

# SMRVIS: Point cloud extraction from 3-D ultrasound for non-destructive testing

Lisa Tang

## Abstract

We propose to formulate point cloud extraction from ultrasound volumes as an image segmentation problem. Through this convenient formulation, a quick prototype exploring various variants of the Residual Network, U-Net, and the Squeeze and Excitation Network was developed and evaluated. This report documents the experimental results compiled using a training dataset of 5 labeled ultrasound volumes and 84 unlabeled volumes that got completed in a two-week period as part of a challenge submission to a CVPR competition entitled “Deep Learning in Ultrasound Image Analysis”. Source code is shared with the research community at this GitHub URL <https://github.com/lisatwyw/smrvis>.

*Keywords:* *Ultrasound volumes; Non-Destructive Testing; Pipe; Manufacture Defects; Residual Networks; U-net++; W-net; Attention Net; Recurrent-Residual U-net*

## 1 Introduction

As part of a submission to a computational challenge entitled “Deep Learning in Ultrasound Image Analysis” hosted by a CVPR workshop, this report presents experimental results compiled using a training dataset of 5 labeled ultrasound volumes and 85 unlabeled volumes. Due to resource and time constraints, we focus on finding a workable solution that can be quickly prototyped and tested. Accordingly, we propose to formulate three-dimensional mesh estimation from ultrasound volumes as an image segmentation problem. To this end, we explored W-unet, R2Unet, SE-Unet, and Attention Unet and Unet++, which are, in a nutshell, variants of the U-net [10]. Source code is shared with the research community at <https://github.com/lisatwyw/smrvis>. We nicknamed this proposed framework as SMRVIS partly to draw an analogy to our previous work called SMRFI that formulated shape matching as a feature image registration problem. Instead of embedding 2D shapes with feature values and assigning the feature values to the nearest voxel of the embedding space, the present framework simply encodes the positions of the reference mesh vertices with values from 0 to 1.

## 2 Materials

A training set of 90 ultrasound volumes were provided by the challenge organizers. Each scan captures piece(s) of steel pipe, potentially containing artifacts inside these pipes. Corresponding Surface mesh of the pipe (pieces) were created by an “experienced data analyst” [4]. Five of these surface meshes (corresponding to volumes 1 to 5) were provided to the challenge participants and herein referred to as reference masks. Figure 1-5 show examples of the surface renderings of the annotations of the pipes. As the reference labels for the remaining 85 volumes were not provided at the time of challenge, these volumes were mainly used in this study as test samples.

## 3 Methods

### 3.1 Preprocessing

Each of the reference mesh was first encoded into an image representation. To do so, the vertices of each mesh were read into memory. Next, a binary mask was created to encode the mesh vertices, taking into account the voxel spacing of the ultrasound volumes. To facilitate the learning, the point cloud mask was ‘thickened’ so that the edges of the binary mask softens. An alternative approach would be to apply Gaussian blur [5] but we found this simpler approach more memory and time efficient.

To read in the corresponding ultrasound volume, a meta file was created for each of the raw ultrasound files (example in Sec. C; script is also provided under [https://github.com/lisatwyw/smrvis/blob/main/utils/write\\_meta.sh](https://github.com/lisatwyw/smrvis/blob/main/utils/write_meta.sh)) and python package SimpleITK was used (example code snippet follows). To ensure that the

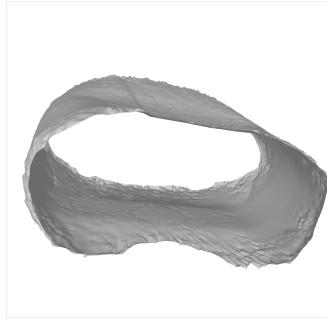


Figure 1: Reference mesh of volume 1.



Figure 2: Reference mesh of volume 2.

fitted models could be robust to noise, We only preprocessed the input ultrasound data with two steps: down-sampling and rescaling their intensity values to [0,1].

```

import SimpleITK as sitk

vols={}
i = 1 # sample_id
filename='train_data/training/volumes/scan_%03d.mhd'% i

header=sitk.ReadImage(filename)
vols[i]=sitk.GetArrayFromImage(header)

from plyfile import PlyData, PlyElement
plydata,verts,faces=[{}],[],[]

filename='train_data/training/meshes/scan_%03d.ply' % i
vx = plydata[i][ 'vertex'][ 'x']
vy = plydata[i][ 'vertex'][ 'y']
vz = plydata[i][ 'vertex'][ 'z']
verts[i] = [ (vx[d],vy[d],vz[d]) for d in range(len(vx)) ]
num_faces= plydata[i][ 'face'].count
faces[i] = [ plydata[i][ 'face'][d][0] for d in range(num_faces) ]

```

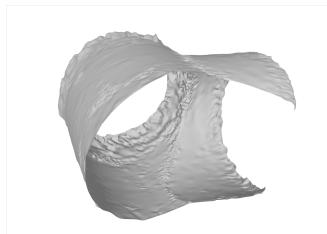


Figure 3: Reference mesh of volume 3.

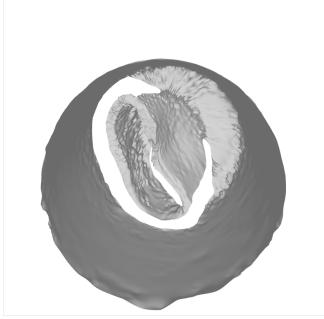


Figure 4: Reference mesh of volume 4.

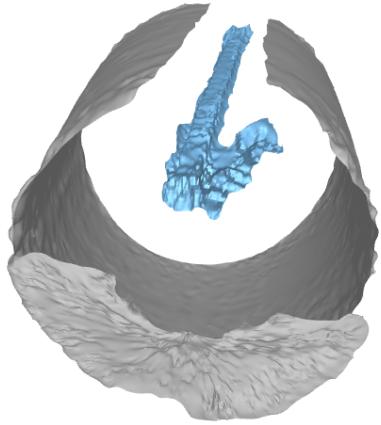


Figure 5: Surface rendering of a pipe (grey) containing artifacts (blue).

```

import pyvista
# we need to start the frame buffer even if plotting offline
pyvista.start_xvfb()
plotter = pyvista.Plotter(off_screen=True)
mesh = pyvista.read(ply_filename)
plotter.add_mesh(mesh, opacity=.3, color='grey')
plotter.add_title('Estimated mesh for volume # %d %i')
plotter.show(screenshot = out + '_screenshot.png')

```

## 3.2 Models

Based on observations and evidence from prior studies [7, 8], we elected to explore variants of the U-net, namely the W-net, the recurrent-residual U-net (R2Unet), SE-Unet, and the Attention U-net. In short, the R2Unet employs recurrent connections to improve the quality of the feature representations [9].

## 3.3 Model-training

### 3.3.1 Data augmentation

We augmented the training dataset by perturbing the input training set on-the-fly with random translations, rotations, and mirroring on the x- and y-axis.

### 3.3.2 Learning rate

We explored values of 0.0001, 0.008, 0.001, and 0.1 as the initial learning rate in the context of 3 learning rate scheduling schemes (cyclical decay, cosine decay, and polynomial decay).

### 3.4 Evaluation metrics and model selection

Following the Challenge’s evaluation protocol, we employ Chamfer distance (CD) measure, which is defined as:

$$CD(\mathcal{S}, \mathcal{T}) = \frac{1}{|\mathcal{S}|} \sum_{x \in \mathcal{S}} \min_{y \in \mathcal{T}} \|x - y\|_2^2 + \frac{1}{|\mathcal{T}|} \sum_{y \in \mathcal{T}} \min_{x \in \mathcal{S}} \|x - y\|_2^2 \quad (1)$$

where  $\mathcal{S}$  and  $\mathcal{T}$  denote source and target point clouds, respectively.

To enable computation without needing a graphics card, we approximated the distance using 9000 points randomly drawn from each point cloud.

## 4 Mesh surface extraction

We employ Python packages `Veko` and `Pvista` to respectively extract isosurfaces and visualize the extracted isosurfaces for volume rendering of the mesh surfaces. Figure 1-5 illustrate reference meshes rendered for volumes 1-5.

## 5 Implementation details

All experiments were conducted in a virtual environment with Python 3.10, Tensorflow 2.12 and Torch 1.13.0. graphical processing unit (GPU) cards explored include NVIDIA Tesla V100, Tesla T4, and P1000-SXM2 (CUDA Version 12.0).

