
Test set A – Dice= 0.5400 (B-scan level) – Mean uncertainty: 0.0014 (B-scan level)

Uncertainty estimates are inversely correlated with performance

Test set A
(early AMD, CNV, RVO, DME)

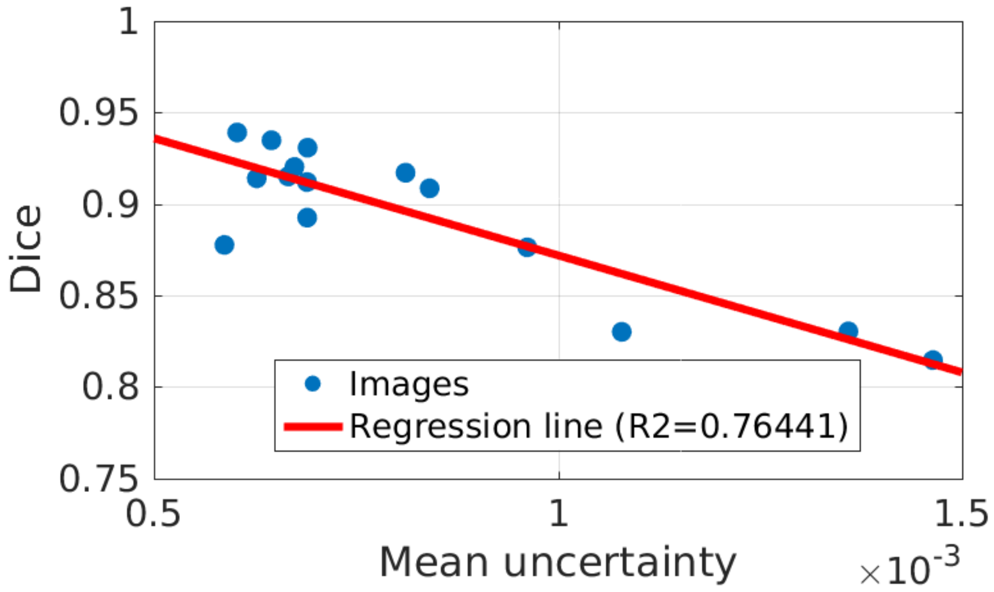


Test set B
(late AMD, GA)



Uncertainty estimates are inversely correlated with performance

Test set A
(early AMD, CNV, RVO, DME)



Test set B
(late AMD, GA)

Conclusions

First deep learning approach for **photoreceptor segmentation** in **pathological OCT scans**

Averaging multiple **MC samples** allows to **increase performance in abnormal areas** without affecting results in healthy regions

Epistemic uncertainty can be used to **assess results' quality** and to **identify areas that might need for manual correction**

Thanks for your attention!

Do you have any questions?

Web https://ignaciorlando.github.io

jose.orlando@meduniwien.ac.at

@ignaciorlando

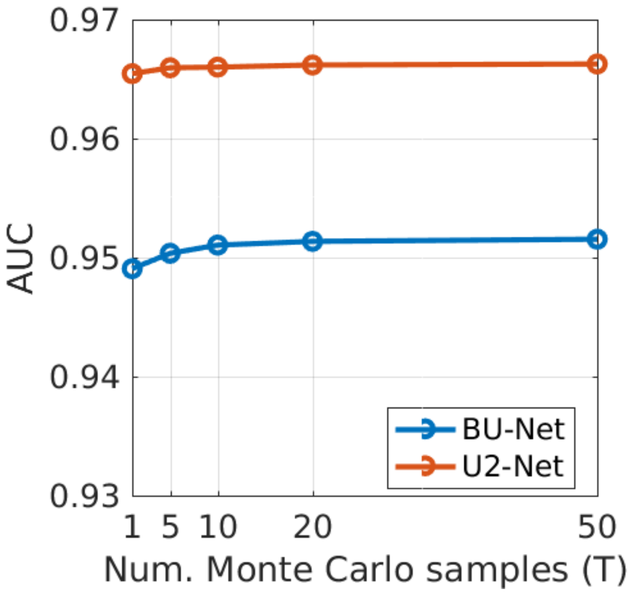
U2-Net:
A Bayesian U-Net Model with Epistemic Uncertainty Feedback for Photoreceptor Layer Segmentation in Pathological OCT Scans

