Coronary Ostia Localization Using Residual U-Net with Heatmap Matching and 3D DSNT

Milosz Gajowczyk[*,a,b], Patryk Rygiel[*,a,b], Piotr Grodek[a], Adrian Korbecki[a,c], Michal Sobanski[a,c], Przemyslaw Podgorski[a,c], Tomasz Konopczynski[a],

[a] Hemolens Diagnostics Ltd., Wroclaw, Poland, [b] Wroclaw University of Science and Technology
[c] Wroclaw Medical University [*] Equal contribution

HEMOLENS

Wrocław University of Science and Technology

Introduction

Objective: Localization of coronary ostia landmarks in Computed Tomography Angiography (CTA) to facilitate various automatic diagnostic procedures.

Challenge: The anatomical differences between patients and differences in image acquisition.

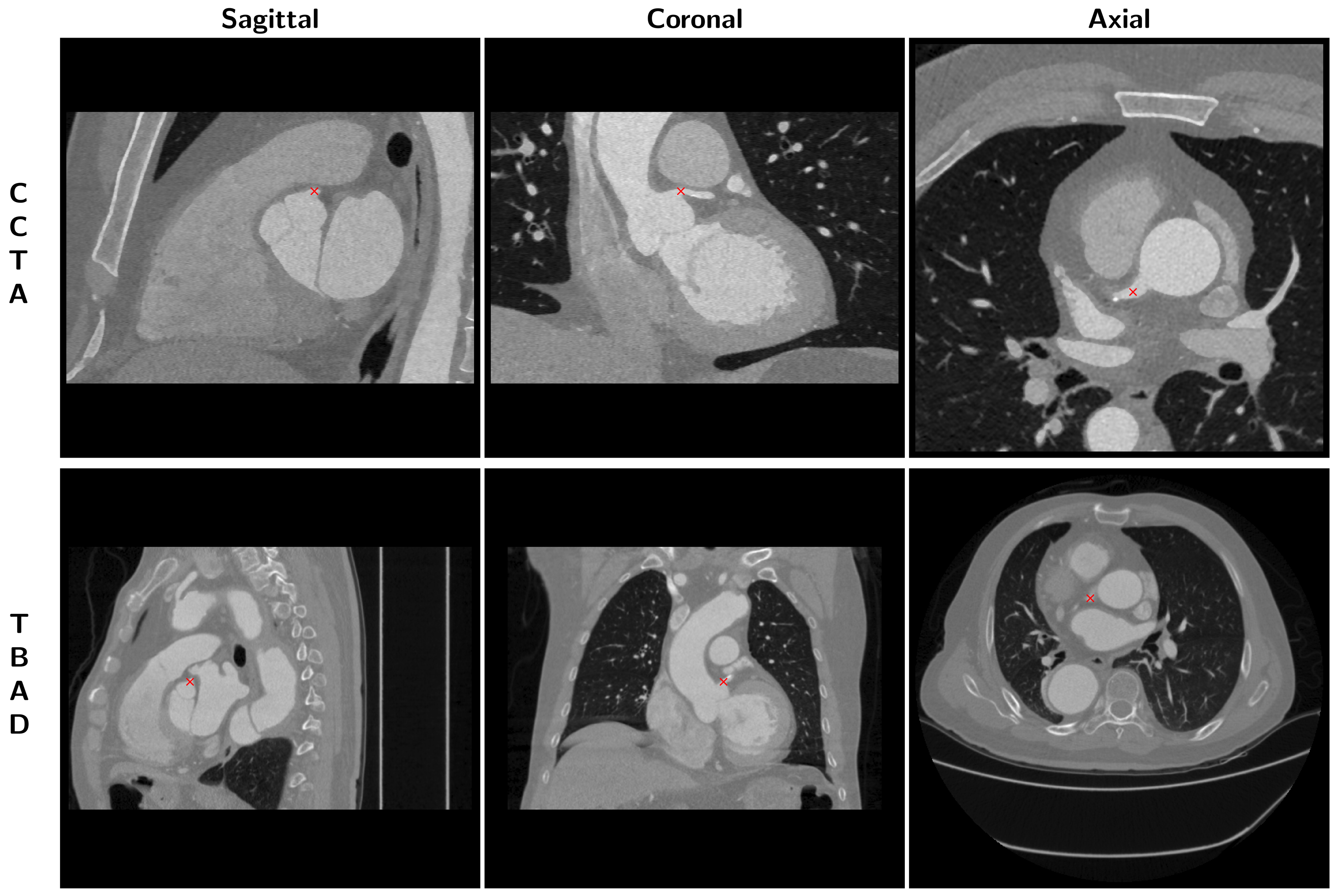
Contribution:

- We propose one-step method of coronary ostia landmark localization that utilizes a residual U-Net with heatmap matching and 3D Differentiable Spatial to Numerical Transform (DSNT).
- We evaluate the method on two datasets a Coronary Computed Tomography Angiography (CCTA) dataset containing 201 scans and a publicly available ImageTBAD dataset containing 77 CTA scans annotated with coronary ostia landmarks.
- We extend ImageTBAD dataset with coronary ostium annotations and share these annotations with the public for further research https://www.synapse.org/#!Synapse:syn35789568/wiki/.

Datasets

The experiments were performed on two datasets. CTA ImageTBAD dataset is significantly lower compared to the CCTA dataset.Despite that, we decide to use it, as it is the only publicly available CT dataset known to the authors with good enough quality for the task of ostia localization. Datasets' details are given in the table below.

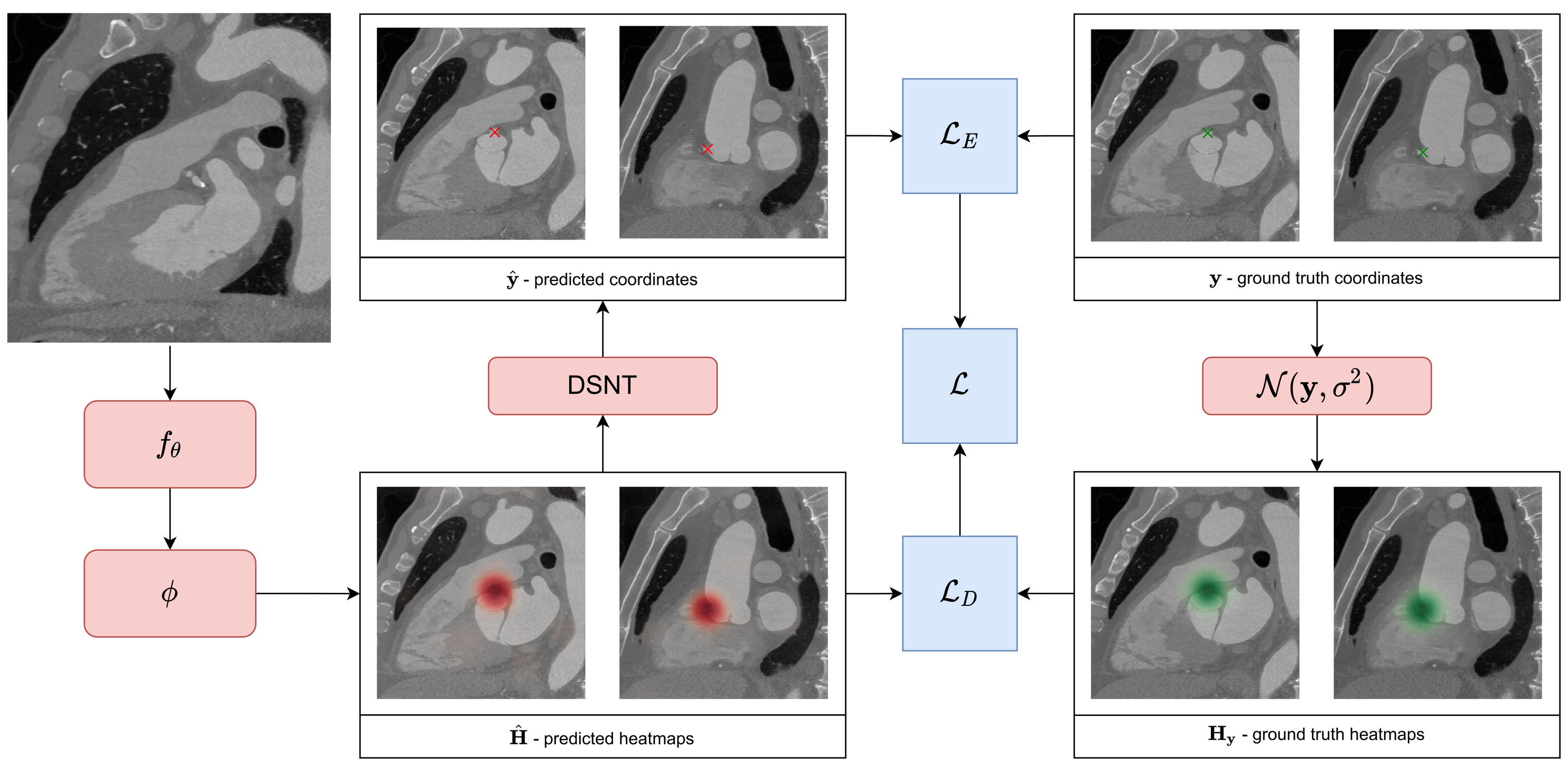
	CCTA (private)	TBAD (public)
# volumes	201	100 (77)
# females	87	31
# males	114	69
# medical centers	8	1
# scanner models	7	2
average patient's age	69.3 ± 10.8 years	52.5 ± 11.3 years
volume size	$512 \times 512 \times (130 - 420)$	$512 \times 512 \times (135 - 416)$