## 6 Results & Discussions

When each ultrasound volume of 1281 slices was down-sampled to an axial resolution of  $256 \times 256$ , inference on the GPU (V100) took less than one minute while on CPU took  $431.2 \pm 0.9$  seconds (7 minutes).

Figure 8-Figure 11 illustrate point and the point cloud extracted for a randomly selected volume and the point cloud from the corresponding reference mesh label.

Figure 12 and 6 illustrate some of the results computed from a random subset of the test volumes (volumes numbered 0, 6-89).

We next present quantitative results. Table 6-6 report the Chamfer distances between the reconstructed and reference mesh surfaces. Table 6 presents trials that failed to converge. Table 6-6 present trials that yielded satisfactory Chamfer distance (below 95.0). In generating these tables, samples from select volumes were used as the training and validation set while a left out volume was used as a test (unseen) sample; these are marked in the tables as ‘(val)’ and ‘(tst)’ to denote validation and test set, respectively.

Model	Settings	Volume 1	Volume 2	Volume 3	Volume 4	Volume 5
-	-	167.6	167.6	167.6	167.6	167.6
R2Unet	BCE	105.8	103.8	102.7	105.8 (val)	118.6 (tst)
SE-Unet	BFCE	107.9	108.1	102.3	104.8 (val)	114.9 (tst)
R2Unet	Hard sigmoid/ 2-1 ( $t=0.32$ )	147.4 (tst)	137.6 (val)	144.4	134.1	149.8

Table 1: Baseline for subsequent evaluations based on Chamfer distance. For reference, treating every voxel position as part of the reconstructed mesh yields the worse possible distance as 167.6. Results in this table were derived from trials using volume 4 and 5 to construct the validation and test sets, respectively.

Based on the results from the trials summarized by these tables, volumes 4 and 5 had the lowest and highest Chamfer distance, respectively. This may be explained by Figure 5, which shows that volume 5 captured a pipe with an object inside, rendering an obstacle to achieving low Chamfer distances between the extracted and the reference point clouds.

Results from Table 6 suggest that the choice on the encoding schemes of the reference mesh labels did not impact performance in significant ways. Results from A suggest that taking the average of contiguous slices maybe superior to taking the solution of the most confident slice.

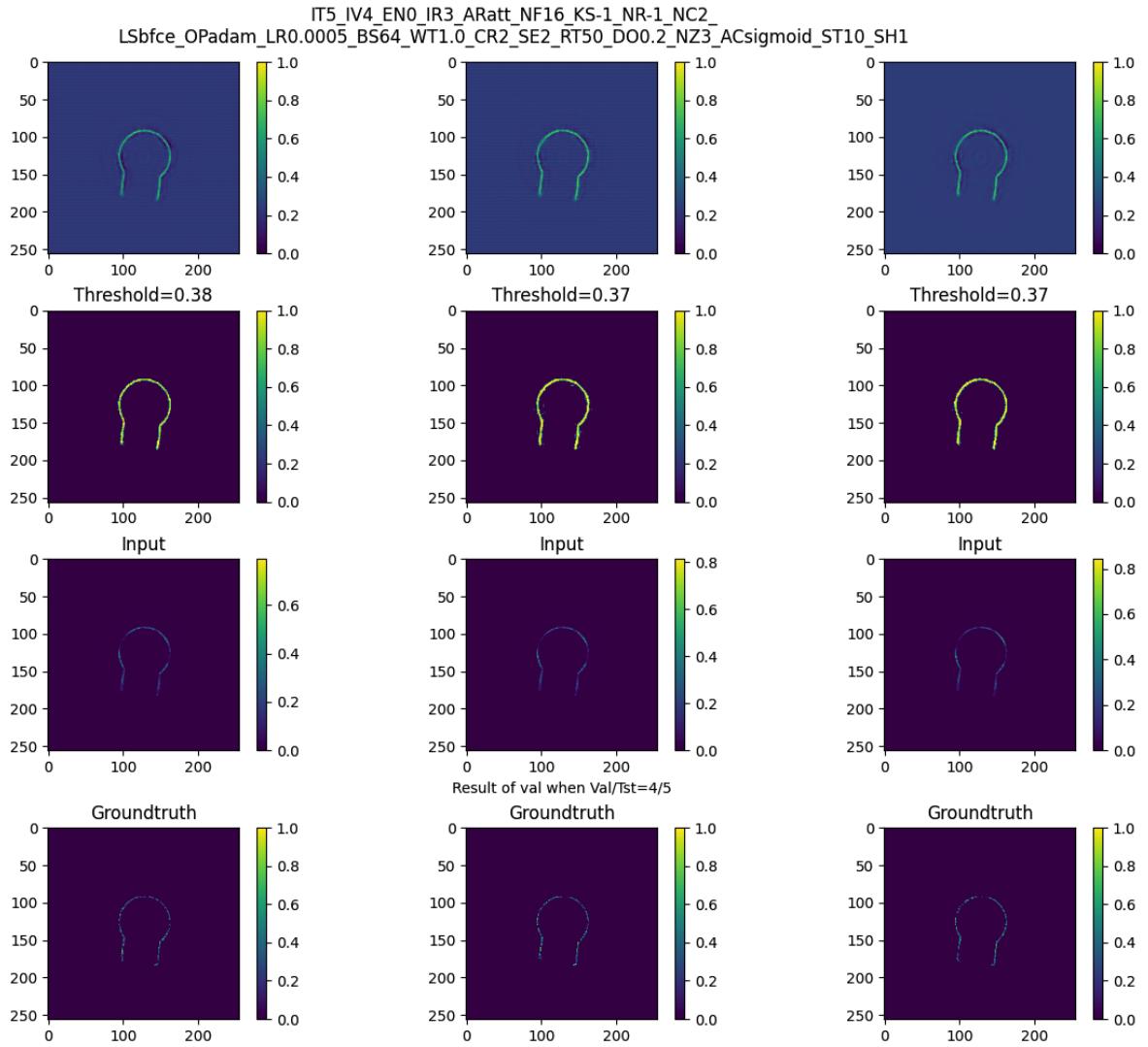


Figure 6: Example 3-slice input to each model. Top to bottom: probabilistic output, thresholded predicted label, input data, and its corresponding reference label.

In summary, the models did not appear to have over-fitted to the training set as including the hardest sample (volume 5) did not lead to reduction in error.

## 7 Conclusion

In this brief note, we explored the feasibility of surface mesh reconstruction via point cloud estimation as an image to mask generation problem. This initial framework opens door to possibility of leveraging (pretrained) deep and wide networks published in the wild. The source code developed during the course of this experimental prototyping period is posted at <https://github.com/lisatwyw/smrvis>. We hope the research communities will find this quick prototype consisting of a few Python scripts approachable.

Model	Settings	$t$	Volume 1	Volume 2	Volume 3	Volume 4	Volume 5
Att-Unet	Trained on 1,2,3	0.19	84.0	75.1	75.9	68.4 (val)	94.0 (tst)
		0.28	82.1	73.4	72.5	66.3 (val)	93.2 (tst)
		0.37	83.2	73.7	70.5	66.1 (val)	93.2 (tst)
R2-Unet	Trained on 3,4,5	0.15	81.5 (tst)	73.3 (val)	73.3	65.1	91.8
		0.22	77.9 (tst)	73.1 (val)	72.3	65.3	91.8
	Trained on 1,4,5	0.2	83.1	74.6 (tst)	75.0 (val)	66.4	93.4
	Trained on 1,2,3	0.3	80.4	73.7 (tst)	73.5 (val)	65.8	93.3
		0.4	73.8	73.9 (tst)	72.1 (val)	65.9	93.4
		0.14	82.0	74.8	73.0	65.5 (val)	92.1 (tst)
	Trained on 3,4,5	0.21	81.8	71.4	72.7	65.4 (val)	92.2 (tst)
		0.25	82.1	71.2	72.1	65.6 (val)	92.2 (tst)
SE-Unet	Trained on 3,4,5	0.35	NIL	71.7	71.3	65.3 (val)	91.1 (tst)
		0.2	82.0 (val)	72.7	73.0 (tst)	65.9	92.1
		0.3	80.3 (val)	72.4	72.7 (tst)	65.3	92.3
	Trained on 1,2,3	0.4	78.3 (val)	73.3	71.1 (tst)	65.6	92.8
		0.20	81.8	72.7	71.9	64.7 (val)	92.4 (tst)
		0.30	80.1	72.2	71.2	65.0 (val)	93.0 (tst)
Wnet	Trained on 1,2,5	0.32	79.6	72.3	70.9	65.1 (val)	93.2 (tst)
		0.338	72.8	72.6	72.3 (val)	65.7 (tst)	92.7
		0.375	71.7	72.8	72.1 (val)	65.9 (tst)	93.0
		0.413	69.9	73.1	71.9 (val)	66.1 (tst)	93.4

Table 2: Effects of training set on validation and test volumes. Performance evaluation of different networks based on Chamfer distance. For each model, the best test score selected by the lowest validation error (i.e. lowest Chamfer distance) is highlighted with blue cells.