José Ignacio Orlando, Philipp Seeböck, Hrvoje Bogunović, Sophie Klimscha, Christoph Grechenig, Sebastian Waldstein, Bianca S. Gerendas, Ursula Schmidt-Erfurth

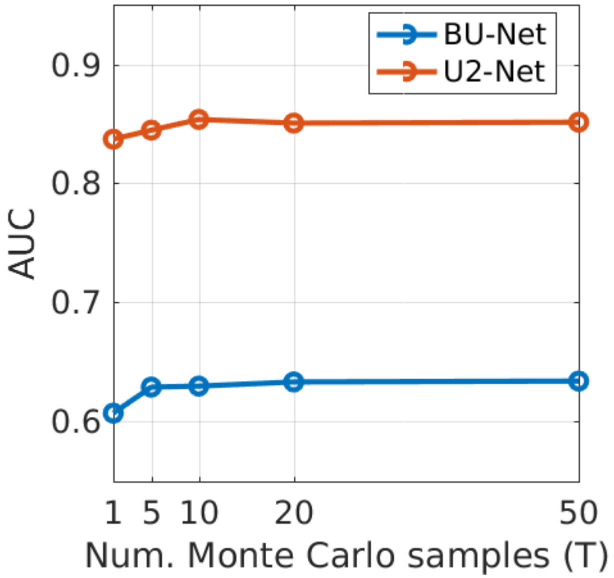
OPTIMA Ophthalmic Image Analysis
Christian Doppler Forschungsgesellschaft
MEDICAL UNIVERSITY OF VIENNA

How many MC samples are necessary?

