Dissertation Defense | Argued by PhD Filipe Soares

17th July 2025 | FEUP

CHARACTERIZATION OF RETINAL FLUID IN OPTICAL COHERENCE TOMOGRAPHY IMAGES

Master in Biomedical Engineering

Authored by David Castanho Terroso

Supervised by Prof. Tânia Melo & Co-supervised by Prof. Ana Maria Mendonça

SUMMARY

01 PROBLEM

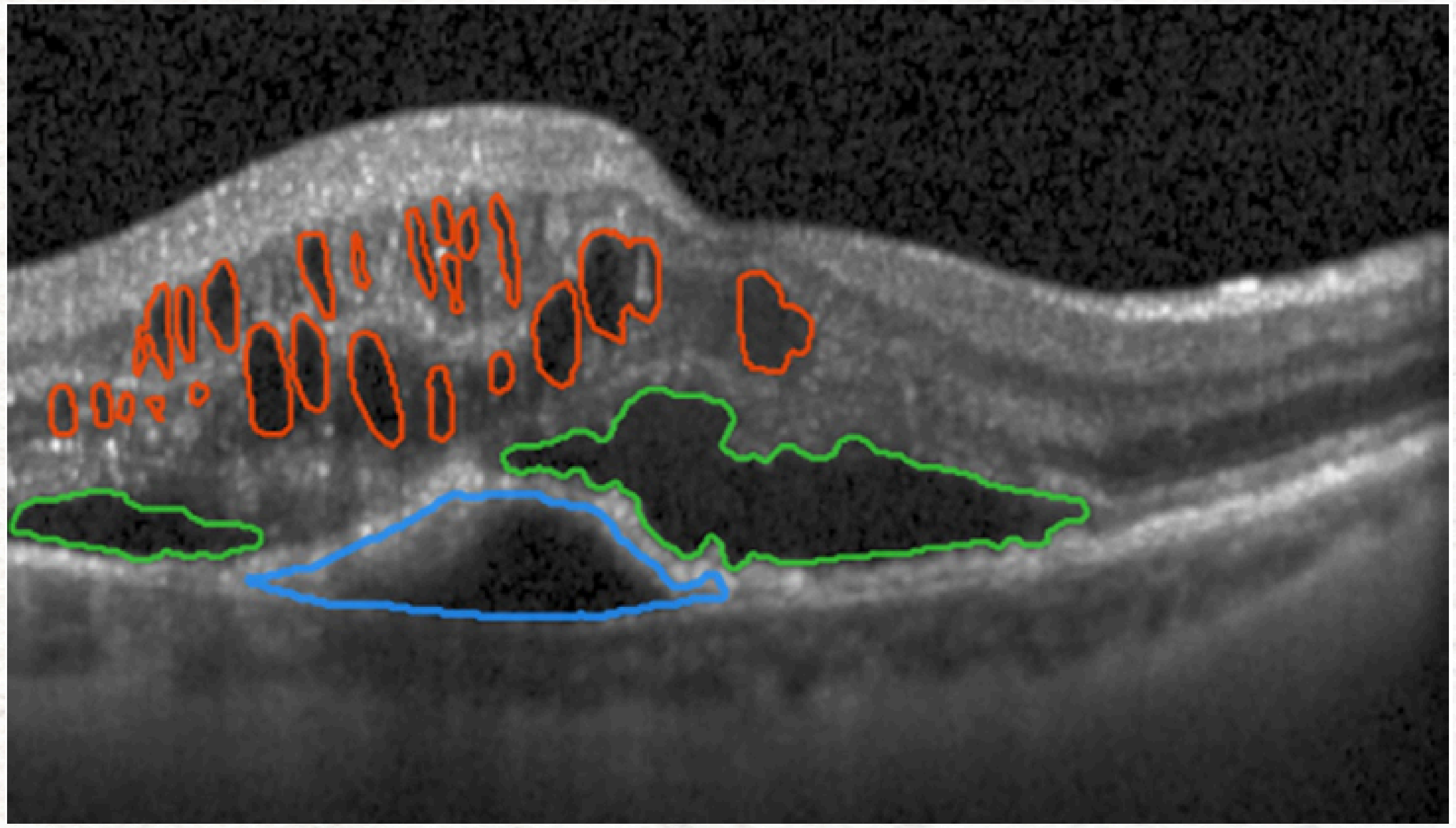
Some Leading Causes of Blindness

- Age-related Macular Degeneration (AMD)
- Diabetic Macular Edema (DME)
- Macular Edema Secondary to Retinal Vein Occlusion (RVO)

Characterized by leaking **fluid** onto the retina

Which can be classified in:

- Intraretinal Fluid (IRF)
- Subretinal Fluid (SRF)
- Pigmental Epithelium Detatchment (PED)



IRF in red, SRF in green, and PED in blue [1]

Characterization of fluid:

- Important biomarker;
- Laborious task;
- Conditioned by medical expertise and OCT quality and resolution.

[1] H. Bogunović et al., *Chapter 14 - OCT fluid detection and quantification*, pp. 273–298. Academic Press, 2019.

02 DISSERTATION OBJECTIVES

- Analyze retinal OCT scans, classifying the retinal fluids in three different types (IRF, SRF, and PED) and quantifying their respective volumes.
- Increase the inter-slice resolution of the OCT volumes, with the aim of improving the fluid volume estimation.

1	Develop different 2D deep learning models for the segmentation of different retinal fluid types (IRF, SRF, and PED) in OCT volumes.
2	Evaluate the performance of the best segmentation model in different datasets and estimate the volume of each fluid using its predicted masks.
3	Use a generative model for synthesizing intermediate slices in OCT volumes, in order to improve their inter-slice resolution of the volume, and assessing the quality of the super-resolved volume.
4	Investigate the impact of intermediate slice synthesis on the fluid volume estimation.

03 LITERATURE REVIEW

03.A FLUID SEGMENTATION

Search Query (on Title or Abstract):

"OCT" AND "segmentation" AND ("deep learning" OR "CNN" OR "neural network")

Databases
- PubMed: **398**
- IEEE: **105**
- ScienceDirect: **125**
- ACM: **80**

Excluded: 688
- Duplicates: **133**
- Abstract Screening: **499**
- Full-text Screening: **56**

Criteria:

✗ Before 2019;
✗ Without full-text available in English;
✗ Books, thesis, conference, or review articles.

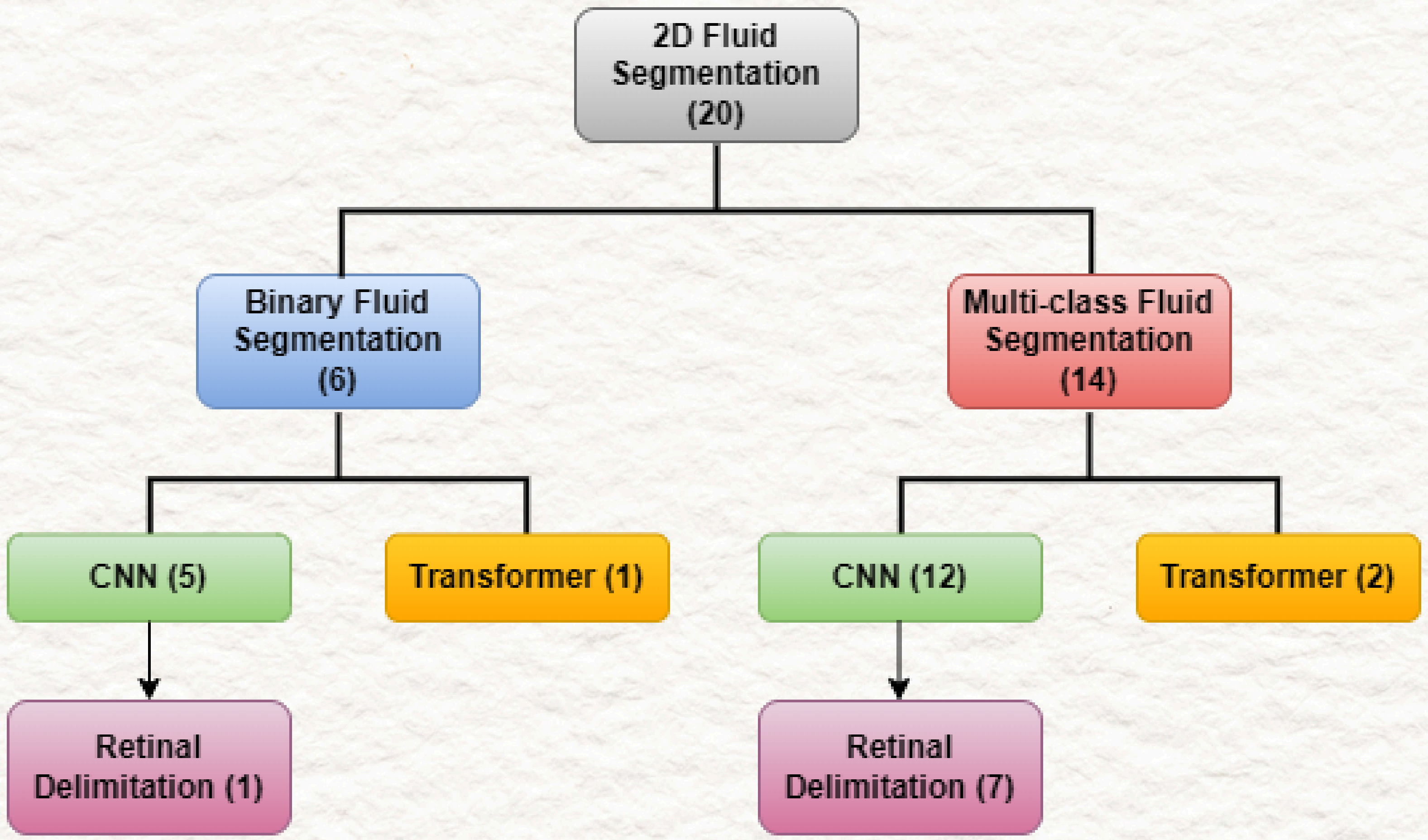
Included: 20

Criteria:

✓ Articles focused on 2D or 2.5D segmentation in OCT volumes.

03 LITERATURE REVIEW

03.A FLUID SEGMENTATION

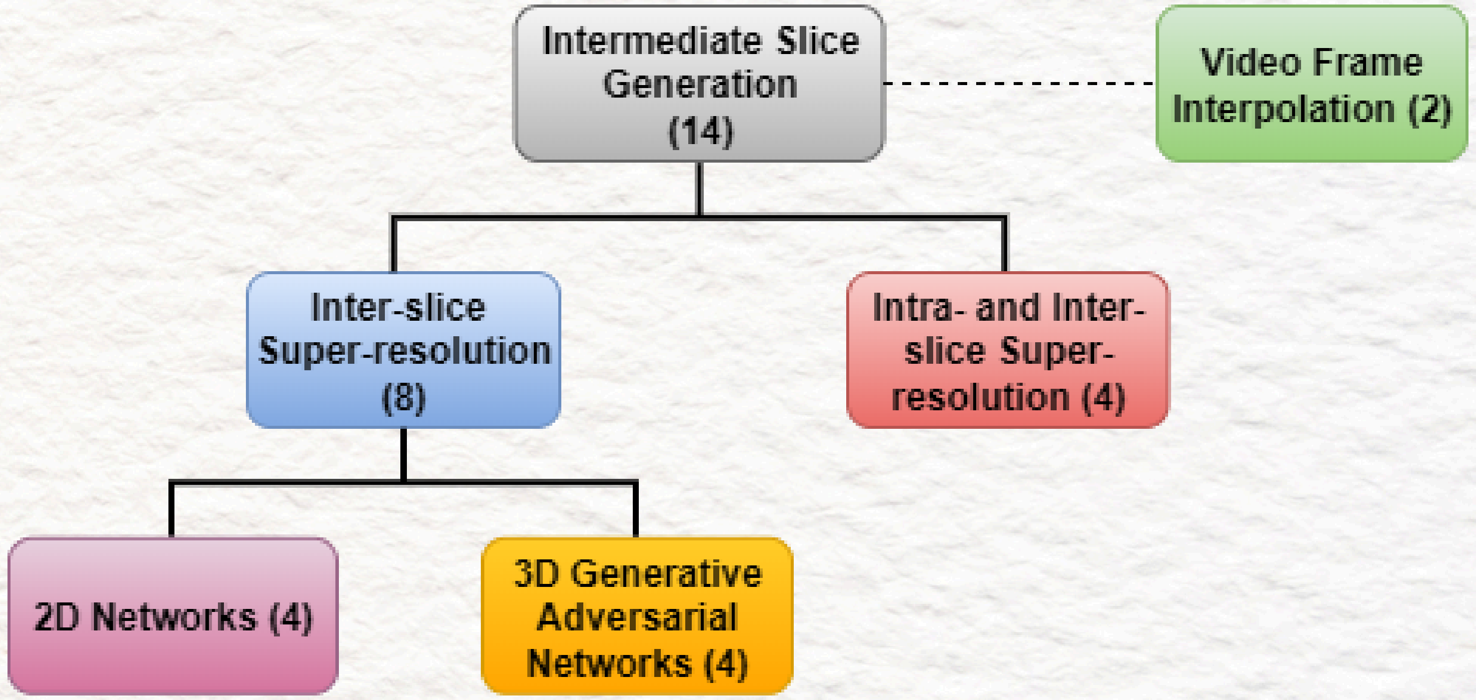


- Convolutional Neural Networks (CNNs)
  - Inspired by the U-Net;
  - Main approach to the segmentation problem;
  - Implementations differ on the modules, blocks, or type of convolutions used.
- Transformers
  - Combined with CNNs, replacing the encoding path or the bottleneck part of the network;
  - All Transformers work alongside with CNNs.
- Retinal Delimitation
  - Common practice to discard irrelevant information;
  - Easier training and more anatomically coherent results.