Model	Settings	threshold	Volume 1	Volume 2	Volume 3	Volume 4	Volume 5
Att-Unet	Saturated mask	0.2	82.1	73.0	72.7 (val)	66.6 (tst)	92.5
		0.3	81.3	72.4	72.3 (val)	66.1 (tst)	92.4
		0.4	78.9	72.5	70.8 (val)	65.7 (tst)	92.7
	Solid mask	0.22	81.1	72.7	72.4 (val)	65.7 (tst)	92.3
		0.33	79.2	72.8	71.6 (val)	65.5 (tst)	92.6
		0.37	78.8	73.1	71.3 (val)	65.5 (tst)	92.6

Table 3: Comparison of the encoding schemes of the reference mesh labels.

## Acknowledgements

We sincerely thank Dark Vision Technologies Inc. for provision of the ultrasound dataset and hosting this exciting challenge. The author also expresses deep gratitude to Tong Tsui Shan and Kim Chuen Tang for their support.

## References

- [1] Rao J, Wang J, Kollmannsberger S, Shi J, Fu H, Rank E. Point cloud-based elastic reverse time migration for ultrasonic imaging of components with vertical surfaces. *Mechanical Systems and Signal Processing*. 2022 Jan 15;163:108144.
- [2] Verykokou S, Ioannidis C. An Overview on Image-Based and Scanner-Based 3D Modeling Technologies. *Sensors*. 2023 Jan;23(2):596.
- [3] Virkkunen I, Koskinen T, Jessen-Juhler O, Rinta-Aho J. Augmented ultrasonic data for machine learning. *Journal of Nondestructive Evaluation*. 2021 Mar;40:1-1.
- [4] Dark Vision Technologies Inc. Nov 2022. The Ultrasound Dataset Challenge. Retrieved Mar 2023 from <https://www.cvpr2023-dl-ultrasound.com/>.
- [5] Tang L, Hamarneh G. SMRFI: Shape matching via registration of vector-valued feature images. In *2008 IEEE Conference on Computer Vision and Pattern Recognition 2008 Jun 23 (pp. 1-8)*. IEEE.
- [6] Gut D, Tabor Z, Szymkowski M, Rozynek M, Kucybala I, Wojciechowski W. Benchmarking of deep architectures for segmentation of medical images. *IEEE Transactions on Medical Imaging*. 2022 Jun 6;41(11):3231-41.

Model	Settings	$t$	Volume 1	Volume 2	Volume 3	Volume 4	Volume 5
Att-Unet	Hard sigmoid	0.20	116.4	107.0	106.2	95.7	114.0
		0.30	81.4	73.8	74.3 (val)	66.4 (tst)	93.0
		0.40	78.5	73.7	73.2 (val)	65.9 (tst)	93.1
R2-Unet	Linear	0.28	76.2	73.0	72.6	65.1	93.8
SE-Unet	Hard sigmoid	0.18	81.9	73.3	74.6 (val)	66.8 (tst)	92.4
		0.27	80.3	73.0	72.9	65.8	92.7
		0.37	81.2	74.3	71.4 (val)	66.1 (tst)	93.4
	tanh	0.26	79.7	73.8	71.5	66.0	93.5
Wnet	Hard sigmoid (BFCE)	0.34	81.6	72.7	69.6	64.3	92.4
		0.37	80.1	72.8	69.1 (val)	64.1 (tst)	92.1
	Hard sigmoid with BCE	0.34	81.6	72.7	69.6	64.3	92.4
		0.37	80.1	72.8	69.1 (val)	64.1 (tst)	92.1

Table 4: Comparisons on the loss and activation functions.

- [7] Kugelman J, Allman J, Read SA, Vincent SJ, Tong J, Kalloniatis M, Chen FK, Collins MJ, Alonso-Caneiro D. A comparison of deep learning U-Net architectures for posterior OCT retinal layer segmentation. *Scientific Reports.* 2022 Sep 1;12(1):14888.
- [8] Lisa YW Tang, Harvey O Coxson, Stephen Lam, Jonathon Leipsic, Roger C Tam, and Don D Sin, “Towards large-scale case-finding: training and validation of residual networks for detection of chronic obstructive pulmonary disease using low-dose CT,” *The Lancet Digital Health*, vol. 2, no. 5, pp. e259–e267, 2020.
- [9] Alom MZ, Hasan M, Yakopcic C, Taha TM, Asari VK. Recurrent residual convolutional neural network based on u-net (r2u-net) for medical image segmentation. *arXiv preprint arXiv:1802.06955.* 2018 Feb 20.
- [10] Sonal Gore, Ashwin Mohan, Prajakta Bhosale, Prajakta Joshi, Ashley George. Brain Tumor Segmentation Using Deep Neural Networks. 2021 Retrieved May 2023 <https://github.com/Brain-Tumor-Segmentation>.
- [11] Szegedy, Christian, Vincent Vanhoucke, Sergey Ioffe, Jon Shlens, and Zbigniew Wojna. ”Rethinking the InceptionNet architecture for computer vision.” In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 2818-2826. 2016.
- [12] He, Kaiming, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. ”Deep residual learning for image recognition.” In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770-778. 2016.
- [13] Zagoruyko, Sergey, and Nikos Komodakis. ”Wide residual networks.” *arXiv preprint arXiv:1605.07146* (2016).
- [14] Iandola, Forrest N., Song Han, Matthew W. Moskewicz, Khalid Ashraf, William J. Dally, and Kurt Keutzer. ”SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and < 0.5 MB model size.” *arXiv preprint arXiv:1602.07360* (2016).

## A More details on the models

Model	Settings	$t$	Volume 1	Volume 2	Volume 3	Volume 4	Volume 5
Att-Unet	Max	0.20	81.0	74.4	73.7	66.6	93.0
		0.30	79.9	74.0	73.3	66.3	93.0
		0.40	79.4	74.0	73.1 (val)	66.3 (tst)	93.2
	Mean	0.40	78.9	72.5	70.8 (val)	65.7 (tst)	92.7
SE-Unet	Max	0.33	79.9	73.4	72.4 (val)	65.8 (tst)	93.1
		0.36	80.4	73.9	71.7 (val)	65.9 (tst)	93.4
	Single slice only	0.33	80.6	73.9	72.0 (val)	65.8 (tst)	93.0
		0.36	81.2	74.4	71.3 (val)	66.3 (tst)	94.2
SE-Unet with tanh	Mean	0.37	81.2	74.3	71.4 (val)	66.1 (tst)	93.4
	Mean	0.23	77.4	73.8	71.8	65.9	93.6
		0.26	77.4	74.0	71.2 (val)	66.0 (tst)	94.8
	Max	0.23	78.2	73.7	72.2	66.0	93.3
		0.26	78.0	73.9	71.8	66.1	93.5
W-net	Single slice only	0.23	79.5	73.7	72.0	65.8	93.5
		0.26	79.6	73.9	71.5 (val)	66.0 (tst)	93.5
		0.338	89.0	80.7	84.2 (val)	74.5 (tst)	98.1
	Max	0.375	81.6	76.5	78.1 (val)	69.5 (tst)	94.8
		0.413	77.5	75.2	75.4 (val)	67.7 (tst)	94.6
W-net	Single slice only	0.338	82.9	77.3	79.4	70.2	95.8
		0.413	77.5	76.5	77.5 (val)	68.6 (tst)	95.2
	Mean (from Table 6)	0.413	69.9	73.1	71.9 (val)	66.1 (tst)	93.4

Table 5: Comparisons on the choice of aggregation scheme. Selection based on the validation set suggests that aggregation by mean achieves distance of 66.1 on the test volume, which is a slight improvement compared to alternative choices (67.7- 68.6) for the Wnet. For the SE-Unet, there is no distinctive impact due to aggregation choices (Chamfer distance ranged from 65.8-66.3).