Validation set A

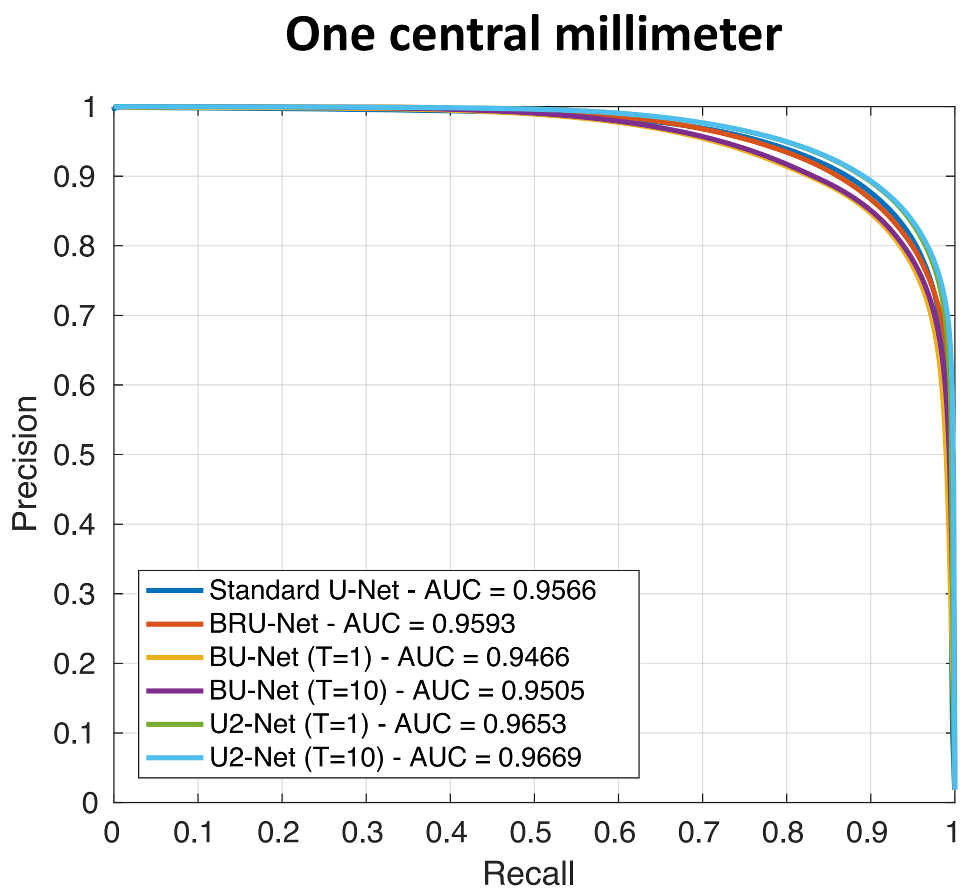


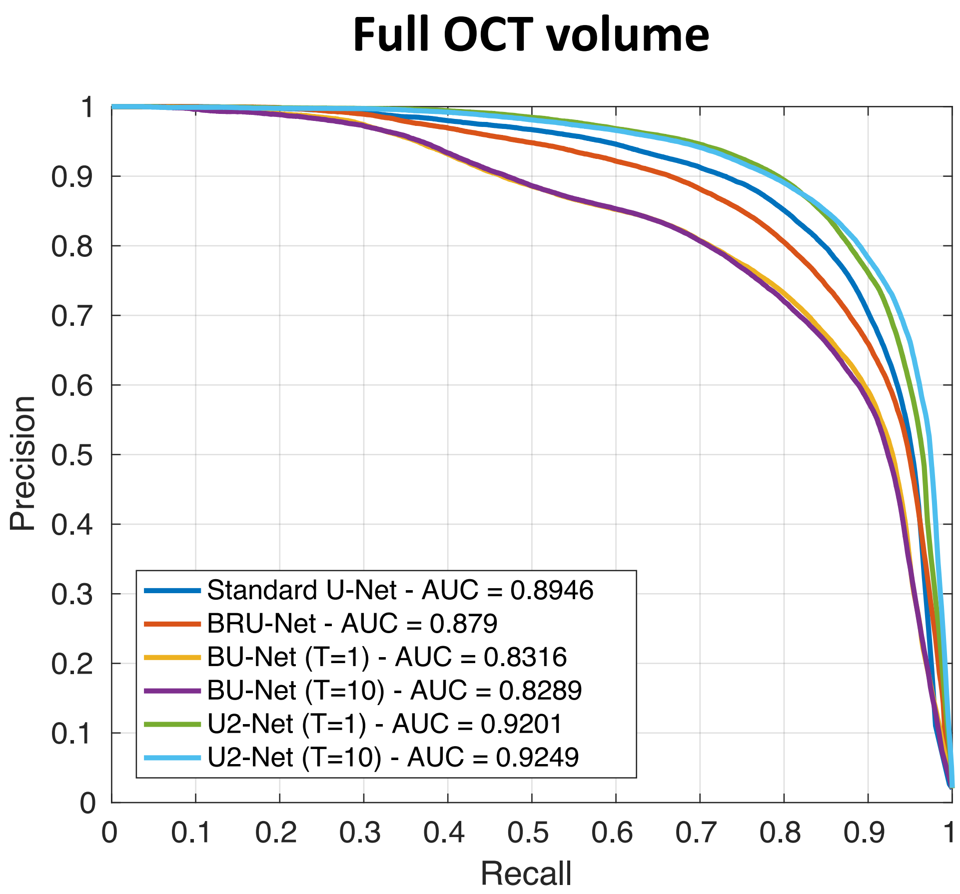
Photoreceptors



Disruptions

Quantitative evaluation