03 LITERATURE REVIEW

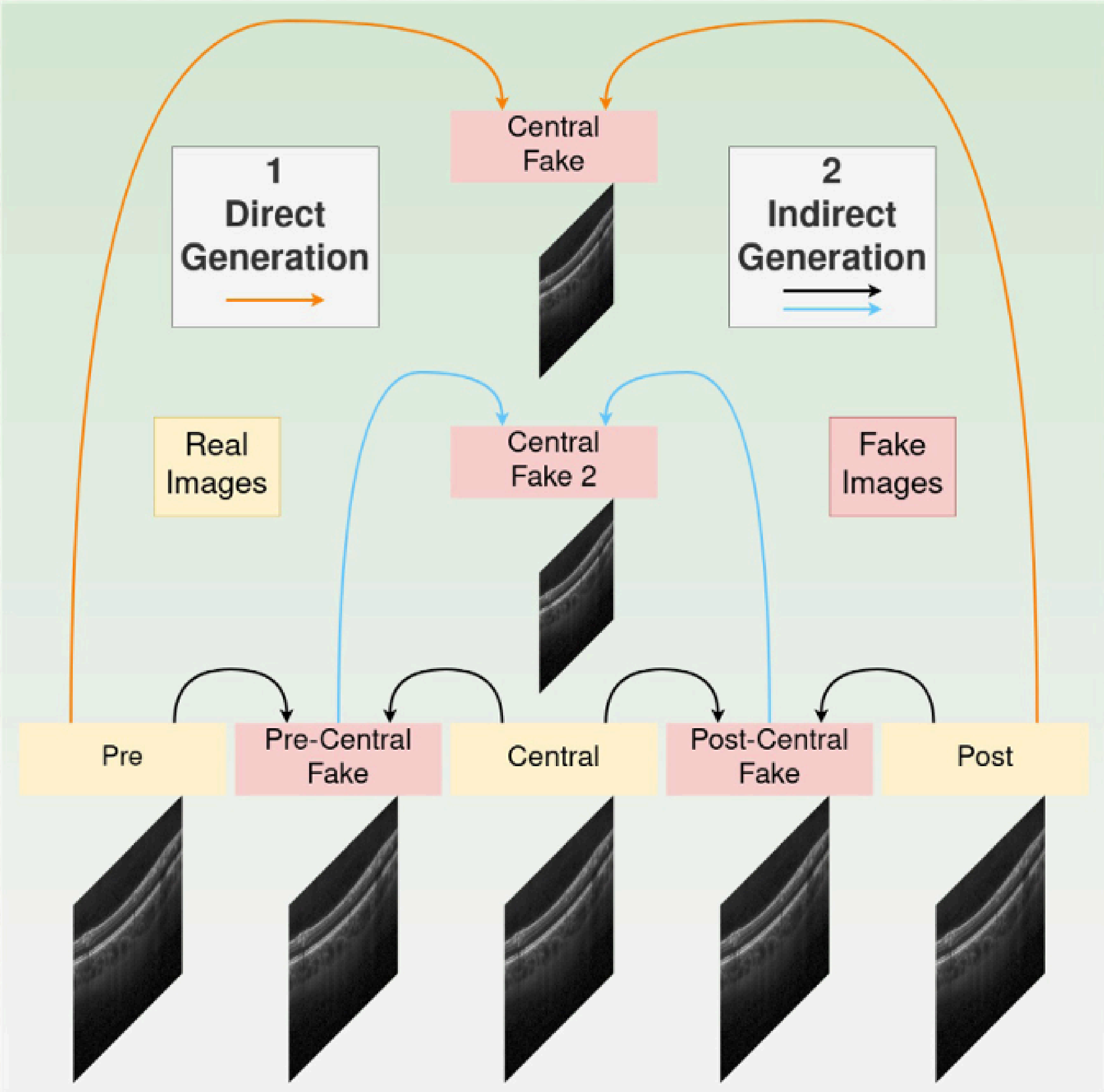
03.B INTERMEDIATE SLICE SYNTHESIS

- Inter-slice Super-resolution with 2D Networks
  - Considers two consecutive slices and learns to generate one intermediate slice;
- Intra- and Inter-slice Super-resolution
  - Usually composed by three networks: two intra-slice, one inter-slice;
  - Intra-slice super-resolution is performed on the orthogonal 2D views;
  - Inter-slice super-resolution is applied simultaneously and mutual supervision between the three axes occurs.
- 3D Generative Adversarial Networks
  - Considers the whole low-resolution volume and outputs a high-resolution one;
  - Computationally expensive.
- Video Frame Interpolation
  - Can be used as inspiration for intermediate slice generation in medical images.

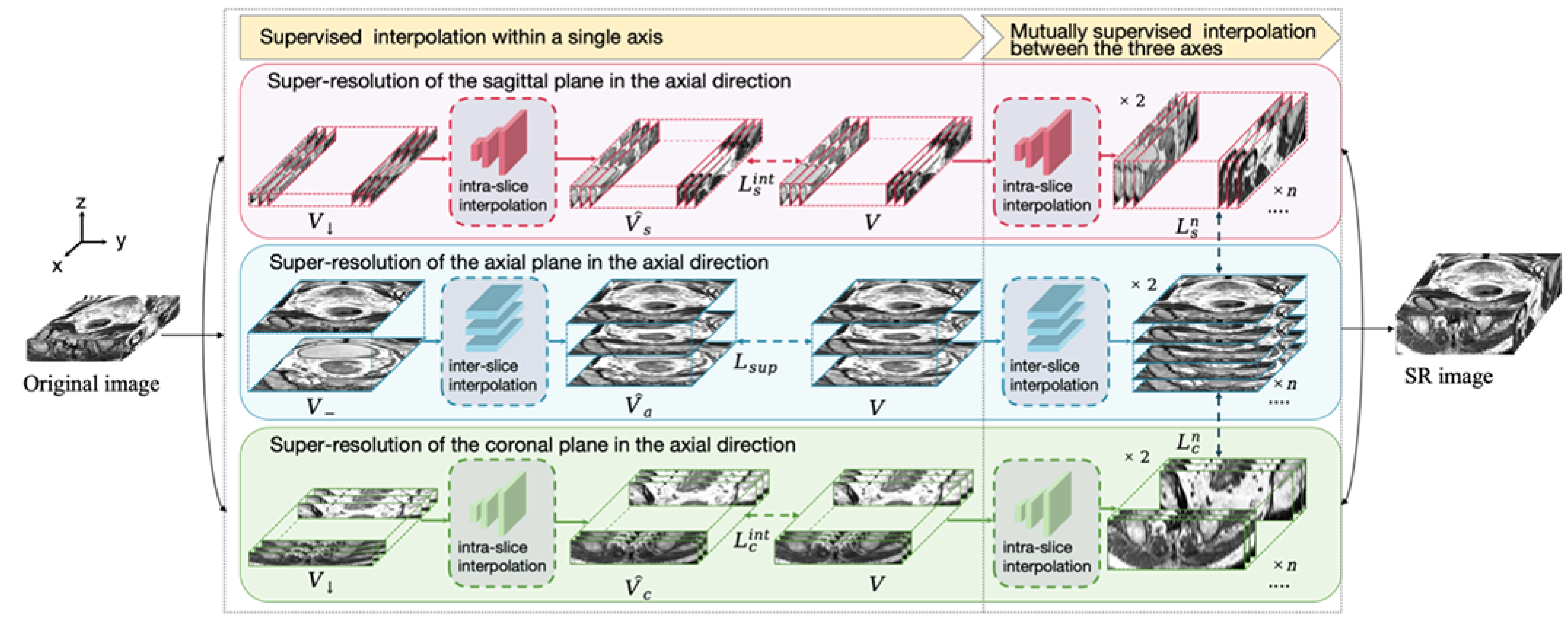


03 LITERATURE REVIEW

03.B INTERMEDIATE SLICE SYNTHESIS



[2] E. López-Varela et al., "Generation of synthetic intermediate slices in 3D OCT cubes for improving pathology detection and monitoring" *Computers in Biology and Medicine*, vol. 163, p. 107214, 2023.



[3] H. Zhang et al., "A novel GAN-based three-axis mutually supervised super-resolution reconstruction method for rectal cancer MR image" *Computer Methods and Programs in Biomedicine*, vol. 257, p. 108426, 2024.

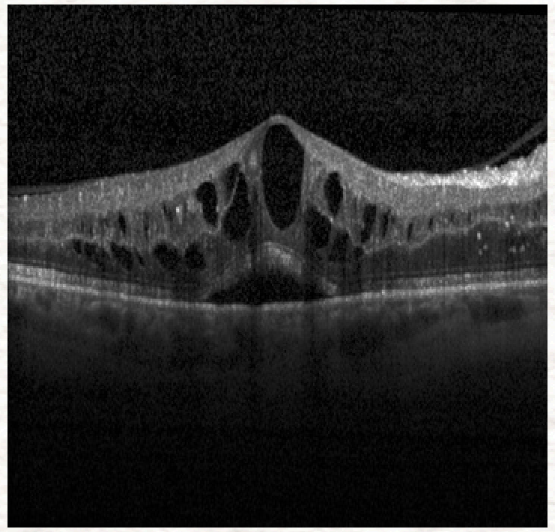
04 MATERIALS AND METHODS

04.A DATASETS

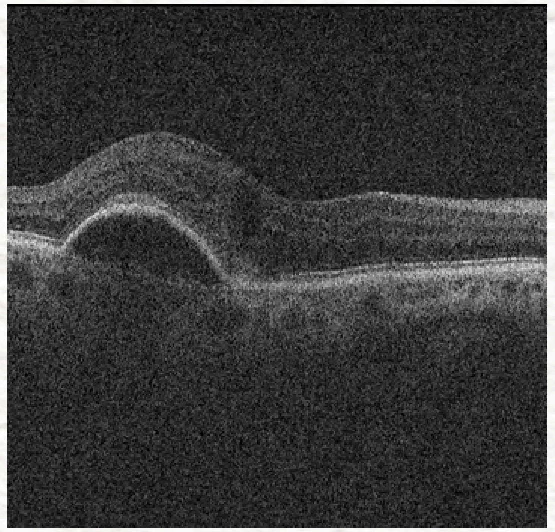
RETOUCH [4]:

- One of the most popular datasets for retinal fluid segmentation in OCT exams;
- Large dataset with 112 volumes and 11334 B-scans;
- Volumes from three different devices (Spectralis, Cirrus, and Topcon);
- Constant volume size for the same vendor.

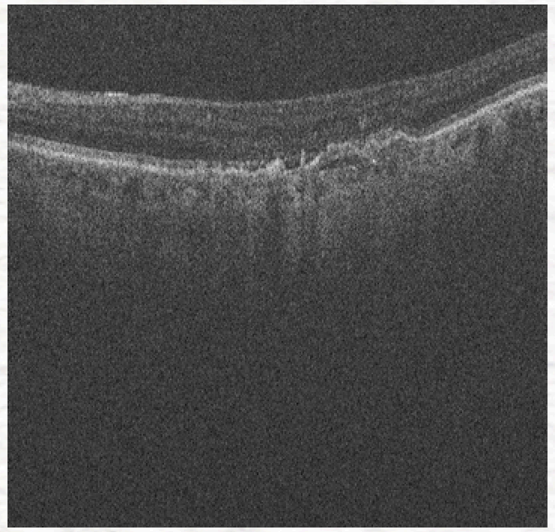
Provides QUANTITY and VARIETY



Spectralis



Cirrus



Topcon

[4] H. Bogunović et al., “RETOUCH: The Retinal OCT Fluid Detection and Segmentation Benchmark and Challenge” *IEEE Transactions on Medical Imaging*, vol. 38, pp. 1858–1874, 2019.

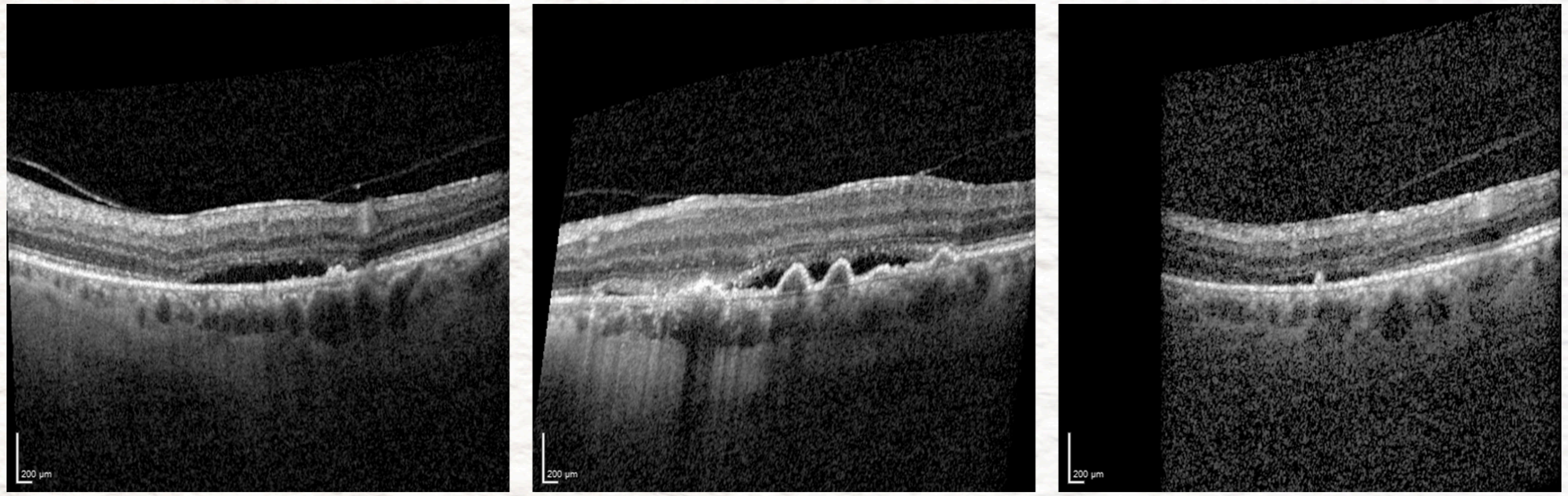
04 MATERIALS AND METHODS

04.A DATASETS

Centro Hospitalar Universitário de São João (CHUSJ):

- Private OCT dataset, with anotated fluids;
- Small dataset, with 6 volumes and 114 B-scans;
- Volumes from Spectralis;
- Varying image sizes, noise levels, and visual characteristics across volumes.