Model	# of filters/ activation	# of layers	# of trainable params.	Training set	# of epochs	$t$
Att-net	16, linear	190	1,989,767	2,3,4 [1/5]	54	0.31
R2-Uet	16, sigmoid	158	1,999,571	3,4,5 [2/1]	35	0.32
U-net++	8, sigmoid	242	514,219	2,4,5 [3/2]	167	0.41
	16, sigmoid	242	2,050,387	1,2,3 [5/4] (0530)	121	0.39
SE-Unet	16, tanh	161	2,054,355	1,2,5 [4/3]*		
W-net	16, hard sigmoid	168	1,159,091	1,2,3 [5/4]	100	0.34

Table 6:

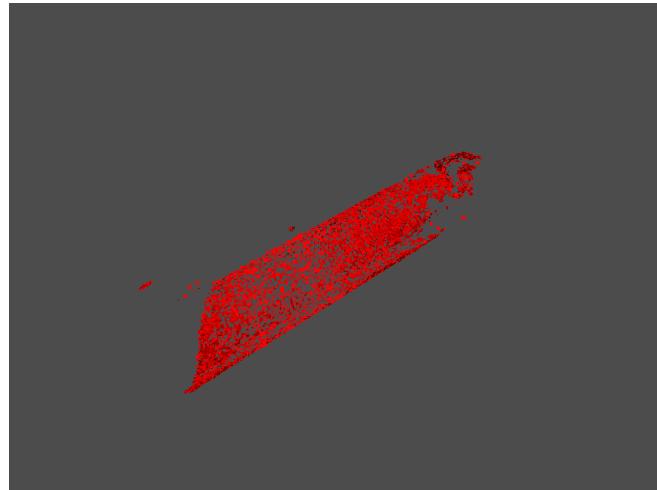


Figure 7: Example of a degenerate case.

## B More visualizations

## C Meta file

```
ObjectType = Image
NDims = 3
BinaryData = True
BinaryDataByteOrderMSB = False
CompressedData = False
TransformMatrix = 1.0 0.0 0.0 0.0 0.0 1.0 0.0 0.0 0.0 0.0 1.0
Offset = 0.0 0.0 0.0
CenterOfRotation = 0.0 0.0 0.0
AnatomicalOrientation = RAI
ElementSpacing = 0.49479 0.49479 0.3125
DimSize = 768 768 1280
ElementType = MET USHORT
ElementDataFile = scan_001.raw
```

## D Training progress

Figure 18 and 19 show the progress of two randomly selected trials involving a R2Unet and SE-Unet model, respectively.

IT4\_IV4\_EN0\_IR3\_ARr2unet\_NF16\_KS-1\_NR2\_NC2  
LSbfce\_OPadam\_LR0.0005\_BS64\_WT1.0\_CR3\_SE2\_RT50\_D00.2\_NZ3\_ACsigmoid\_ST10\_SH1

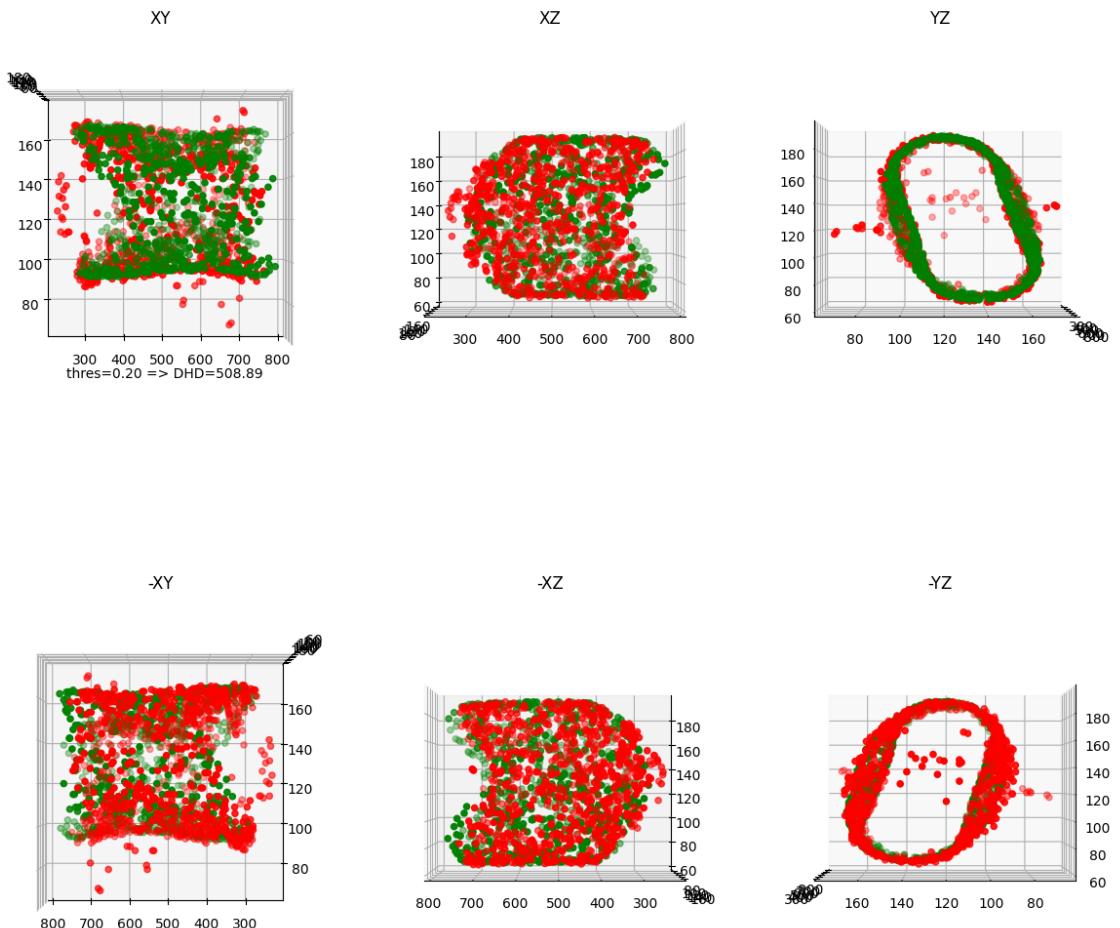


Figure 8: Results for volume 1. Point cloud of the reference mesh label (in green) and the point cloud extracted by the proposed framework using a R2-Unet (in red).

Model	(f to compute t, AG)					
Att	77.5 (tst)	72.9	73.4	65.4	92.8 (val)	(0.9, 0)
	77.1	71.9	73.4	64.5	92.1	(1, 0)
	79.3	73.7	73.3	66.2	93.0	(0.9, 1)
	78.9	73.3	73.3	65.9	92.7	(1, 1)
	76.9	73.2	73.4	65.9	92.9	(0.9, 2)
	75.7	72.7	73.4	65.7	93.0	(1, 2)
R2Unet	81.7	70.9	72.4	65.6 (val/tst)	92.3	(0.9, 0)
	82.8	71.1	72.1	65.6	92.1	(1, 0)
	81.7	71.7	72.5	65.4	92.5	(0.9, 1)
	81.8	72.1	72.2	65.6	92.6	(1, 1)
	82.3	70.5	72.9	65.4	92.1	(0.9, 2)
	83.1	70.8	72.7	65.4	91.9	(1, 2)
SE-Unet	77.4	73.8	71.8 (val)	65.9 (tst)	93.6	(0.9, 0)
	77.4	74.0	71.2 (val)	66.0 (tst)	94.8	(1, 0)
	78.2	73.7	72.2	66.0	93.3	(0.9, 1)
	78.0	73.9	71.8	66.1	93.5	(1, 1)
	79.5	73.7	72.0	65.8	93.5	(0.9, 2)
	79.6	73.9	71.5	66.0	93.5	(1, 2)
Unet++; NF8/SE2/IR3	77.4	72.8	72.6 (tst)	65.1 (val)	92.4	(0.9, 0)
	75.4	73.0	72.4	65.5	92.6	(1, 0)
	78.5	72.9	72.6	65.3	92.5	(0.9, 1)
	76.3	73.1	72.4	65.4	92.6	(1, 1)
	78.3	72.8	72.6 (tst)	64.9 (val)	92.4	(0.9, 2)
	76.1	72.9	72.5	65.4	92.5	(1, 2)
Unet++; NF16/SE1/IR2/RT=60	82.6	74.0	70.9	66.3	93.4	(0.9, 0)
	85.4	74.4	67.9	66.9	93.4	(1, 0)
	82.4	74.0	71.5	66.1	93.5	(0.9, 1)
	83.8	74.4	69.6	66.7	93.4	(1, 1)
	82.1	74.0	71.7	66.0 (val)	93.3 (tst)	(0.9, 2)
	83.1	73.9	70.6	66.5	93.5	(1, 2)
	82.4	73.4	72.7	65.9	93.2	(0.7, 0)
	82.3	73.7	72.1	65.9	93.3	(0.8, 0)
	82.3	73.4	72.9	65.8	93.2	(0.7, 1)
	82.3	73.7	72.3	65.9	93.5	(0.8, 1)
	82.7	73.4	72.8	65.9	93.2	(0.7, 2)
	82.4	73.7	72.3	65.8	93.3	(0.8, 2)
Att-Unet; SE=1; IR=2; RT=45	80.5	73.9	72.0	65.9 (val)	93.7 (tst)	(0.9, 0)
	81.4	74.4	70.6	66.5	94.9	(1, 0)
	80.5	73.9	72.5	66.0	93.4	(0.9, 1)
	80.6	74.3	71.4	66.3	94.7	(1, 1)
	81.9	74.1	71.8	66.1	93.5	(0.9, 2)
	82.0	74.5	70.9	66.4	93.8	(1, 2)
Wnet	82.3	72.4	70.2	64.4 (val)	92.8 (tst)	(0.9, 0)
	81.9	72.4	69.7	63.9 (val)	92.7 (tst)	(1, 0)
	82.6	72.6	70.3	64.7	92.8	(0.9, 1)
	82.4	72.6	69.9	64.5	92.7	(1, 1)