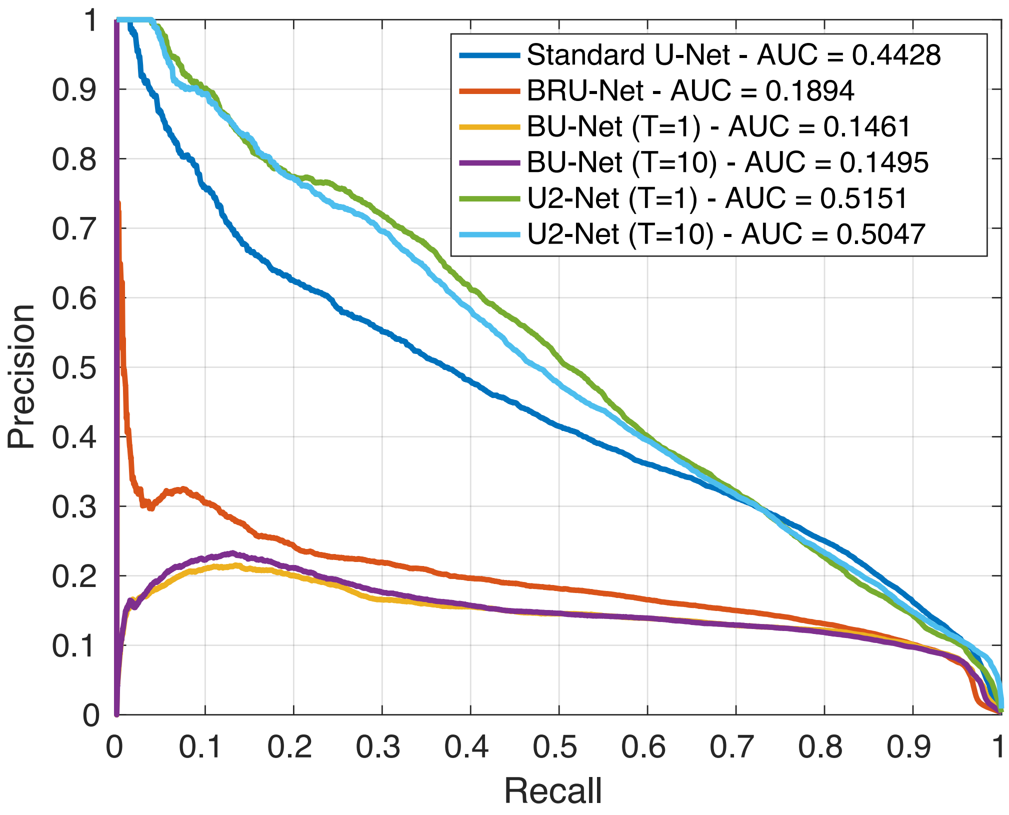




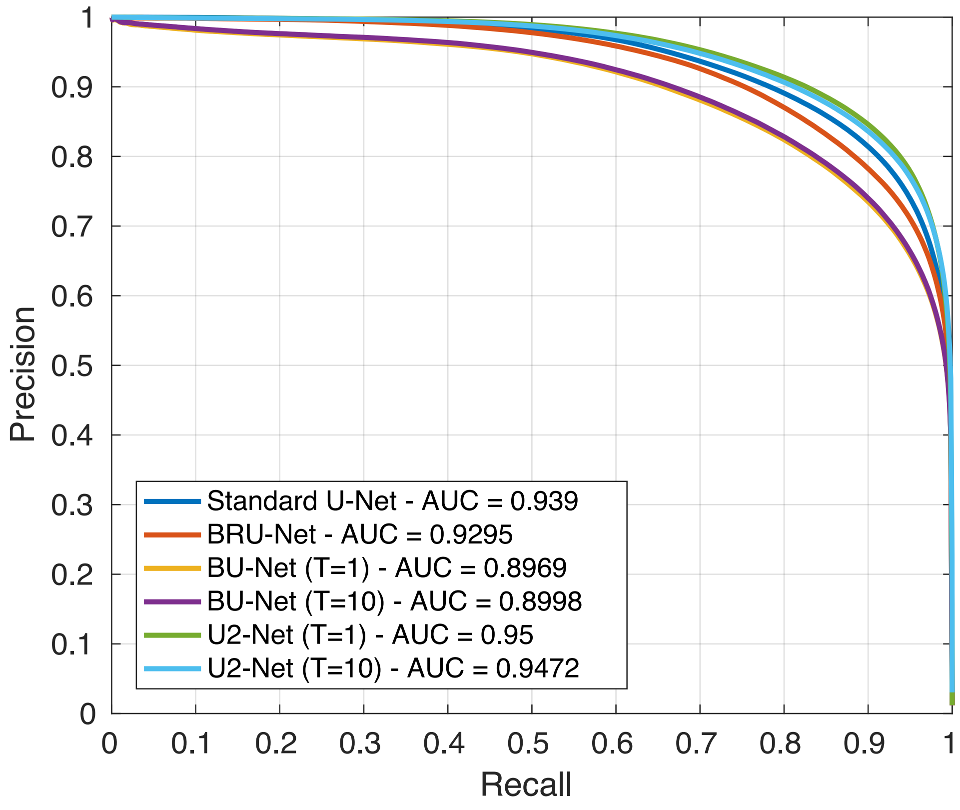
Test set A

Quantitative evaluation

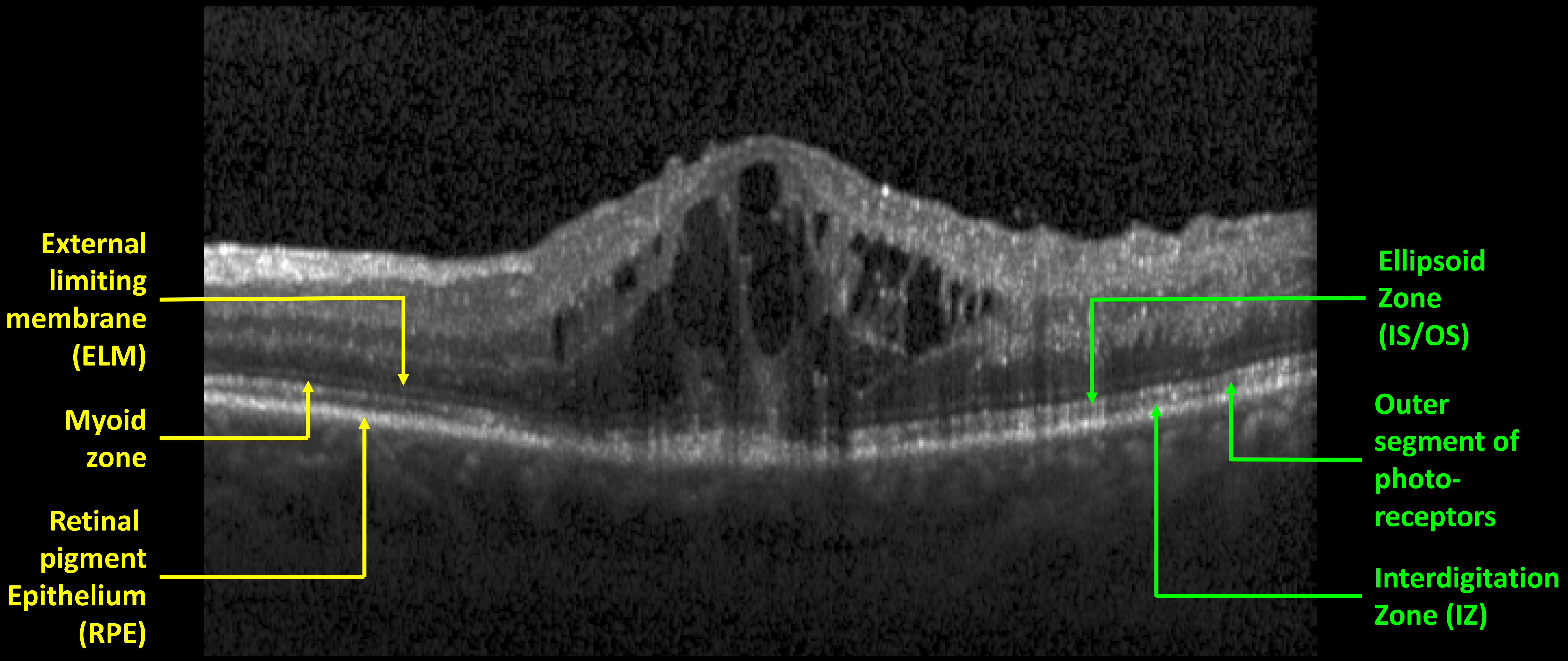
One central millimeter



Full OCT volume



Test set B


External limiting membrane (ELM)
Myoid zone
Retinal pigment Epithelium (RPE)
Ellipsoid Zone (IS/OS)
Outer segment of photo-receptors
Interdigitation Zone (IZ)