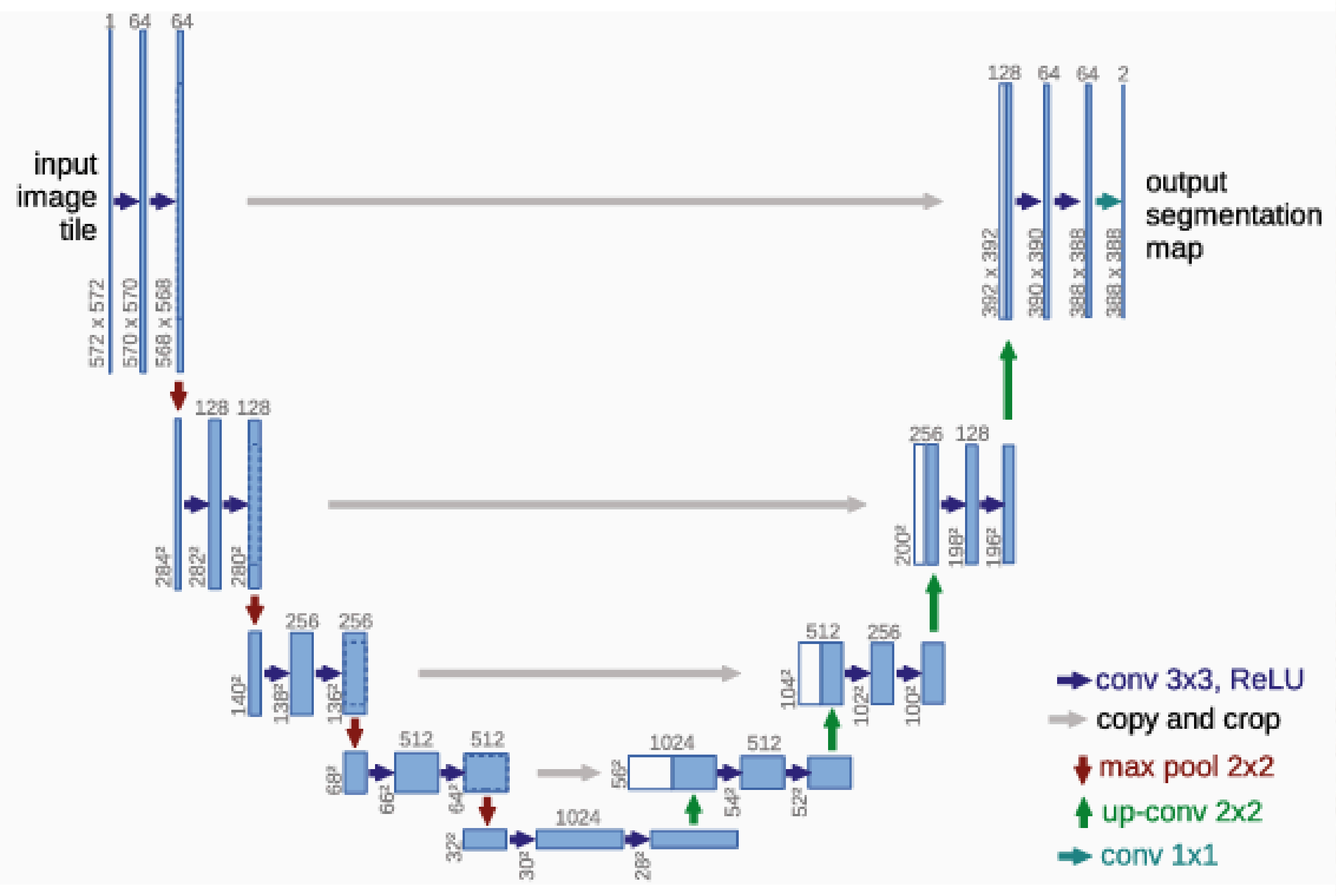
Allows validation on EXTERNAL data



04 MATERIALS AND METHODS

04.B FLUID SEGMENTATION - NETWORK AND LOSS FUNCTION

Selected Network: U-Net [5]



Loss Function [6]

$$\mathcal{L} = \lambda_D \mathcal{L}_D + \lambda_{CE} \mathcal{L}_{CE}$$

0.5

0.5

Dice Loss

Balanced Cross-Entropy Loss

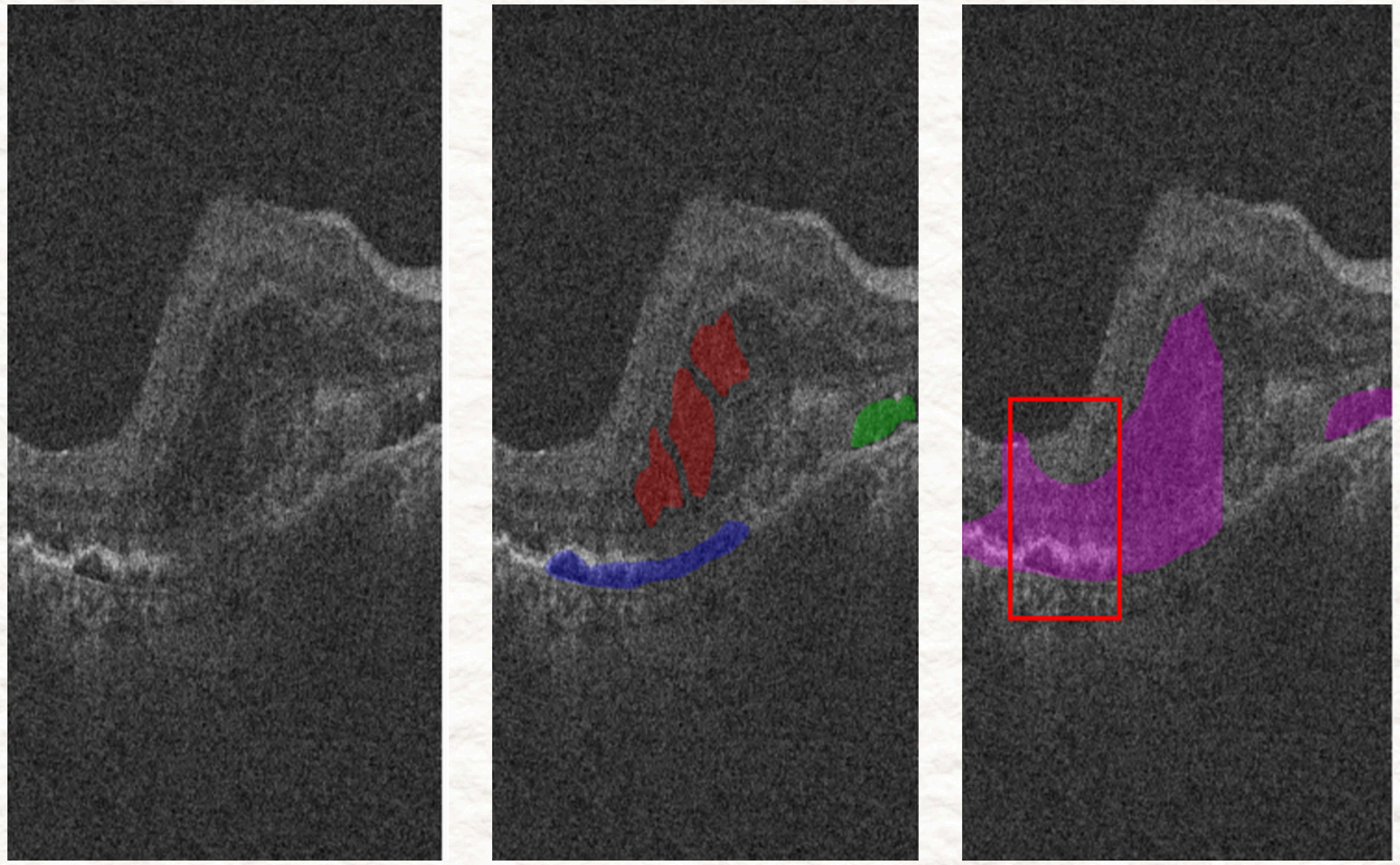
[5] O. Ronneberger, et al., “U-Net: Convolutional Net works for Biomedical Image Segmentation” in *Medical Image Computing and Computer-Assisted Intervention– MICCAI 2015* (N. Navab, J. Hornegger, W. M. Wells, and A. F. Frangi, eds.), pp. 234–241, Springer International Publishing, 2015.

[6] R. Tennakoon et al., “Retinal fluid segmentation in OCT images using adversarial loss based convolutional neural networks” in *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, 2018, pp. 1436–1440. DOI: 10.1109/ISBI.2018.8363842.

04 MATERIALS AND METHODS

04.B FLUID SEGMENTATION - PATCH EXTRACTION

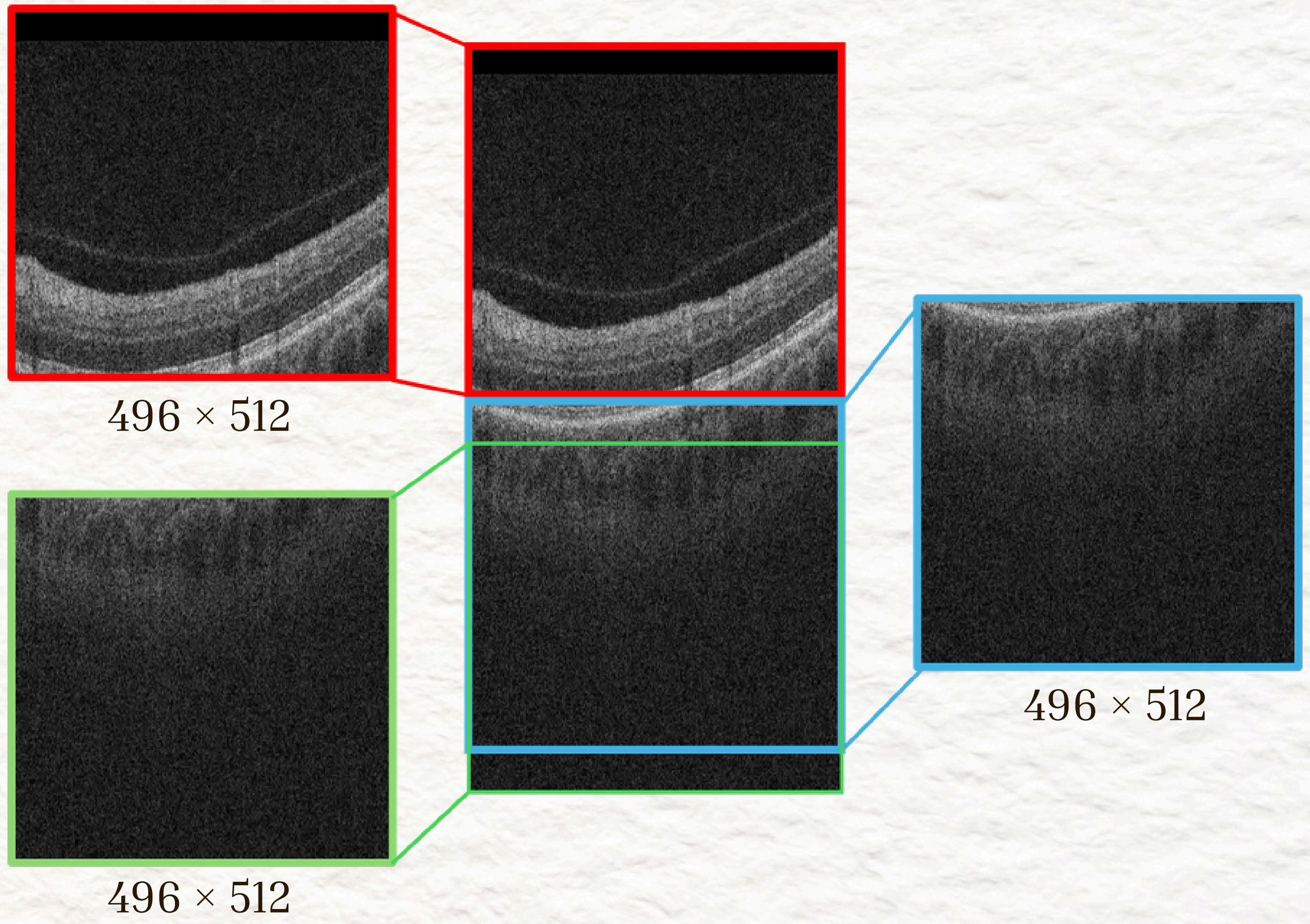
A. Small Patch Extraction



IRF in red, SRF in green, PED in blue, and the region of interest in purple (right).
The red bounding box represents a possible patch.

- 256 x 128 (H x W) patches extracted from the region of interest (ROI);
- The ROI consists of the regions with highest entropy or with fluid;
- 10 patches are **randomly** extracted from this region in each B-scan.

B. Large Patch Extraction

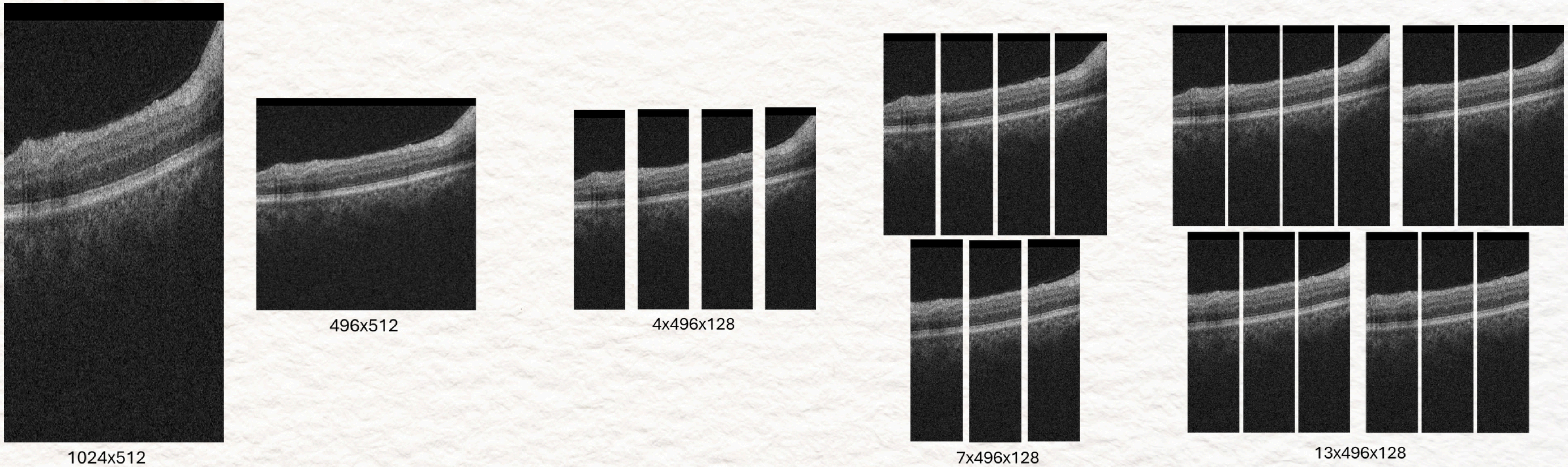


- 496 x 512 patches extracted from each B-scan;
- The patches were extracted from **top to bottom** so that every section of the image was present in at least one patch.