Table 7: Evaluation by Chamfer distance computed between the reconstructed and reference point clouds extracted from each ultrasound volume.

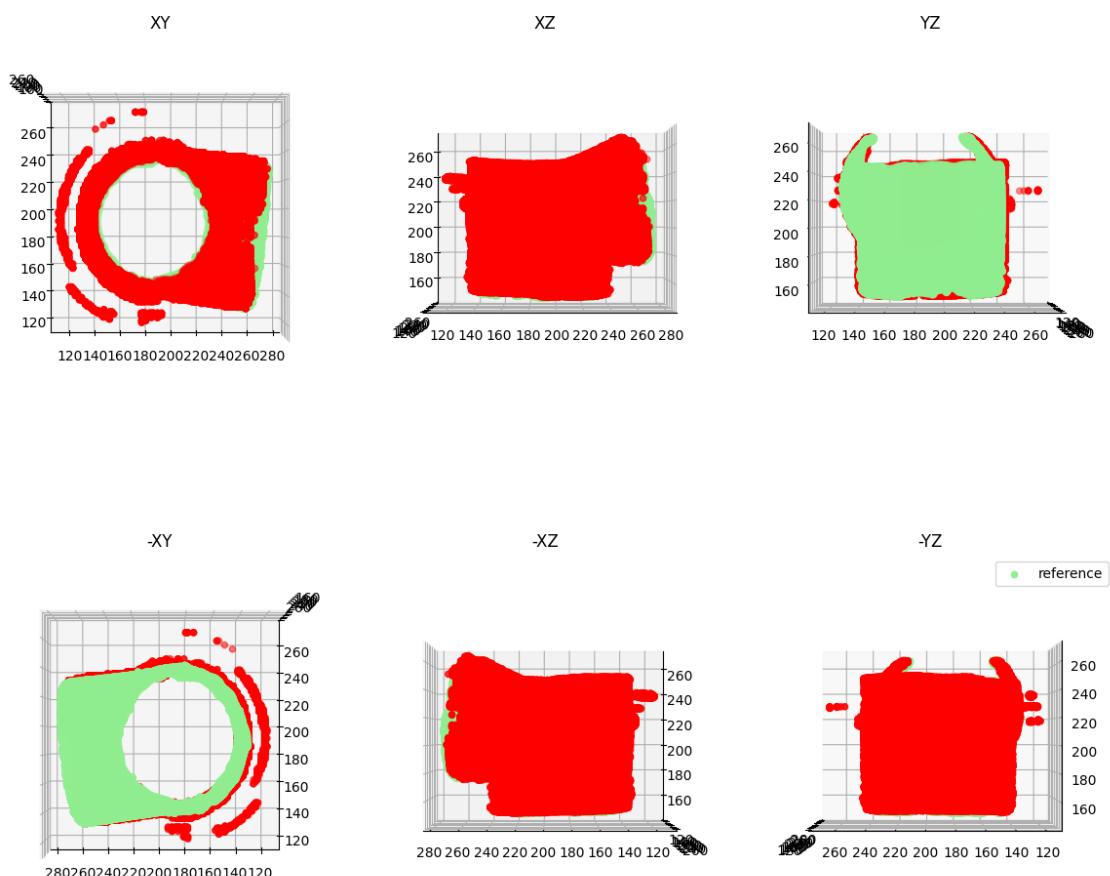


Figure 9: Results for volume 3.

IT5\_IV4\_EN0\_IR3\_Arr2unet\_NF16\_KS-1\_NR2\_NC2  
LSbfce\_OPadam\_LR0.0005\_BS64\_WT1.0\_CR3\_SE2\_RT50\_D00.2\_NZ3\_ACsigmoid\_ST10\_SH1

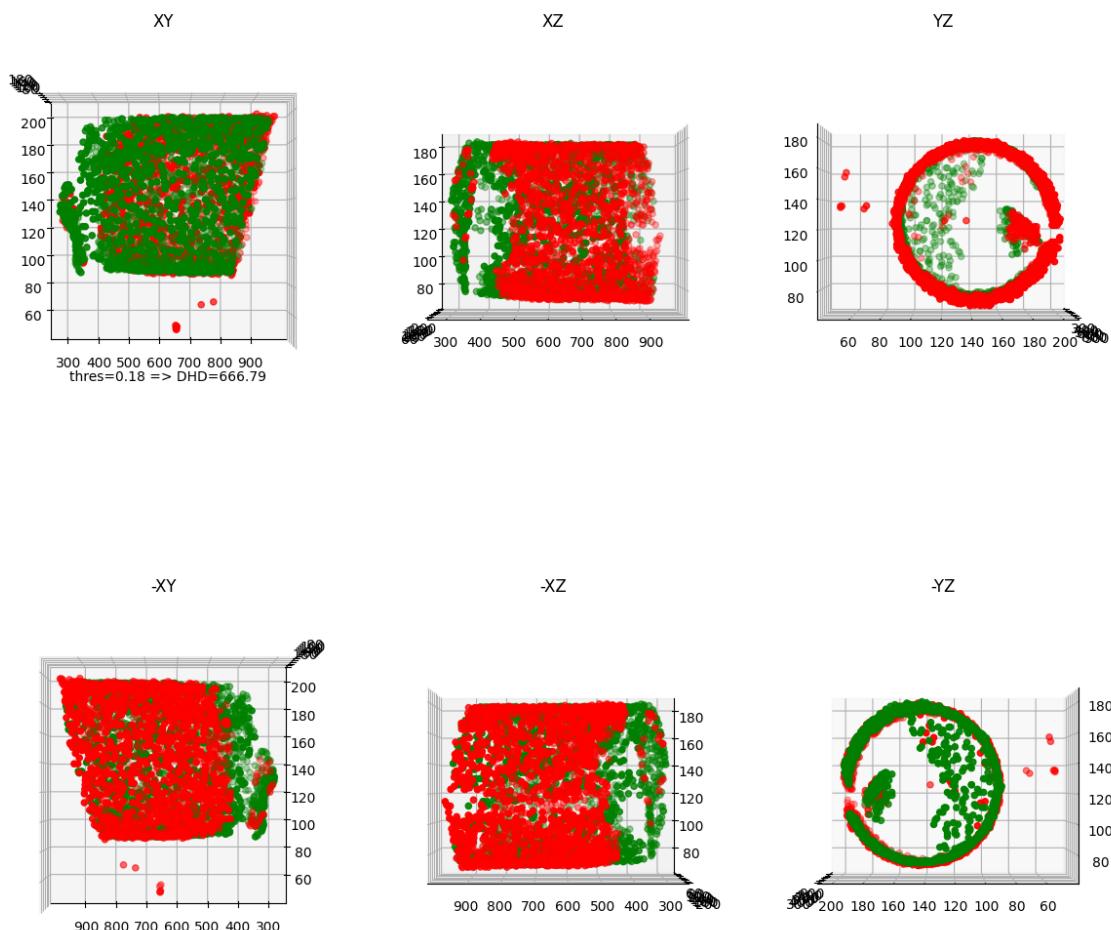


Figure 10: Results for volume 4 using an SE-Unet.

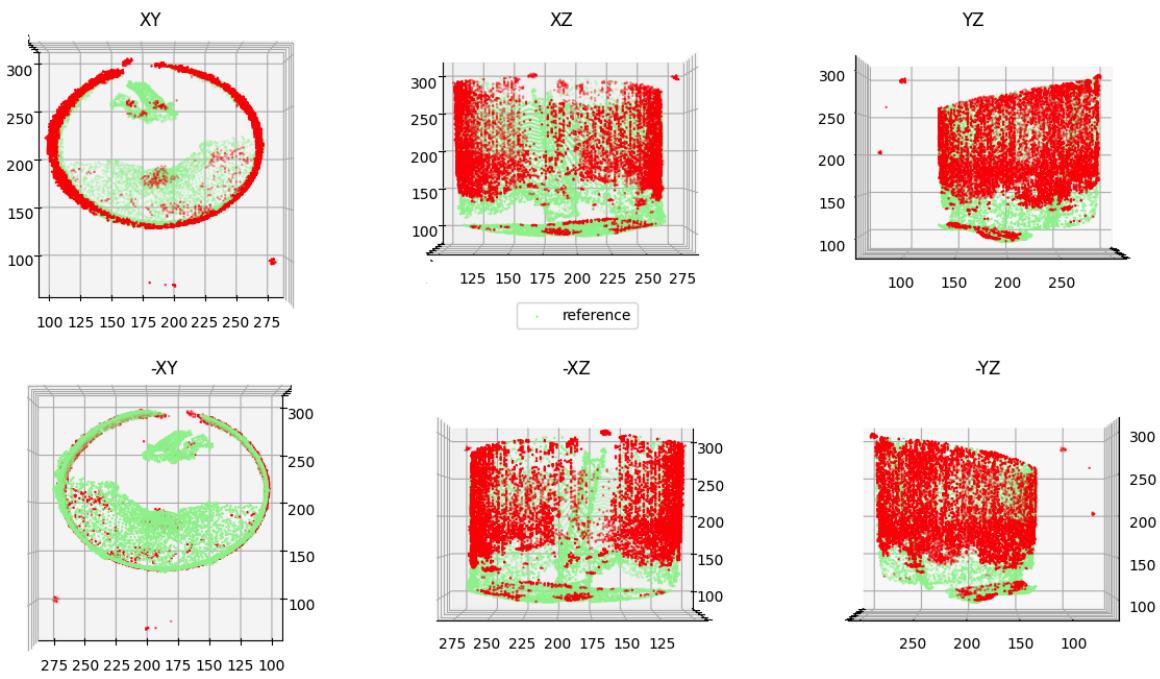


Figure 11: Results from volume 5. Point cloud of the reference mesh label (in green) and the point cloud extracted by the proposed framework using a R2-Unet (in red).

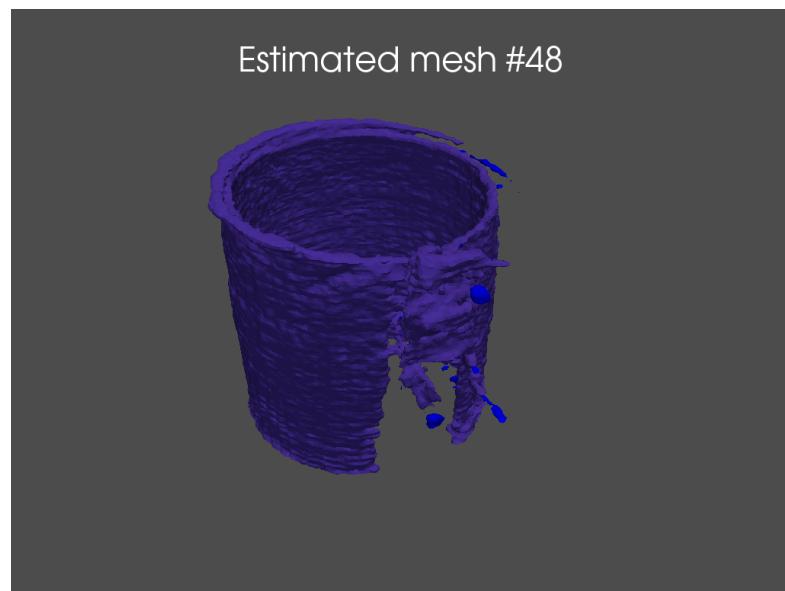


Figure 12: Isosurface rendering of the points extracted for test volume 48 (whose reference label not provided to challenge participants).

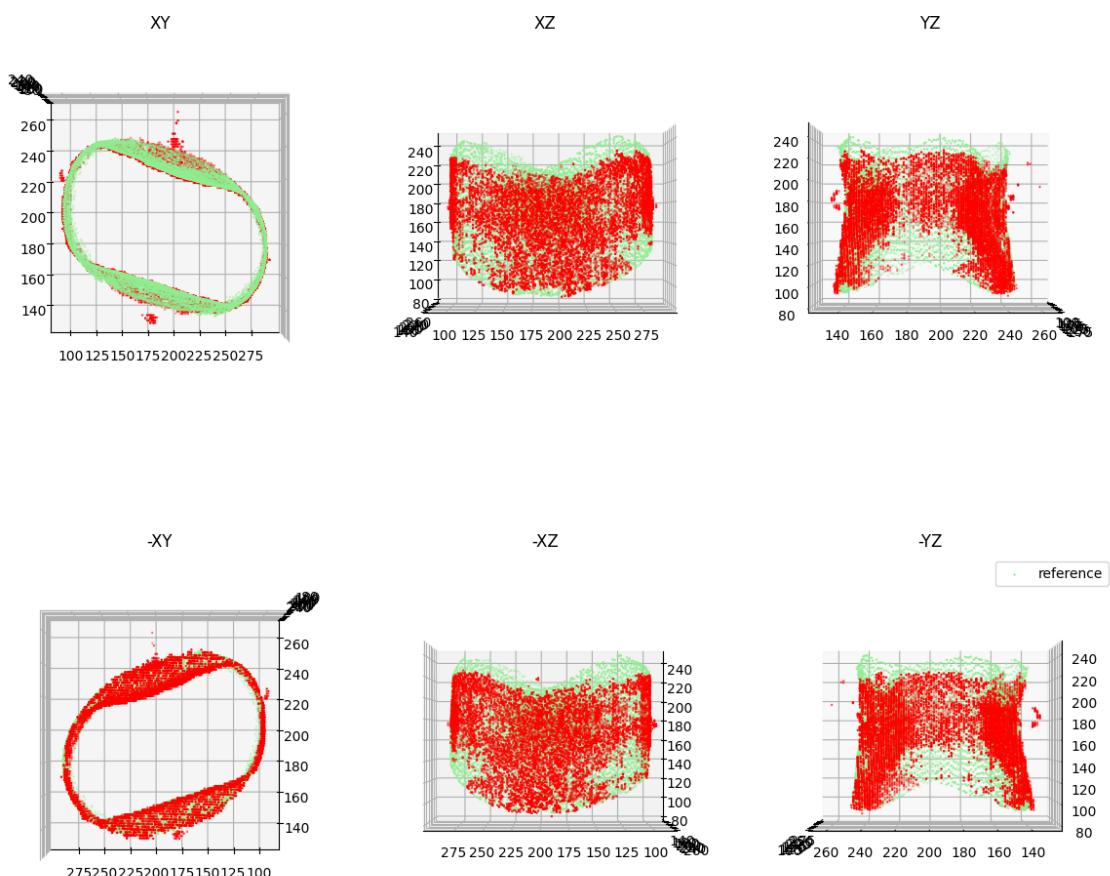


Figure 13: Results for volume 1 by SE-Unet.