04 MATERIALS AND METHODS

04.B FLUID SEGMENTATION - PATCH EXTRACTION

C. Vertical Patch Extraction



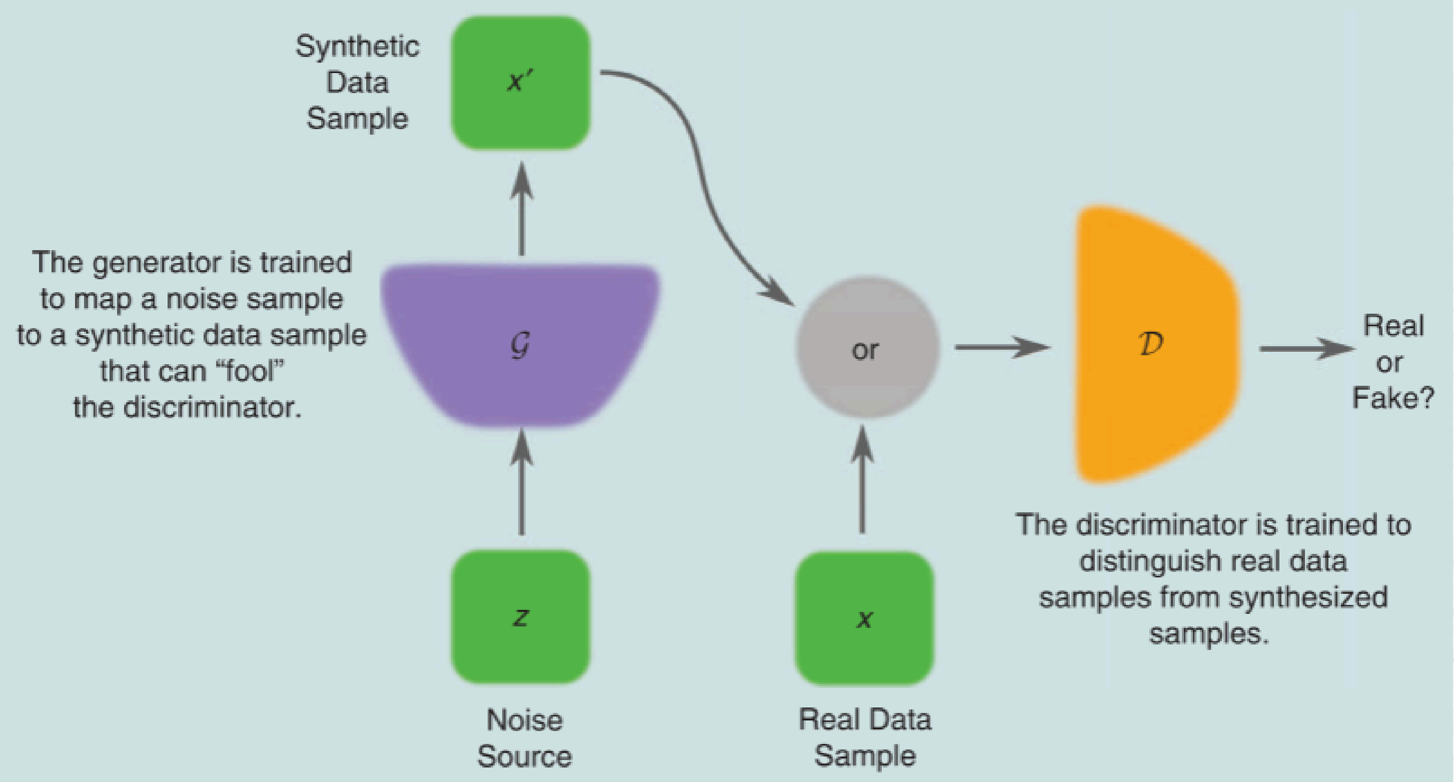
- All B-scans were **reshaped** to 496 x 512;
- Four, seven, or thirteen patches of shape 496 x 128 were extracted from each B-scan;
- The **four** extracted patches were **disjoint**, while the **seven** and **thirteen** patches were **sampled at regular intervals** along the horizontal axis, starting from the left of the image and ending at the last position where a full-width patch could be extracted.

04 MATERIALS AND METHODS

04.B INTERMEDIATE SLICE GENERATION - NETWORKS

Generative Adversarial Network:

- Inspired by the work of Tran and Yang [7];
- The generator receives two consecutive slices and then generates the intermediate slice, while the discriminator must correctly distinguish the real slices from the generated ones;
- As the generator gets better, so does the discriminator.



D is the discriminator and G is the generator [8].

U-Net:

- A **generative U-Net** was also tested for the generatation of an intermediate B-scan, motitvated by the work of Nishimoto et al. [9];
- The training is performed using the entire B-scans and regulated using the MAE.

[7] Q. N. Tran and S.-H. Yang, "Efficient Video Frame Interpolation Using Generative Adversarial Networks" *Applied Sciences*, vol. 10, 18 2020, ISSN: 2076-3417. DOI: 10.3390/app10186245.

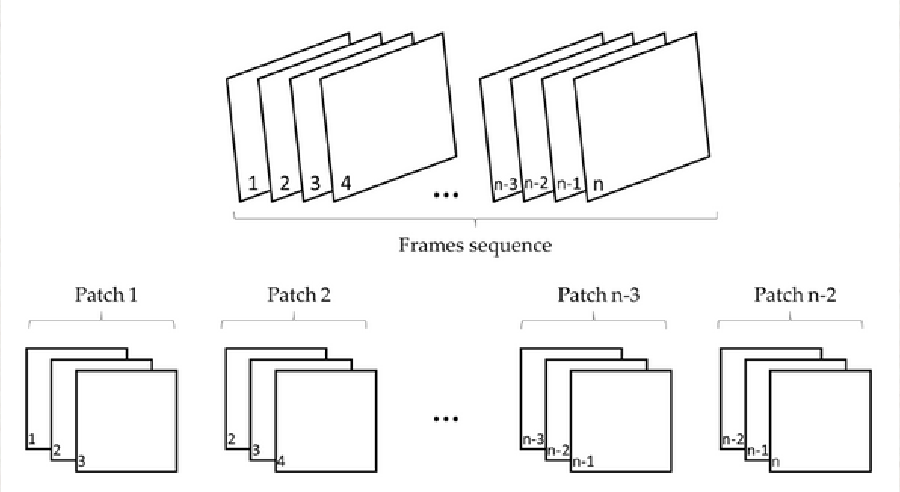
[8] A. Creswell et al., "Generative Adversarial Networks: An Overview" *IEEE Signal Processing Magazine*, vol. 35, pp. 53–65, 2018.

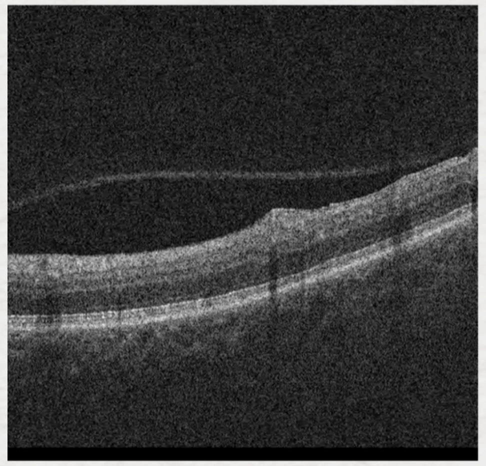
[9] S. Nishimoto et al., "Generating intermediate slices with U-nets in craniofacial CT images" *medRxiv*, p. 2024.05.08.24307089, Jan. 2024. DOI: 10.1101/2024.05.08.24307089.

04.B INTERMEDIATE SLICE GENERATION - PATCH EXTRACTION AND LOSS

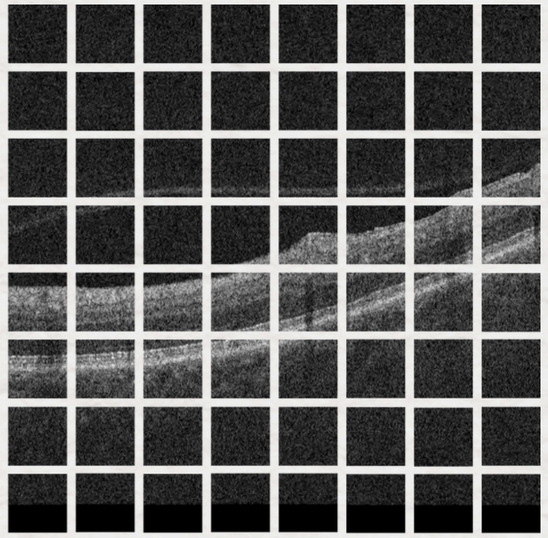
Training Patches:

- The network was trained using **triplets** of **64 x 64 patches**, extracted from B-scans previously **resized** to 496 x 512.





496x512



64x64x64

Generator Loss

Adversarial Loss

Multi-scale Structural Similarity Index Measure Loss (MS-SSIM)

0.05

6.0

$$\mathcal{L}_{\text{Gen}} = \lambda_{\text{adv}} \times \mathcal{L}_{\text{adv}} + \lambda_{\text{MAE}} \times \mathcal{L}_{\text{MAE}} + \lambda_{\text{MS-SSIM}} \times \mathcal{L}_{\text{MS-SSIM}} + \lambda_{\text{GDL}} \times \mathcal{L}_{\text{GDL}}$$

1.0

1.0

Mean Absolute Error

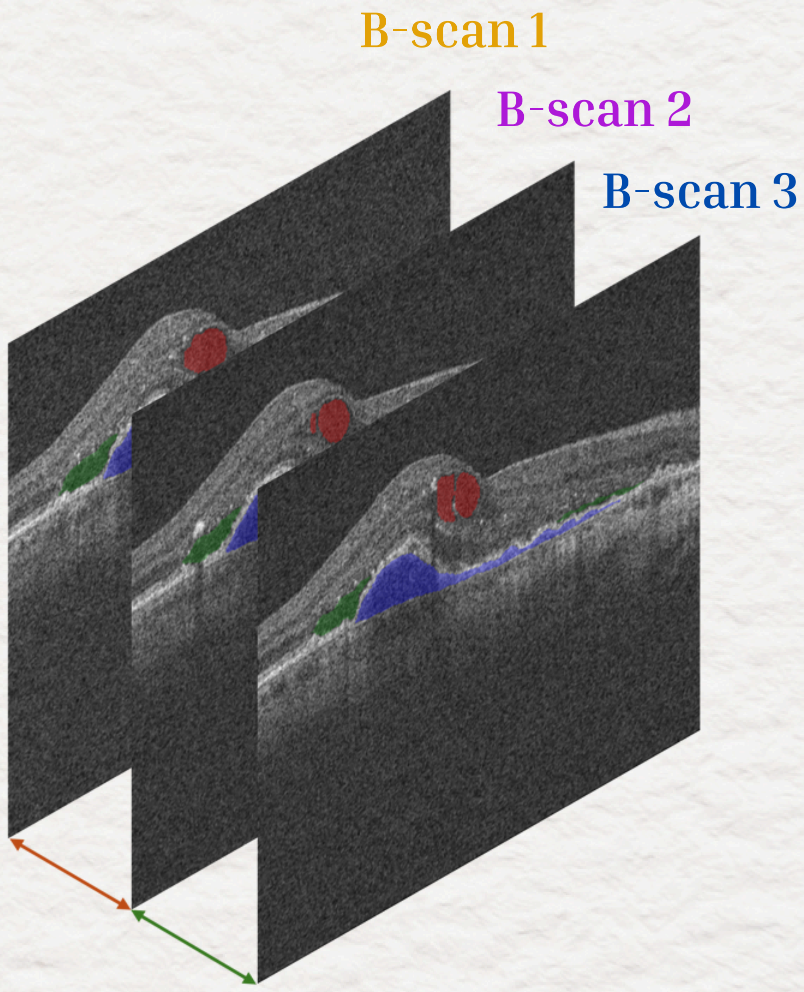
Gradient Difference Loss

Discriminator Loss: Binary Cross-Entropy

04.C FLUID VOLUME ESTIMATION

How the volume is calculated:

- Calculate the area of fluid in a B-scan and multiply by half the distance to the neighbouring slices;
- Sum the fluid volumes estimated from the B-scans that compose an OCT volume;
- The OCT volume metadata is used in the estimation of the real volume represented in a voxel.



IRF Volume from B-scan 2:

IRF area x 0.5 x distance to the previous scan

\+

IRF area x 0.5 x distance to the next scan

IRF Volume from B-scan 1 and B-scan 3:

B-scan 1: IRF area x 0.5 x distance to the next scan

B-scan 3: IRF area x 0.5 x distance to the previous scan

Total IRF Volume in the OCT scan:

IRF Volume from B-scan 1 + IRF Volume from B-scan 2 + IRF Volume from B-scan 3

05 EXPERIMENTS

05.A DATA PARTITION

Cross-validation | **5-fold** validation, applied to **entire OCT volumes**

In Segmentation:

- **Only the volumes with fluid masks** were used (70 OCT volumes);
- Different split was used for each target fluid, in binary segmentation;
- Fold 1 was randomly selected as the reserved fold.

In Generation:

- **All volumes** were used;
- One of the folds in this partition is the reserved fold from the segmentation experiments.

Volumes marked with ***a*** consist of 128 B-scans
Volumes marked with ***b*** consist of 64 B-scans

		Folds				
		0	**1**	**2**	**3**	**4**
Vendors	**Cirrus**	5	4	5	5	5
	Spectralis	5	5	5	5	4
	Topcon	5^a	$4^a + 1^b$	$3^a + 1^b$	4^a	5^a

Volumes marked with ***a*** consist of 128 B-scans
Volumes marked with ***b*** consist of 64 B-scans

		Folds				
		0	**1**	**2**	**3**	**4**
Devices	**Cirrus**	9	4	8	9	8
	Spectralis	9	5	8	9	8
	T-1000	3	3	3	2	3
	T-2000	4^a	$1^a + 1^b$	5^a	5^a	5^a

05.B EXPERIMENTS DESCRIPTION

EXPERIMENTS	*DESCRIPTION*
Experiment 1 | Fluid Segmentation with a <u>Multi-class U-Net</u>	1.1 – Training with patches of size 256 x 128 pixels **randomly extracted from the ROI** (A). 1.2 – Training with **large patches** of size 496 x 512 pixels (B). 1.3 – Training with a **varying number of vertical patches** (four, seven, or thirteen) of size 496 x 128 pixels, extracted from B-scans **resized** to 496 x 512 pixels (C).
Experiment 2 | Fluid Segmentation with <u>three Binary U-Nets</u>	2.1 – Training with the **same loss** as in Experiment 1, using **two different data partitions**. Two merging strategies were tested: probability-based and priority-based. 2.2 – Training with the **balanced cross-entropy loss**.
Experiment 3 | Synthesis of Intermediate Slices using a <u>GAN</u>	**GAN** trained to **generate** a fake B-scan between a pair of two known B-scans.
Experiment 4 | Synthesis of Intermediate Slices with a <u>generative U-Net</u>	**Generative U-Net** trained to **synthesize** a fake B-scan between a pair of two known B-scans.
Experiment 5 | Fluid Volume Estimation in <u>Real</u> OCT Volumes	**Estimation of the fluid volumes** from the **GT** and the **fluid masks** predicted by the segmentation model.
Experiment 6 | Fluid Volume Estimation in <u>Super-resolved</u> OCT Volumes	**Estimation of the fluid volumes** for **generated** and **super-resolved** OCT volumes, obtained through the **GAN** and using the masks predicted by the segmentation model.

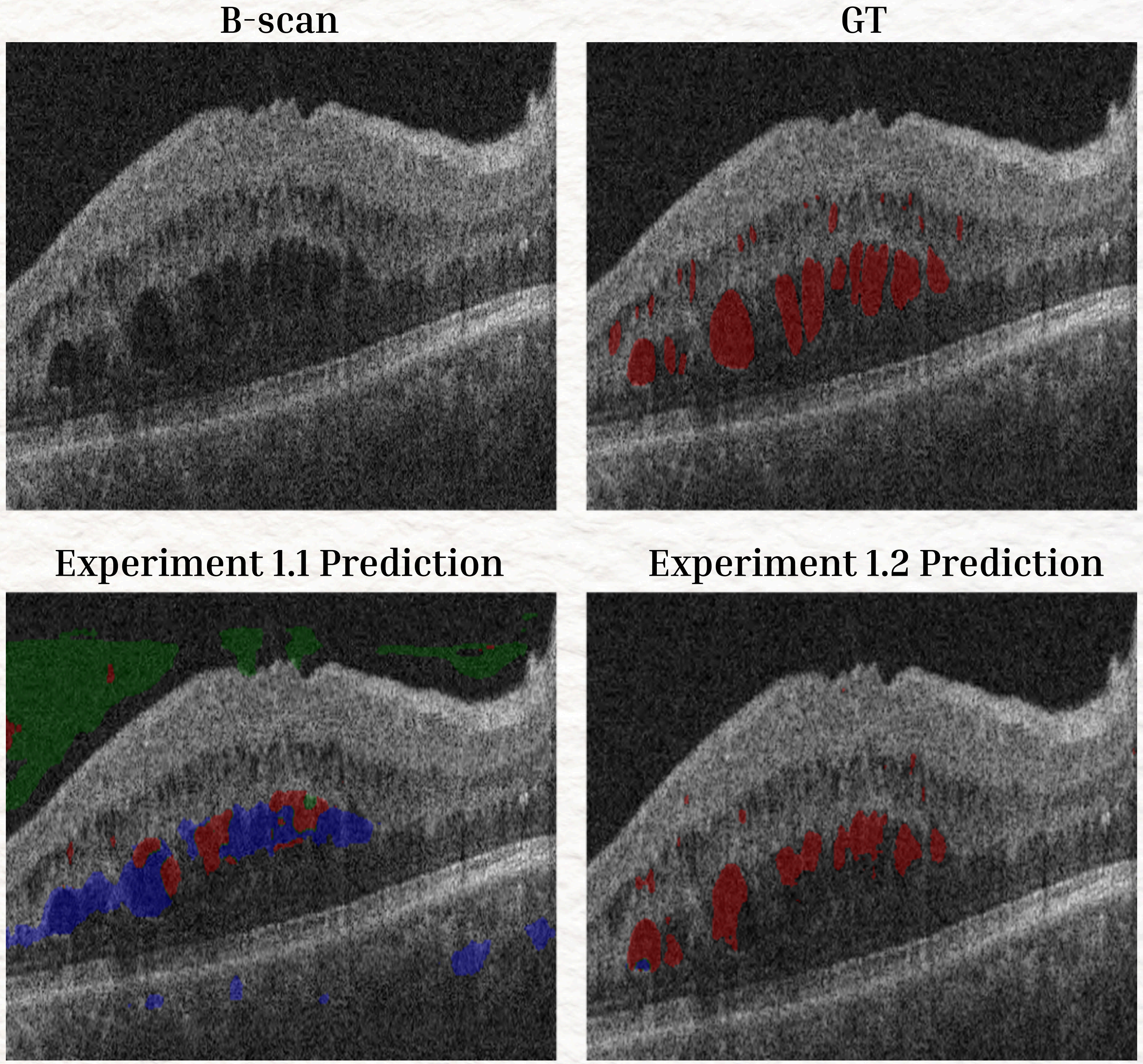
06.A FLUID SEGMENTATION WITH A MULTI-CLASS U-NET

Experiment 1.1 - Small Patch Extraction (A):

- **Poor** segmentation results;
- The patches were **not big enough** to capture the transitions between the retina and the background;
- The model did not understand the **relationships between the retinal anatomy and fluids**.

Experiment 1.2 - Large Patch Extraction (B):

- **Better performance** than in Experiment 1.1;
- The segmentations of smaller fluid regions were **not very accurate**;
- During patch extraction, the retina would often be **split into two patches**, resulting in a poor understanding of the transitions between the retina and the choroid.



IRF in red, SRF in green, and PED in blue

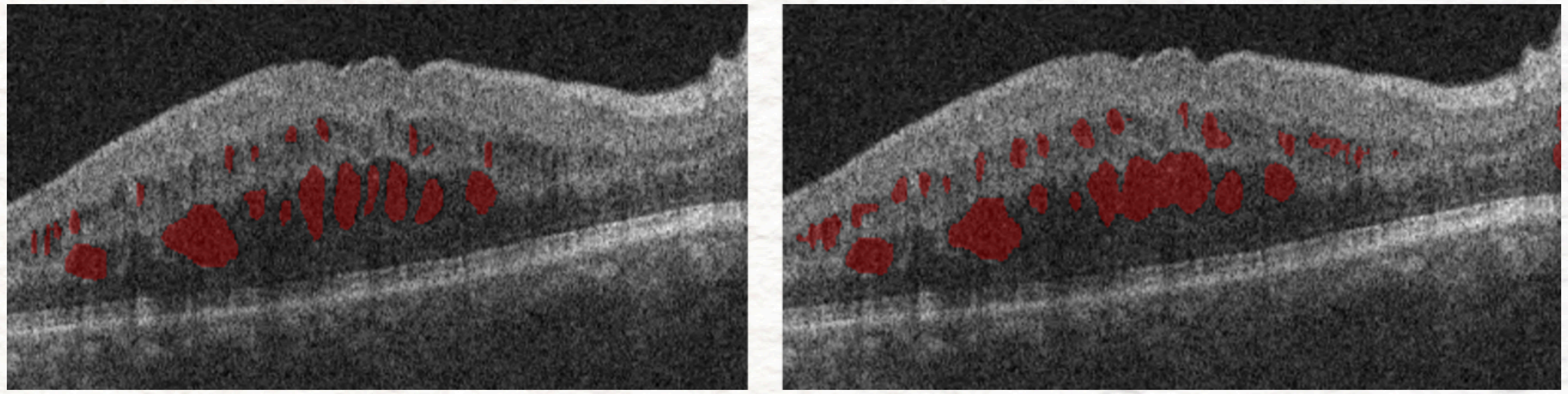
06 RESULTS AND DISCUSSION

06.A FLUID SEGMENTATION WITH A MULTI-CLASS U-NET

Experiment 1.3 - Vertical Patch Extraction (C):

- **Most extensive experiment**;
- **Better performances** than in the previous experiments;
- The vertical patches **understood the anatomy of the retina**;
- The patches were **small enough** for the detailed segmentation of smaller fluid regions.

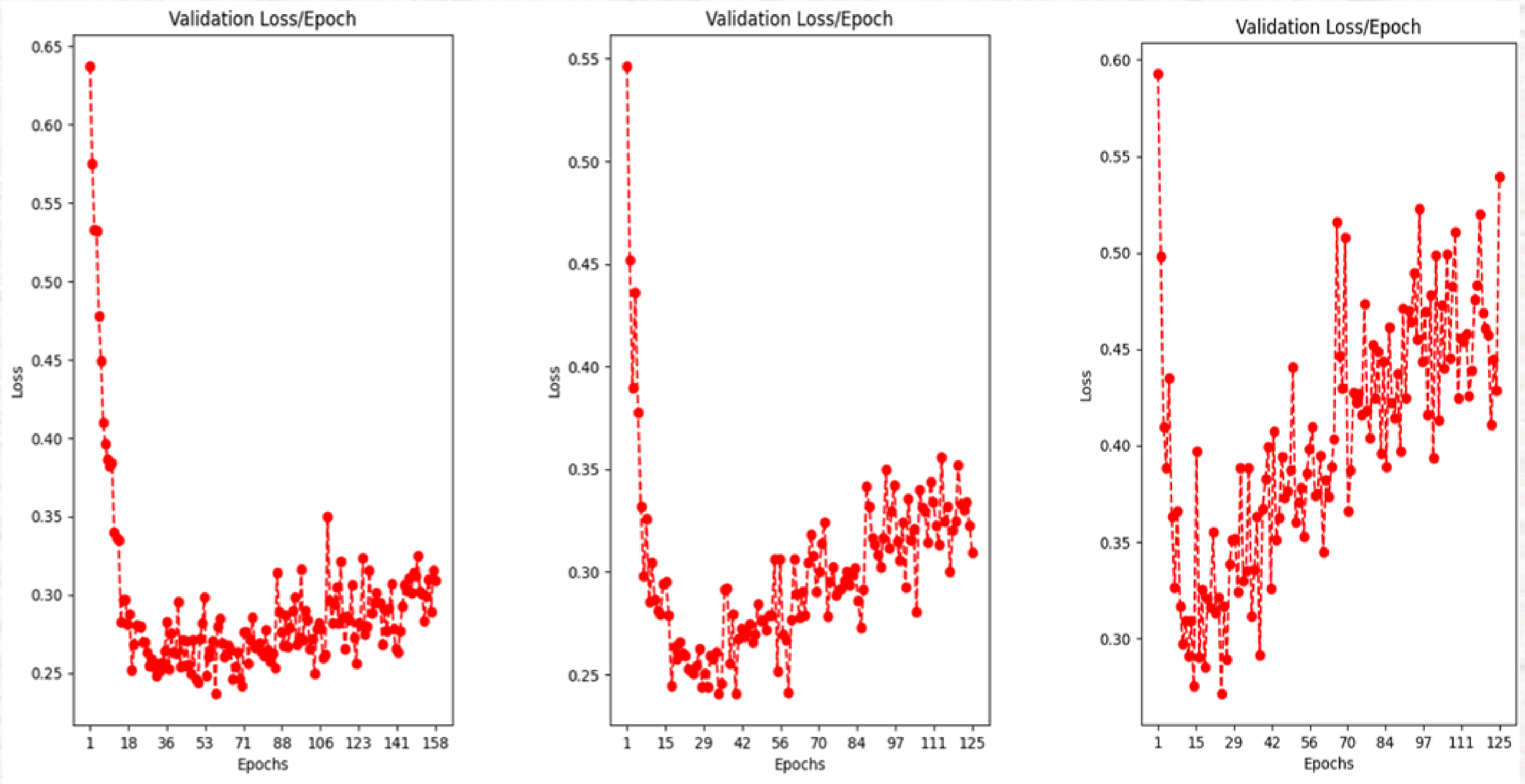
GT | Experiment 1.3 Prediction



IRF in red

The impact of using four, seven, or thirteen patches per B-scan:

- As the number of patches per B-scan increased, the **convergence was faster**, but the **training was slower**;
- The models trained with four and seven patches per B-scan performed better than when trained with thirteen.

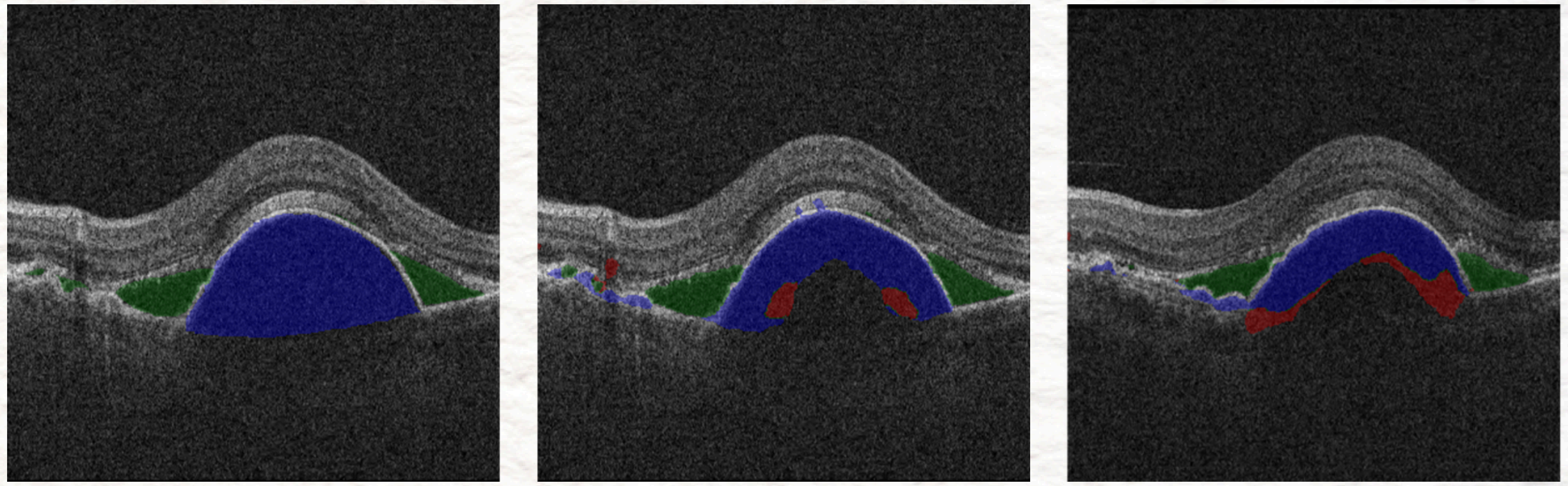


Four | Seven | Thirteen

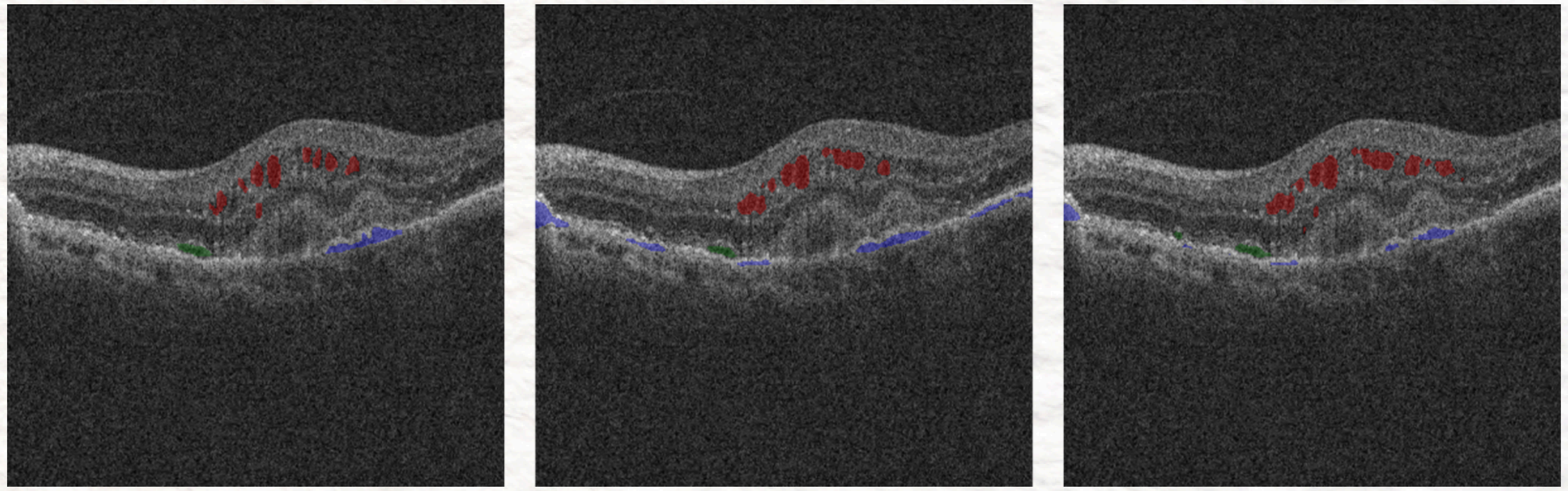
06.A FLUID SEGMENTATION WITH A MULTI-CLASS U-NET

The impact of rotation in the training data:

- Transforming the inputs through random rotations **significantly impacted** the model's performance;
- When the model was trained **without random rotations**, the performance **declined significantly**;
- The performance obtained when using a 5 degree random rotation was **slightly worse** than when using a 10 degree random rotation, but **less prone to oversegmentation** and **more anatomically coherent**.