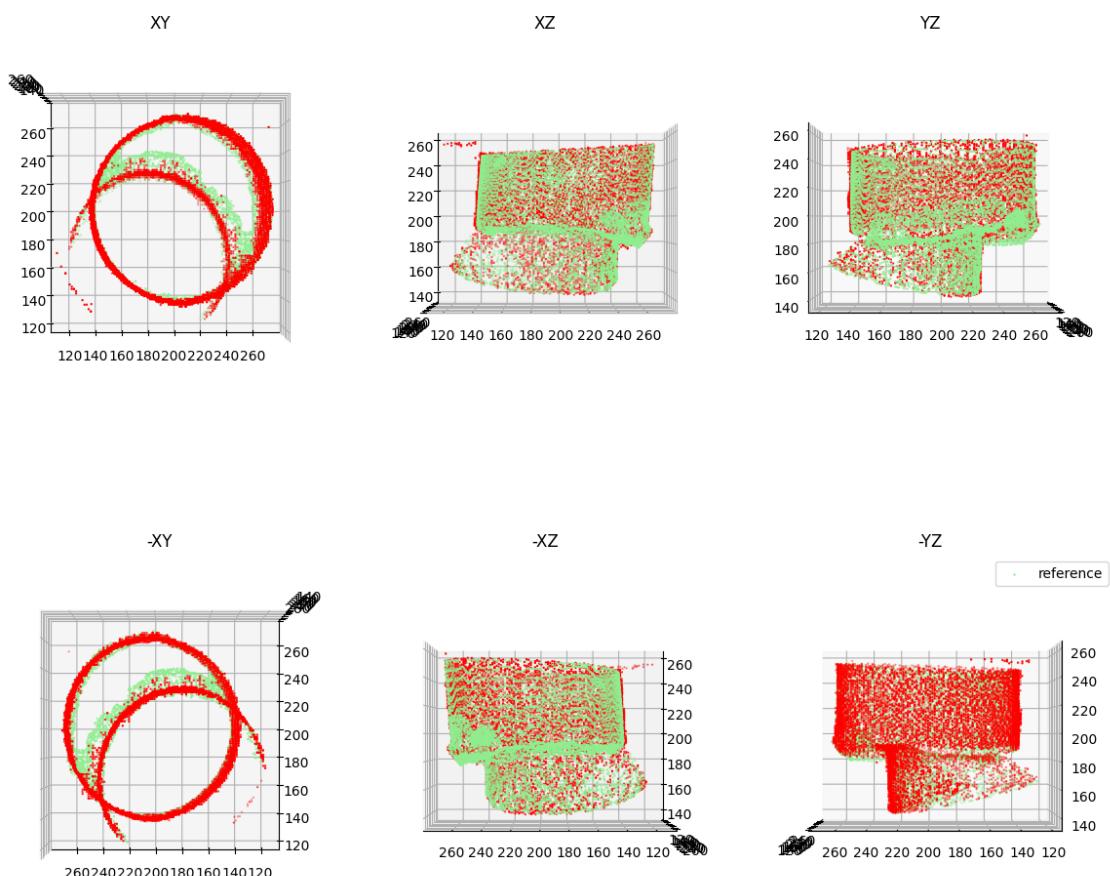


Figure 14: Results for volume 2 using an SE-Unet.

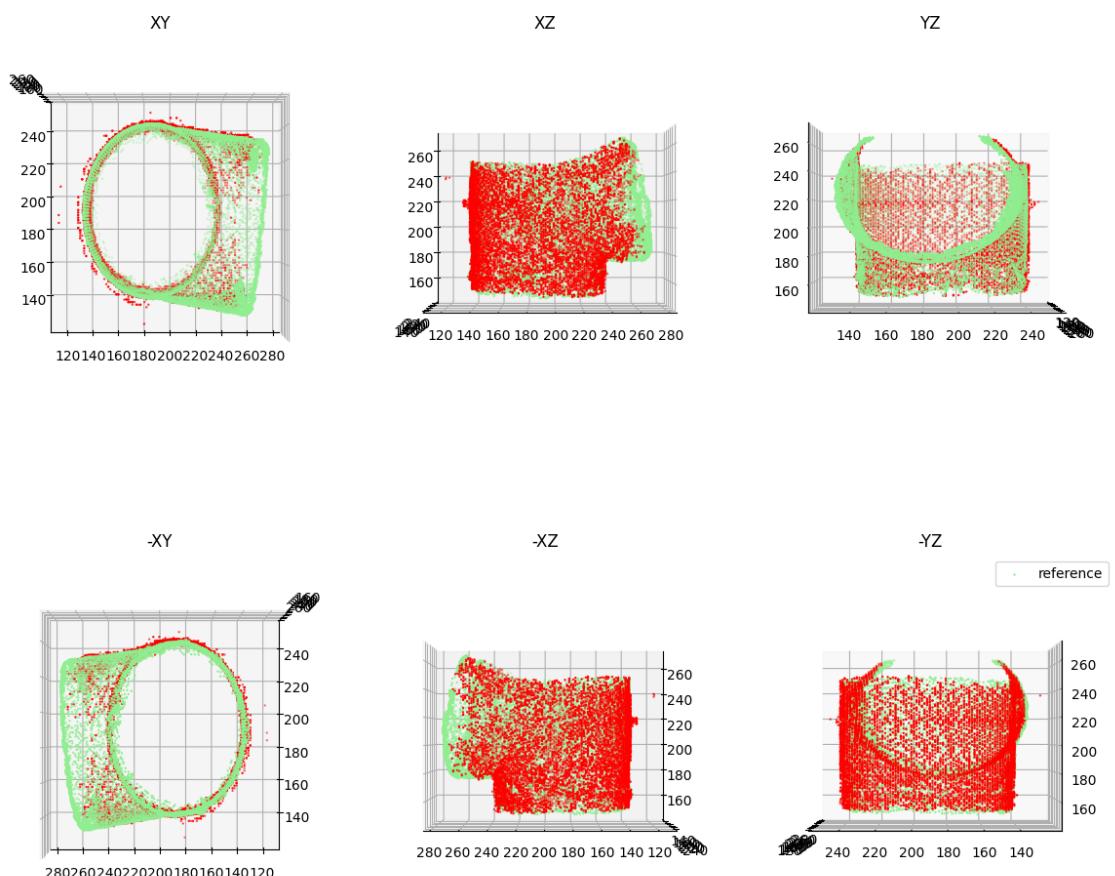


Figure 15: Results for volume 3 using an SE-Unet.

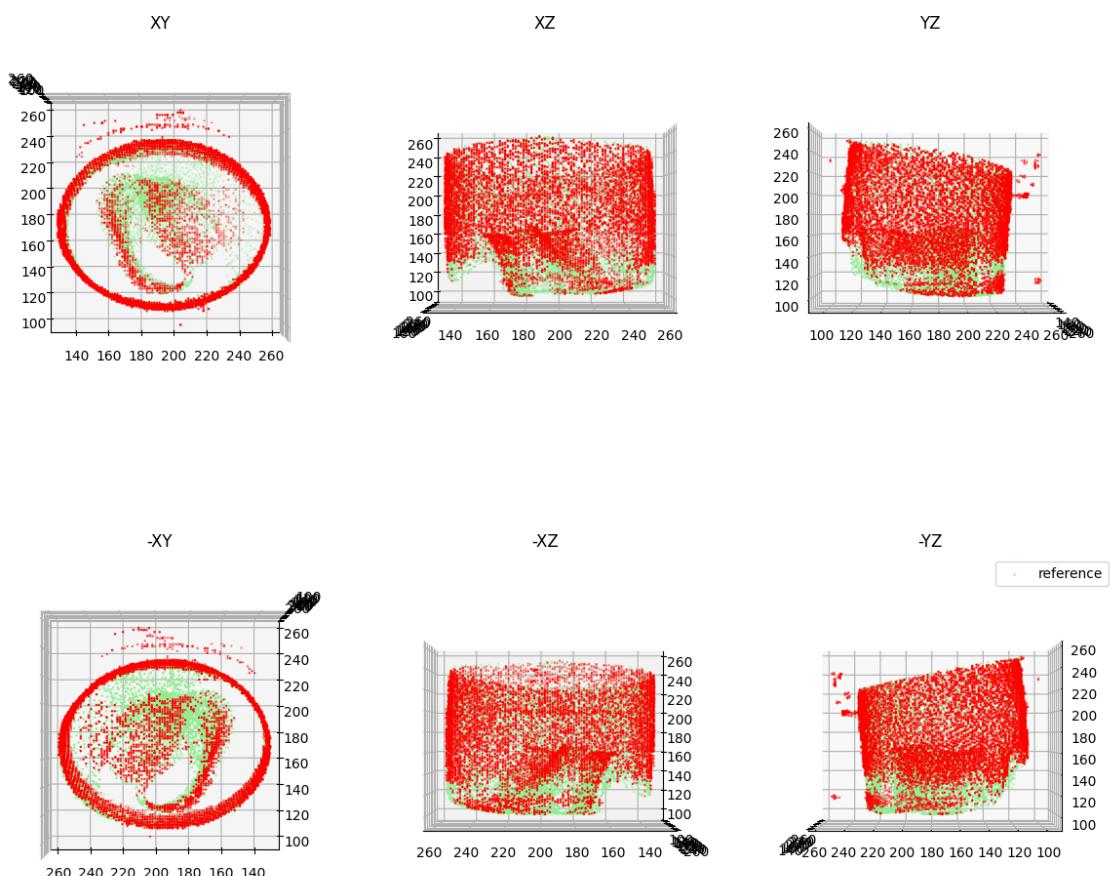


Figure 16: Results for volume 4 using an SE-Unet.