IRF in red, SRF in green, and PED in blue



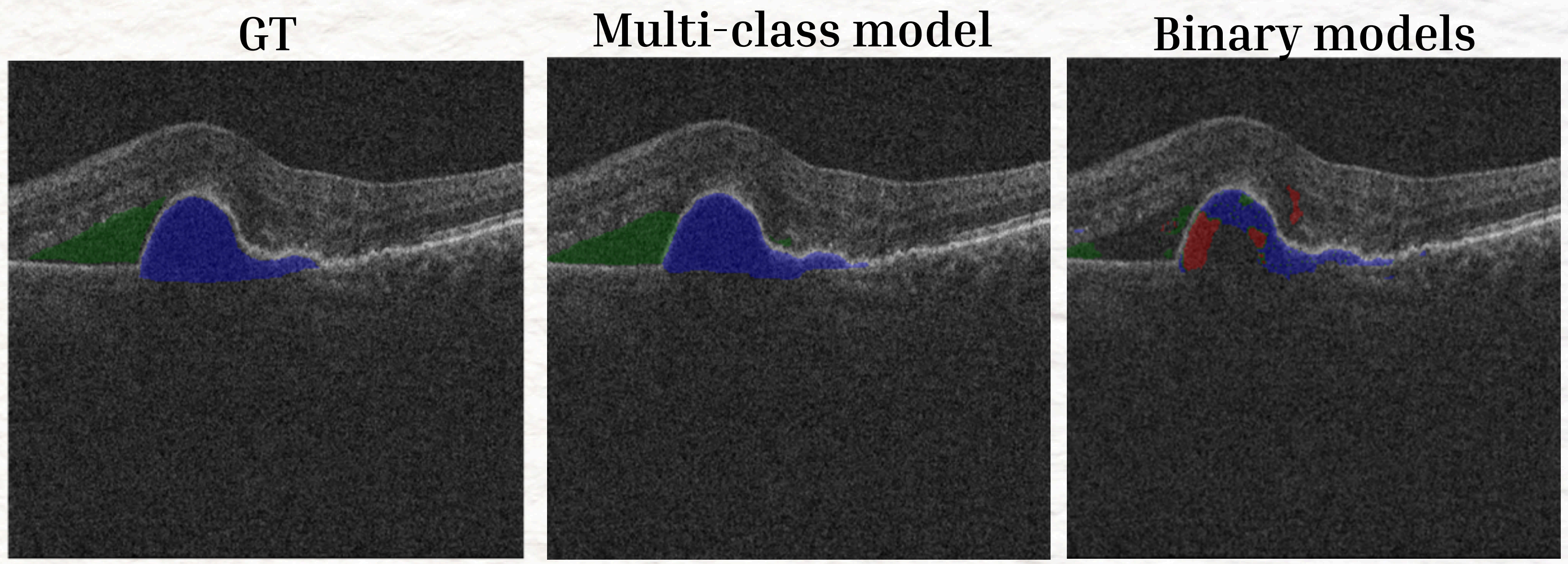
IRF in red, SRF in green, and PED in blue

06 RESULTS AND DISCUSSION

06.B FLUID SEGMENTATION WITH THREE BINARY U-NETS

Experiment 2.1 - Multi-class Loss:

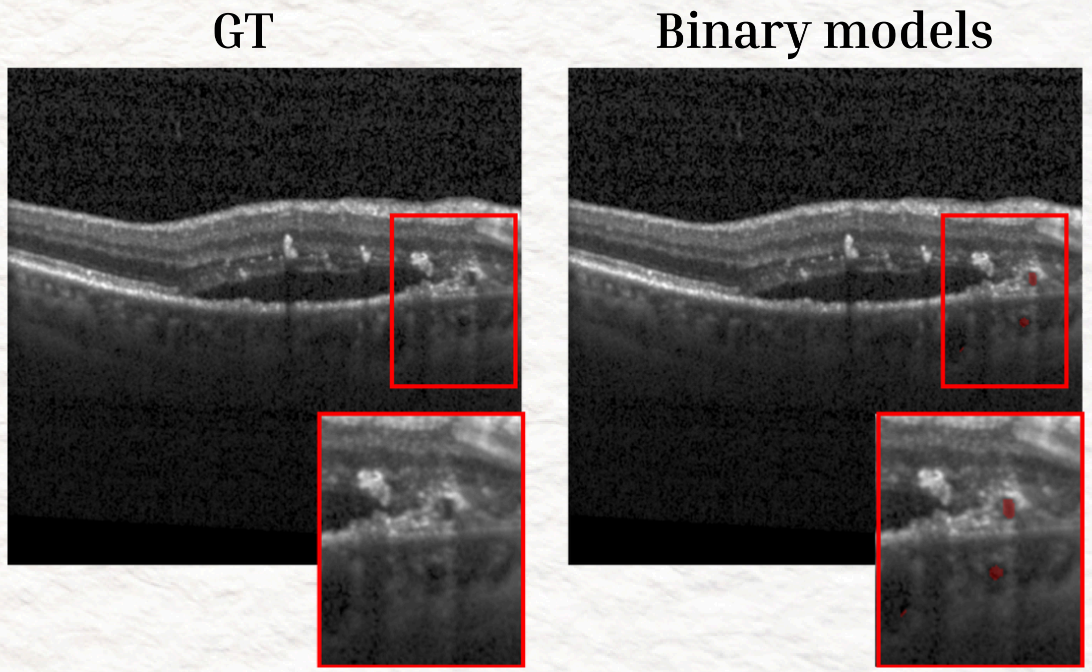
- **Worse performance** than the achieved by the multi-class model;
- The models were **prone to oversegment**, due to the **lack of inter-class competition**;
- **Less aware of the anatomical relationships**, a key capability developed in the multi-class model for the correct pixel classification.



IRF in red, SRF in green, and PED in blue

Experiment 2.2 - Balanced Cross-Entropy Loss:

- **Significantly worse performance**;
- The models segmented, in almost all slices, **small regions of fluid**;
- These small segmentations are **not heavily penalized through the loss function.**



IRF in red

06 RESULTS AND DISCUSSION

06.C FLUID SEGMENTATION - COMPARISON IN RETOUCH

Comparison with the literature:

- The mean Dice coefficients, computed across the four validation folds of the best performing segmentation models, were **compared to the work of Alsaih et al. [10]**;
- In this work, the authors implemented **three networks** (trained with different optimizers) to segment the retinal fluid in the RETOUCH dataset.

Training the network on patches **significantly improves** the fluid segmentation;

Performs equally in IRF, but the larger patches **promote a better SRF** segmentation while **harming the PED** delimitation;

Our implementation achieved **comparable results** with **lower computational cost**.

Dice Coefficients

Runs	IRF	SRF	PED
Multi-class	0.67	**0.80**	0.72
Binary	0.60	0.75	0.44
U-Net trained with entire image [10]	0.29	0.47	0.18
U-Net trained with 128 x 128 patches [10]	**0.68**	0.61	**0.84**

[10] K. Alsaih et al., “Deep learning architectures analysis for age-related macular degeneration segmentation on optical coherence tomography scans” *Computer Methods and Programs in Biomedicine*, vol. 195, p. 105 566, 2020, ISSN: 0169-2607. DOI: 10.1016/j.cmpb.2020.105566.

06 RESULTS AND DISCUSSION

06.D FLUID SEGMENTATION - PERFORMANCE IN CHUSJ

Results in CHUSJ Dataset:

- **Satisfying segmentation** of the retinal fluids;
- Consistently **predicted fluid** in **choroid region**;
- Similarly, the **PED** was **constantly oversegmented**, influenced by the segmentation criteria seen in the training data.

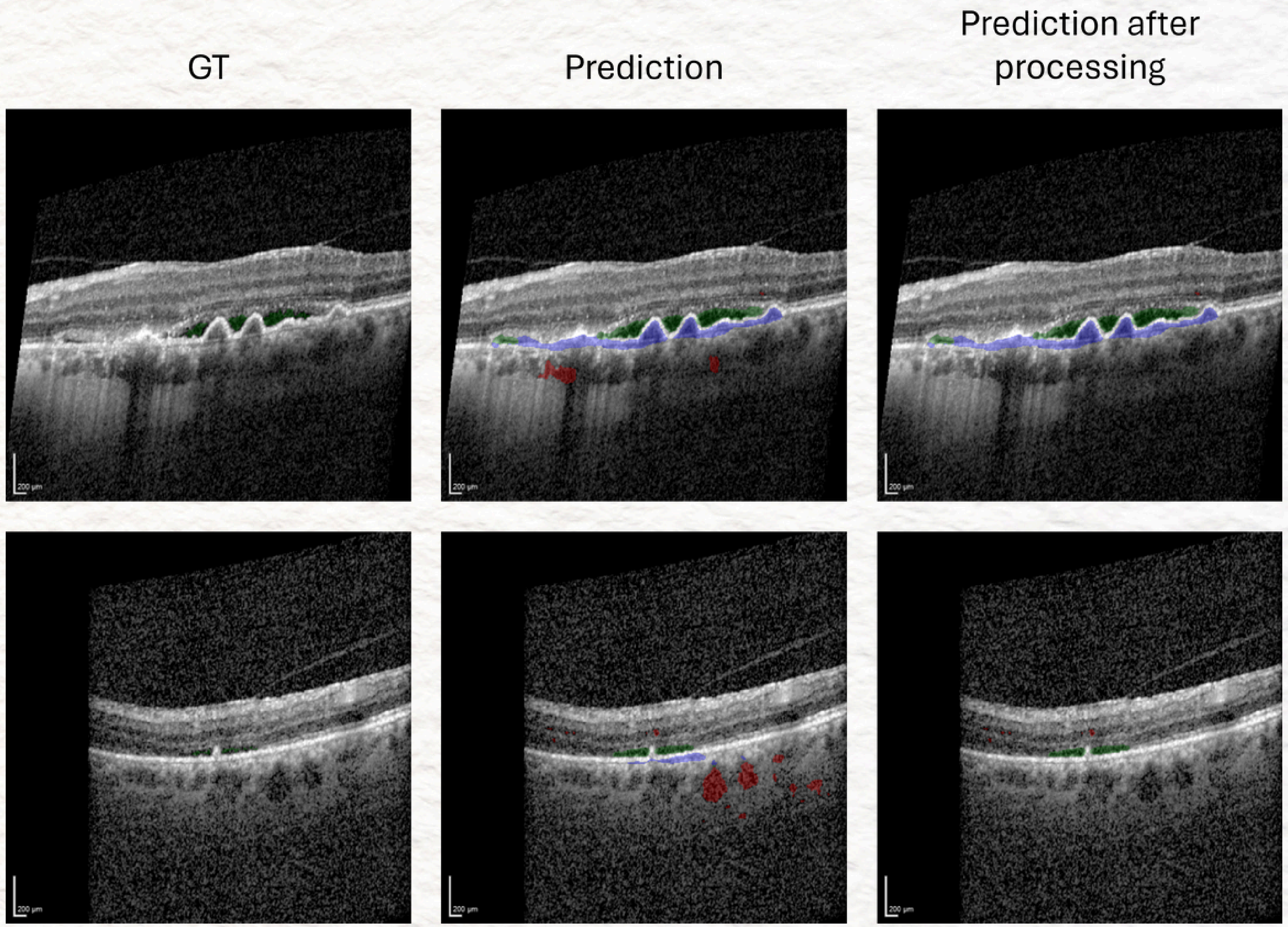
Caused by the differences between datasets

Processing with retinal delimitation:

- The overall **Dice coefficient** was **significantly improved**;
- The **anatomical coherence** of the segmentations was also **enhanced**.

Dice Coefficients

	IRF	SRF	PED
Without Retinal Delimitation	0.17	0.44	0.14
With Retinal Delimitation	0.52	0.68	0.39

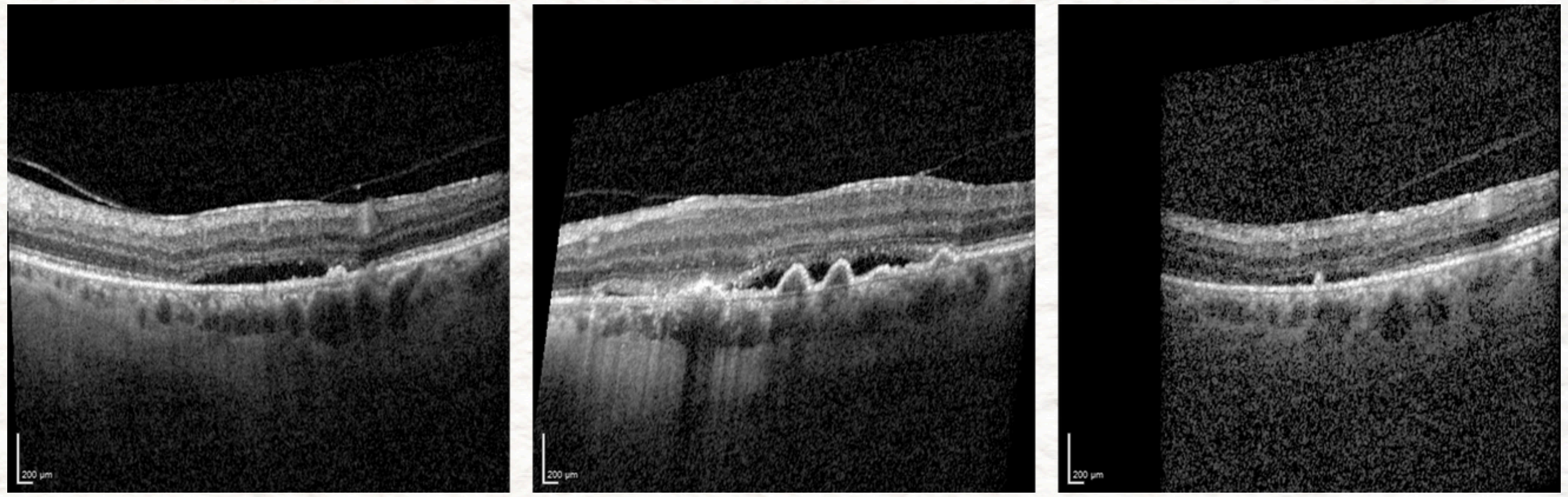


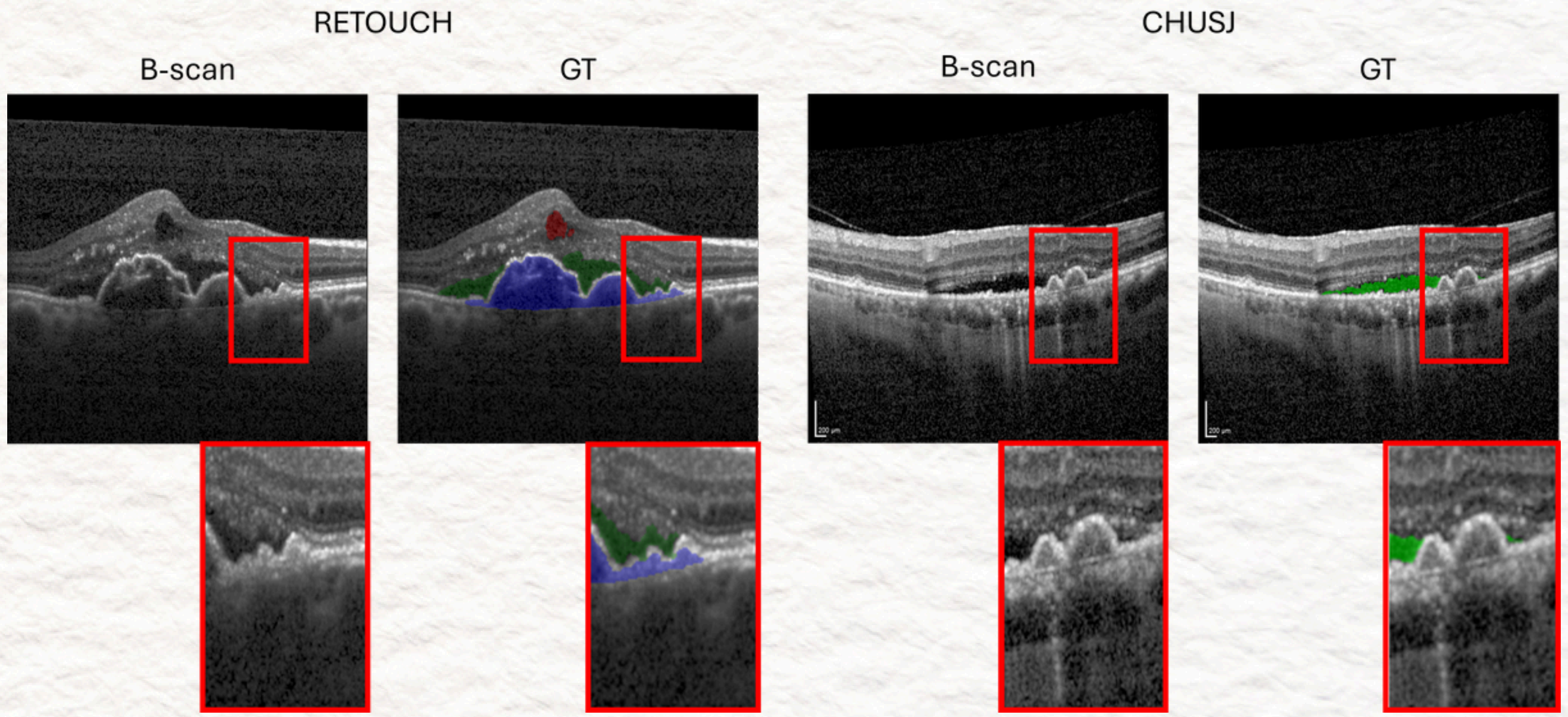
IRF in red, SRF in green, and PED in blue

06.D FLUID SEGMENTATION - PERFORMANCE IN CHUSJ

CHUSJ Dataset Characteristics which Influence the Segmentation Performance:

- The **choroid definition** and **noise levels**, which are significantly different from the training data;
- Only **one clinician** segmented the retinal fluids, while training data was annotated by **two clinicians**;
- The segmentation criteria of PED **differs from the criteria** in the training data.

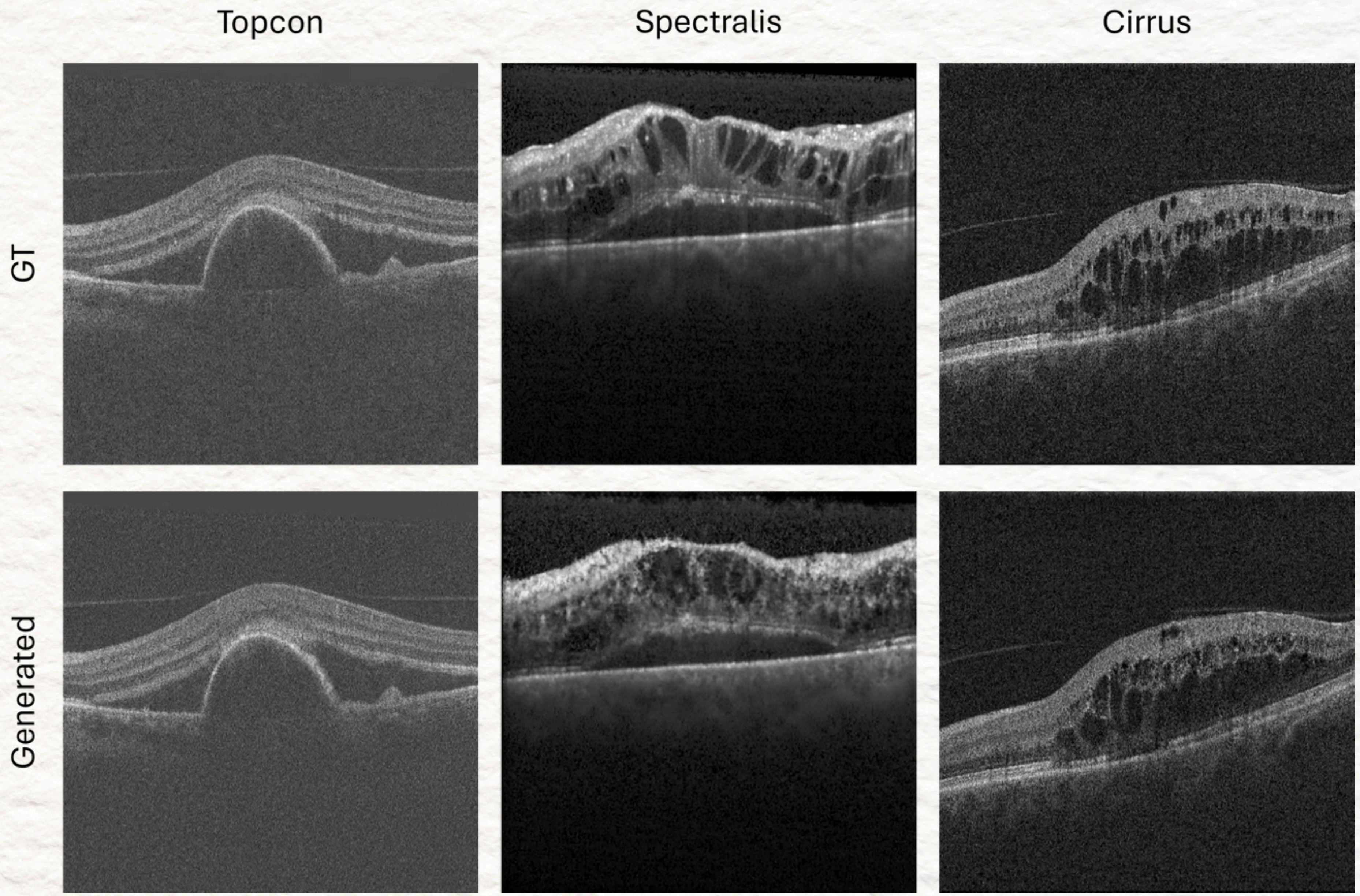




IRF in red, SRF in green, and PED in blue

06.E INTERMEDIATE SLICE GENERATION - GAN

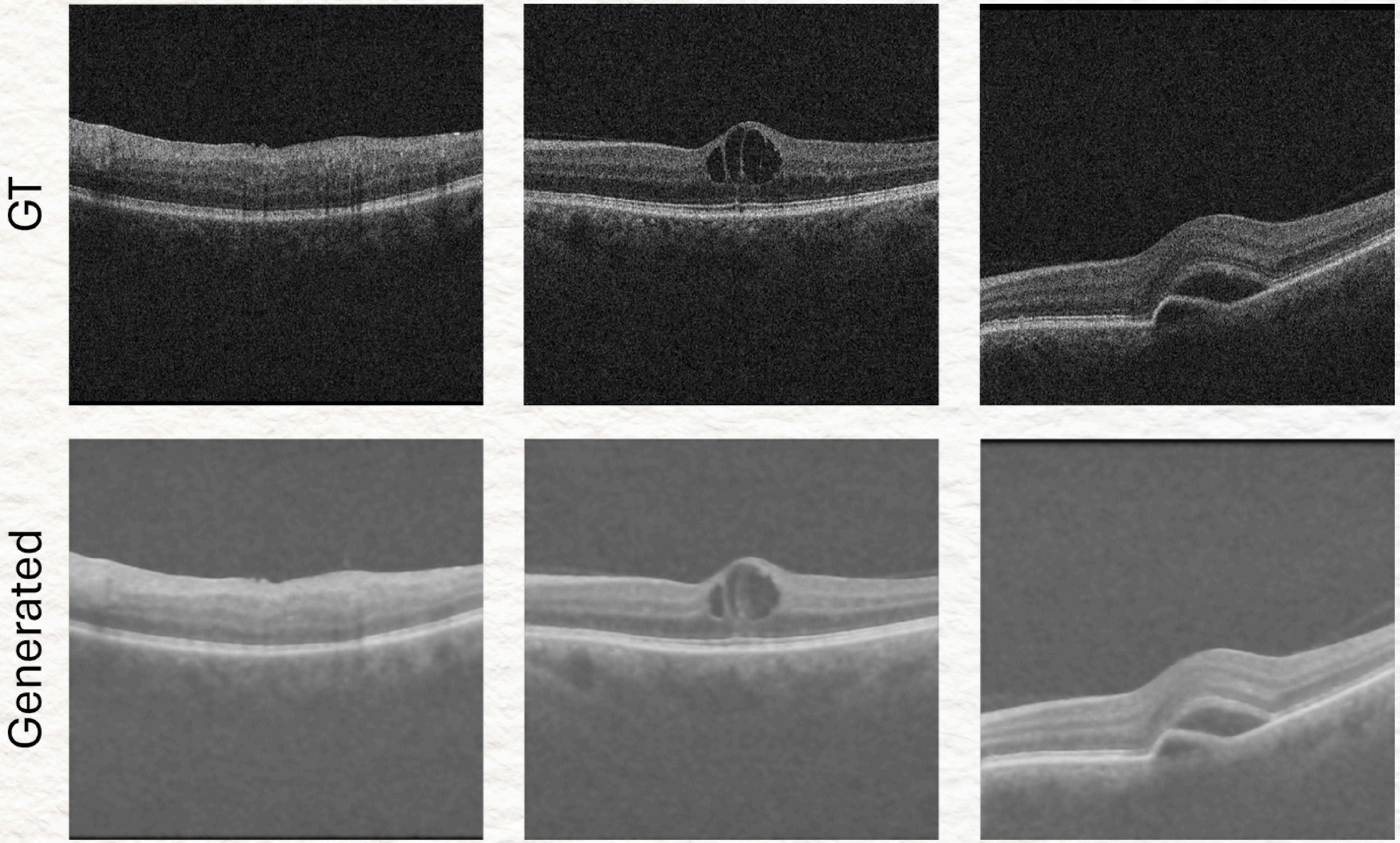
- **Realistic** B-scans were generated;
- The **distance between slices** significantly influenced the generation of intermediate slices;
- **Larger fluids** are **easier** to generate;
- The **generation of IRF** is the **most difficult task**;
- In OCT, losses based on pixel-wise comparisons need to be **complemented** with **strong perceptual losses**.



06 RESULTS AND DISCUSSION

06.F INTERMEDIATE SLICE GENERATION - U-NET

- Produced **smoothed intermediate images**, with **poor details and definition**;
- By training the network using only a **pixel-wise loss** (MAE), the model produces **blurred outputs**;
- The introduction of **perceptual components** to the network loss would lead to a **more accurate** representation of the B-scans.



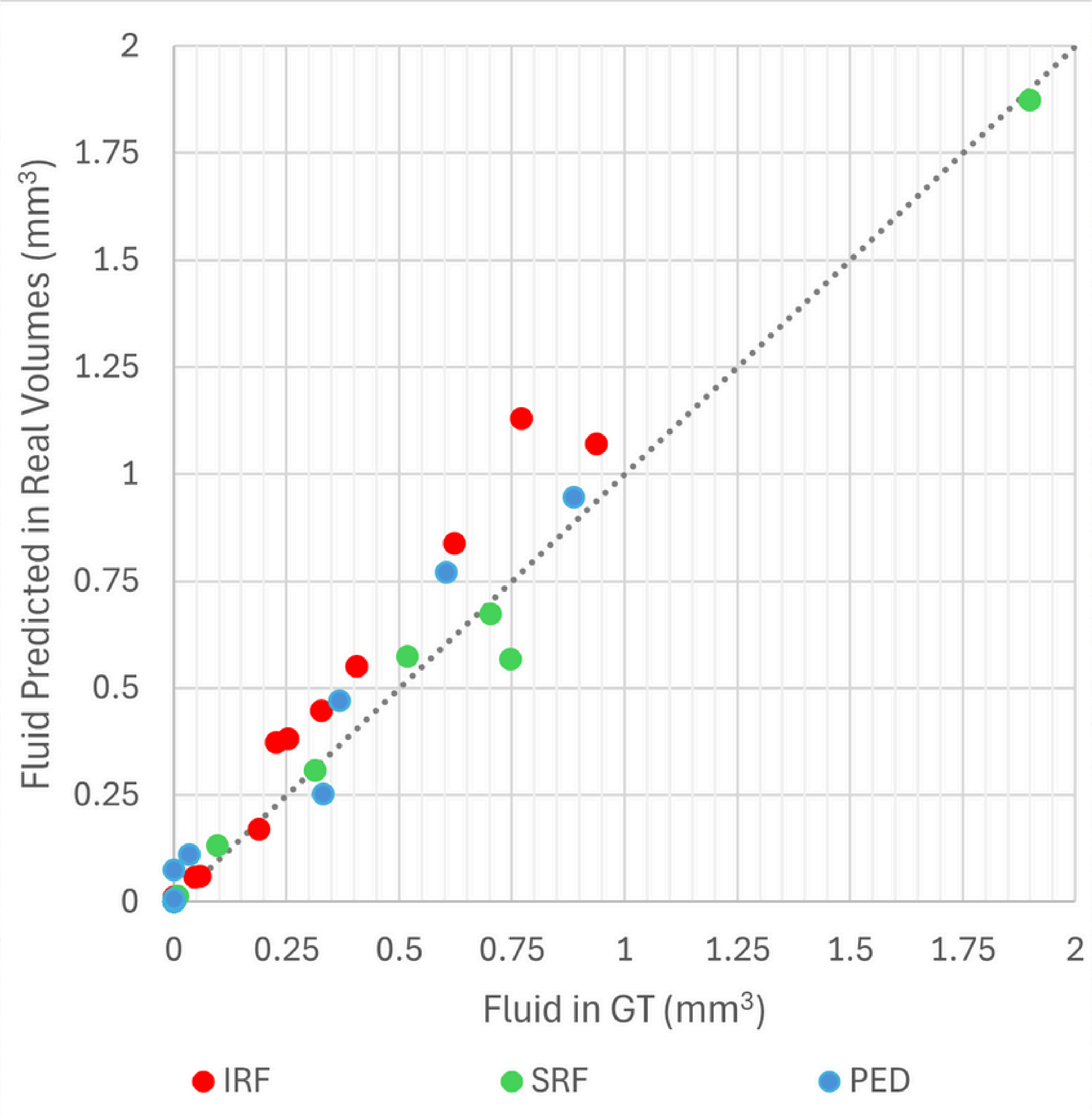
06 RESULTS AND DISCUSSION

06.G FLUID VOLUME ESTIMATION IN ORIGINAL OCT VOLUMES

- **Similar** to the values estimated using the GT masks;
- The volume differences are **limited by the capabilities of the segmentation model**;
- Most errors occur in IRF, where the **model performed worse** and the **fluid regions are smaller**.