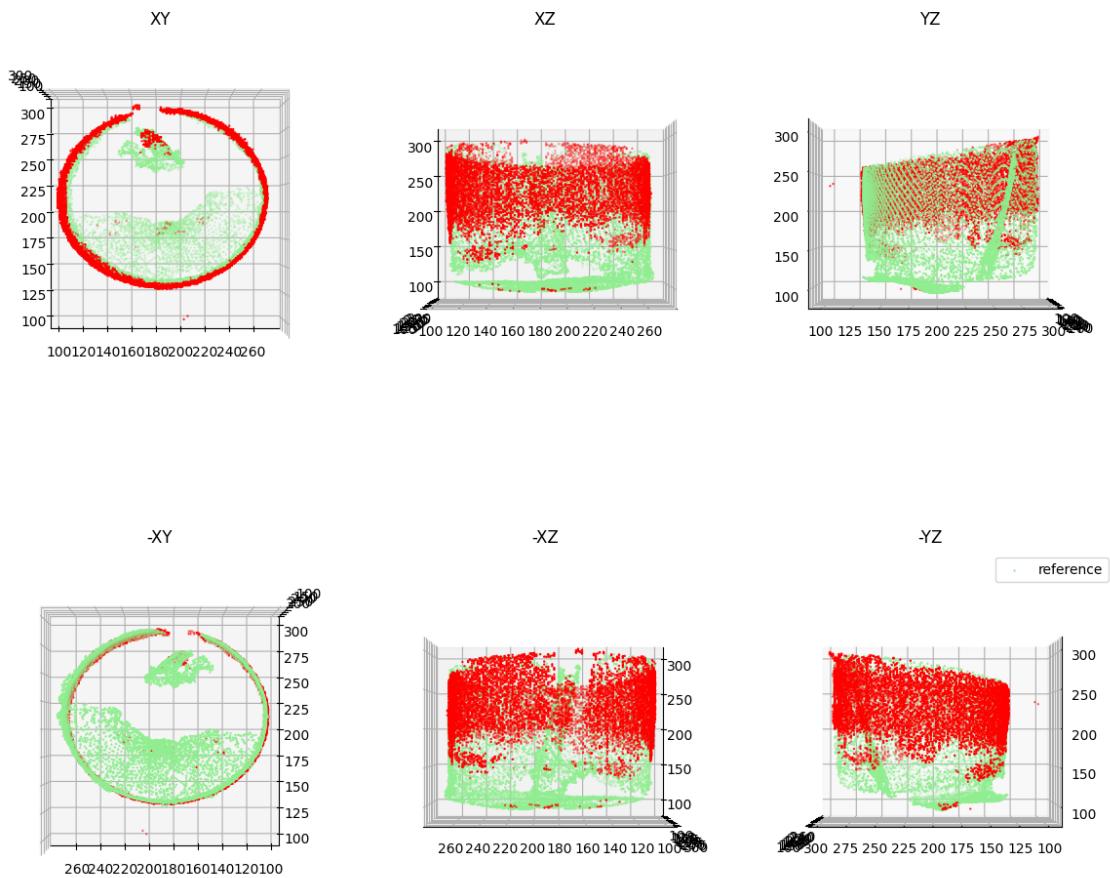


Figure 17: Results for volume 5 using an SE-Unet.

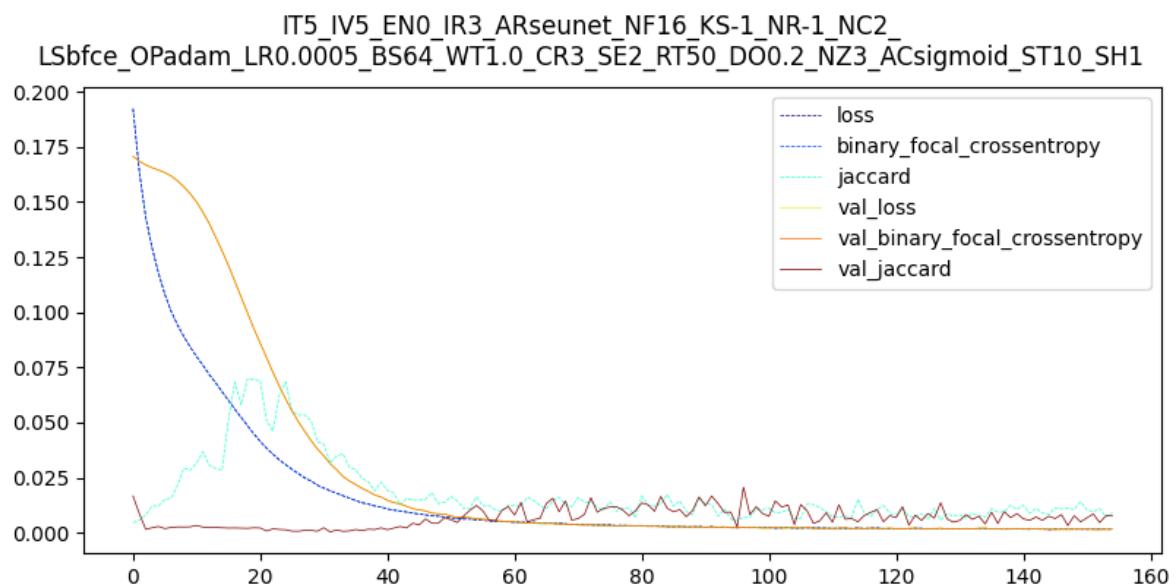


Figure 18: Training progress of a (randomly selected) trial.

IT5\_IV4\_EN0\_IR3\_ARseunet\_NF16\_KS-1\_NR-1\_NC2  
ze\_OPadamw\_LR0.0010\_BS64\_WT1.0\_CR1\_SE2\_RT50\_DO0.2\_NZ3\_ACsigmoid

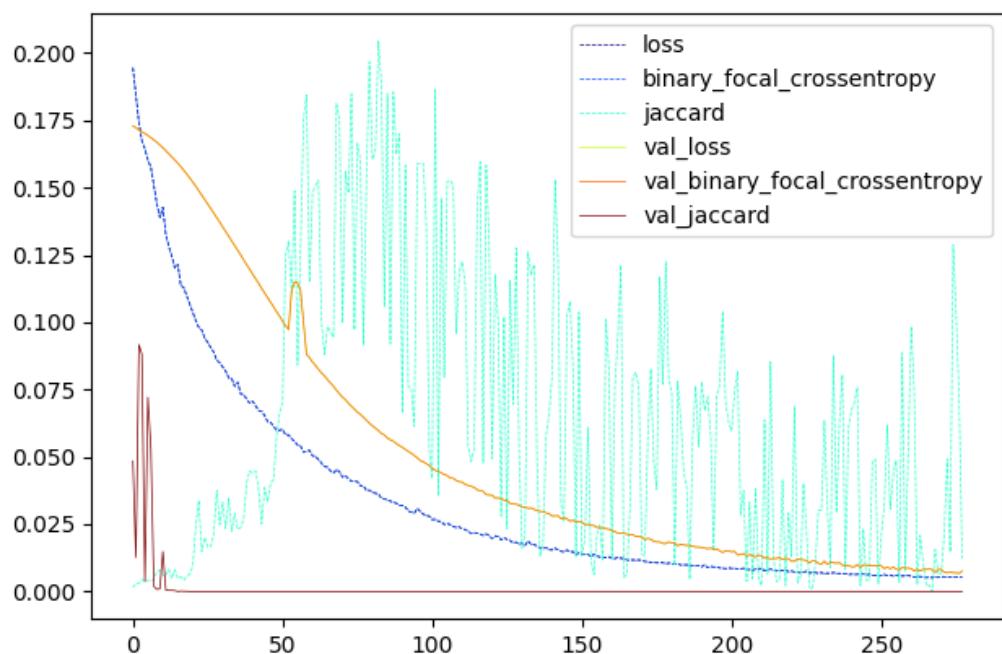


Figure 19: Training progress of a (randomly selected) trial.