Fluid Volume MAE in mm^3 between GT and predicted masks

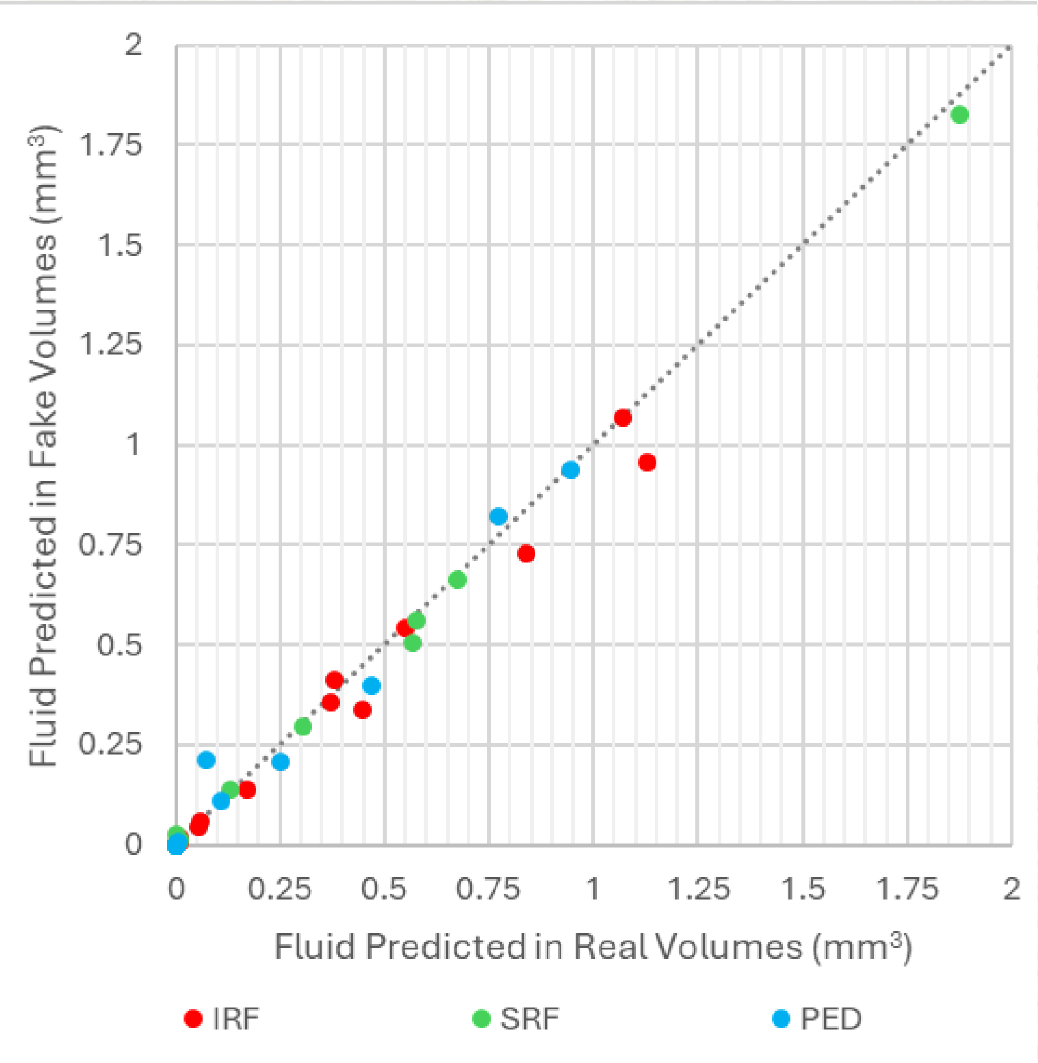
	IRF	SRF	PED
MAE	0.092	0.025	0.041

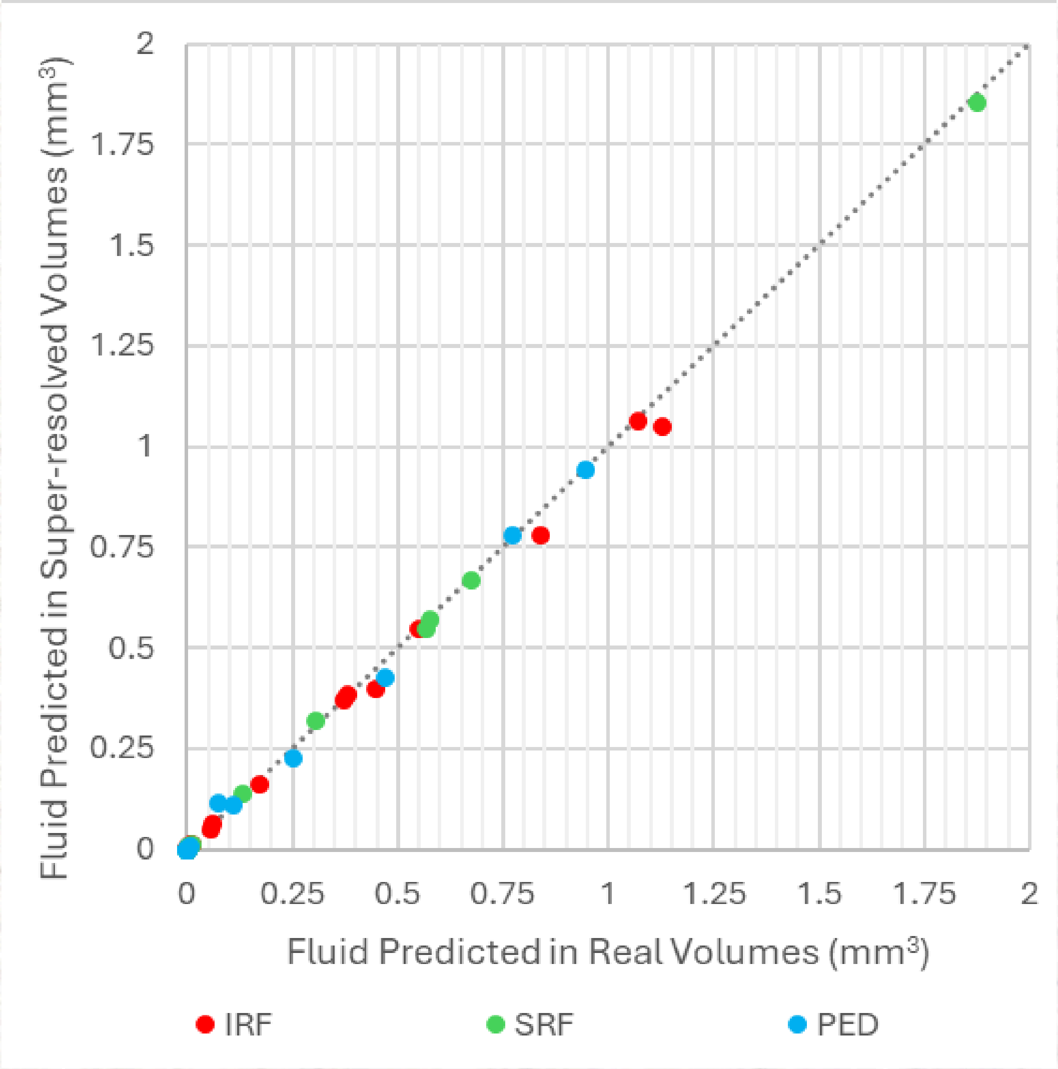


06 RESULTS AND DISCUSSION

06.H FLUID VOLUME ESTIMATION IN SUPER-RESOLVED OCT VOLUMES

- The results in the super-resolved volumes are **really similar** to those predicted through the fluid segmentation of the original OCT volumes;
- Most differences occur in the **segmentation of generated** IRF, which was the **fluid generated with poorer quality**;
- The **Spectralis** volumes were those where most differences appear, since the quality of the generated fluid is poorer, due to the **larger inter-slice distance**.





06 RESULTS AND DISCUSSION

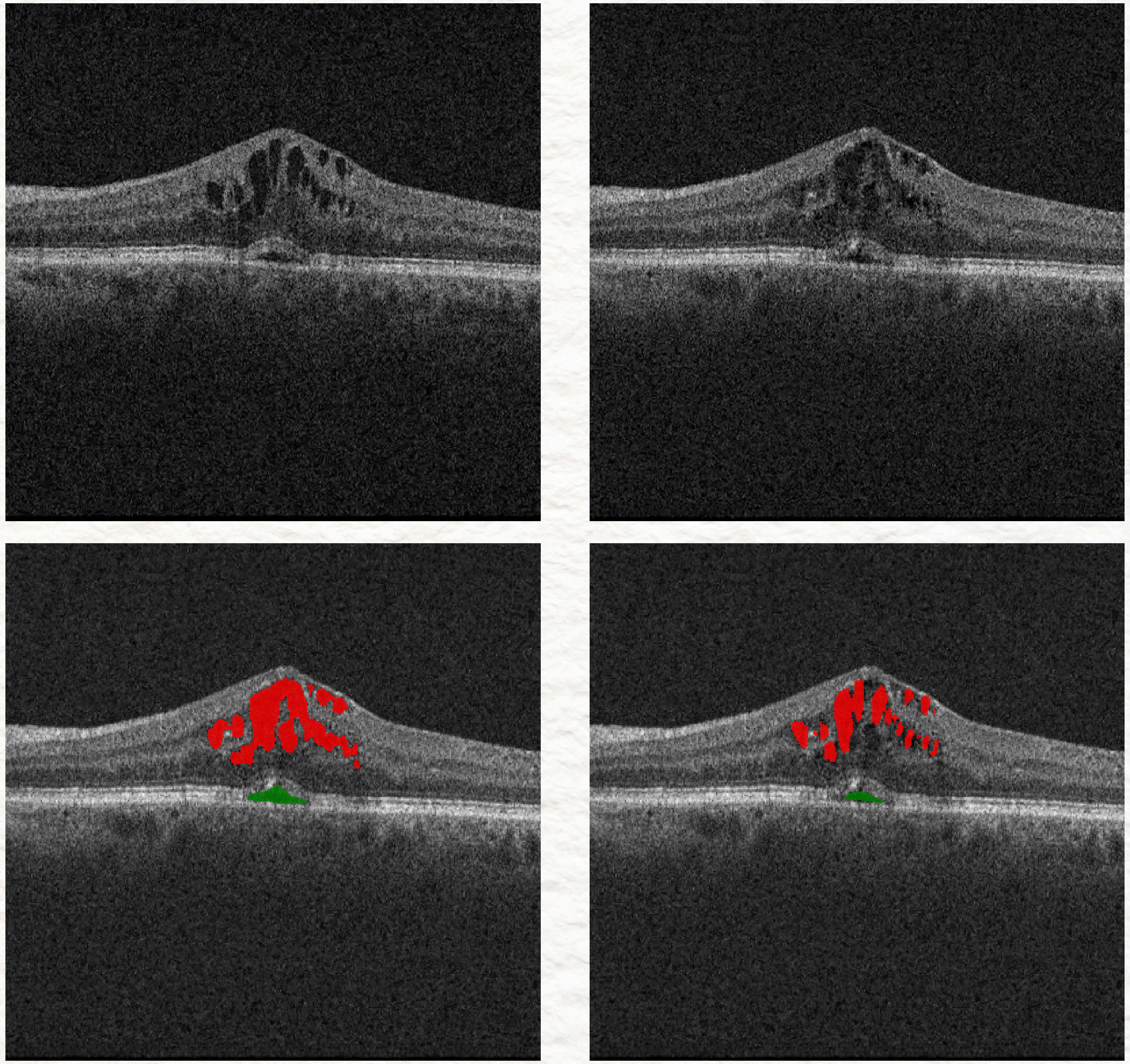
06.H FLUID VOLUME ESTIMATION IN SUPER-RESOLVED OCT VOLUMES

Fluid Volume MAE in mm^3 between predicted masks in the original OCT volumes and in the generated OCT volumes

	IRF	SRF	PED
MAE	0.037	0.012	0.023

Fluid Volume MAE in mm^3 between predicted masks in the original OCT volumes and in the super-resolved OCT volumes

	IRF	SRF	PED
MAE	0.016	0.005	0.009



IRF in red and SRF in green

07 CONCLUSIONS

SEGMENTATION

- The **size of the input patchs** and **how they are extracted** strongly impact the performance of the segmentation model;
- The **random transformations** applied to the input change the model's performance;
- Results **comparable with the literature** were obtained, despite using a **computationally cheaper** approach.

GENERATION

- **Inter-slice distance** is very relevant in intermediate slice generation;
- **Pixel-wise losses** motivate the generative models to produce **accurate representations of speckle noise**, disregarding the appearance of some anatomically important structures;
- **Small fluid regions** are harder to generate, as they change faster between consecutive B-scans.

FLUID VOLUME ESTIMATION

- Inherently **dependent** on the **segmentation** and **generation quality**;
- **More influenced** by the **segmentation model** than the generative model;
- **Less accurate in IRF**, which is hard to segment and generate due to its **anatomical characteristics**.

08 LIMITATIONS AND FUTURE WORK

LIMITATIONS

SEGMENTATION

- Baseline U-Net is **not capable** of both **detailed and broader** segmentation;
- **Generalizes poorly** in other data sources.

GENERATION

- Focuses more on the **accurate representation of speckle** rather than the **generation of fluids and other anatomically relevant structures**.

FUTURE WORK

SEGMENTATION

- Implementation of **more complex networks**, which combine shallow and deep features;
- **Condition the input** of the segmentation network to **include additional information** of the retina;
- Train the model on data from **different sources**.

GENERATION

- Include **stronger perceptual losses** (e.g. LPIPS);
- Introduce **attention blocks** into the generator architecture.

THANK YOU!

ANY QUESTION?