Example slices from datasets with ostium annotation highlighted in red

Method diagram

The input volume is processed by a residual U-Net f_θ and a normalization function ϕ to predict heatmaps $\widehat{\mathbf{H}}$. These heatmaps are then transformed by a DSNT layer to coordinates predictions $\widehat{\mathbf{y}}$. The loss $\mathcal{L}$ consists of the Euclidean distance $\mathcal{L}_E$ between predicted coordinates $\widehat{\mathbf{y}}$ and ground-truth coordinates $\mathbf{y}$, and a divergence $\mathcal{L}_D$ between predicted heatmaps $\widehat{\mathbf{H}}$ and ground-truth heatmaps $\mathbf{H}_\mathbf{y}$. Predicted and ground-truth localization and heatmaps are highlighted in red and green, respectively.



DSNT

The Differentiable Spatial to Numerical Transform (DSNT) is a fully differentiable and non-trainable layer. Due to this fact, numerical coordinates can be obtained during training and can be used in a learning criterion since backpropagation does pass via DSNT. We define its 3D extension as follows:

$$\mathrm{DSNT}(\widehat{H}) = \left[\langle \widehat{H}, X \rangle_F \quad \langle \widehat{H}, Y \rangle_F \quad \langle \widehat{H}, Z \rangle_F \right],$$

where $\widehat{H} = \phi(H)$ is a normalized heatmap H output from the backbone $f_\Theta(\mathbf{x})$ by a normalization function $\phi(\cdot)$, $\langle \cdot, \cdot \rangle_F$ represents Frobenius inner product and X, Y and Z are grids in a form of $n \times m \times l$ matrices used to calculate expected values of landmarks' coordinate and are defined as:

$$X_{i,\,j,\,k} = \frac{2j - (n+1)}{n}, \quad Y_{i,\,j,\,k} = \frac{2i - (m+1)}{m}, \quad Z_{i,\,j,\,k} = \frac{2k - (l+1)}{l}.$$

Loss function

During training we utilize the Euclidean metric $\mathcal{L}_E$ between the network output $\widehat{\mathbf{y}} = \mathrm{DSNT}(\widehat{\mathbf{H}})$ and the ground-truth $\mathbf{y}$. To achieve better results we use the Jensen-Shannon divergence regularization $\mathcal{L}_D$ as a term to calculate, and minimize the divergence between heatmap prediction $\widehat{\mathbf{H}}$ and appropriate target normal distribution $\mathcal{N}(\mathbf{y}, \sigma^2)$. The complete loss is defined as follows:

$$\mathcal{L}(\widehat{\mathbf{H}},\ \mathbf{y}) = \mathcal{L}_E\big(\mathrm{DSNT}(\widehat{\mathbf{H}}),\ \mathbf{y}\big) + \lambda \mathcal{L}_D(\widehat{\mathbf{H}},\ \mathbf{y}),$$

with λ being a regularization coefficient hyper-parameter controlling the strength of the regularization. The Euclidean loss $\mathcal{L}_E$ is defined as:

$$\mathcal{L}_E(\widehat{\mathbf{y}}, \mathbf{y}) = \|\widehat{\mathbf{y}} - \mathbf{y}\|_2,$$

and the Jensen-Shannon divergence regularization $\mathcal{L}_D$ is

$$\mathcal{L}_D(\widehat{\mathbf{H}}, \mathbf{y}) = D_{JS}(\widehat{\mathbf{H}} \parallel \mathcal{N}(\mathbf{y}, \sigma^2)),$$

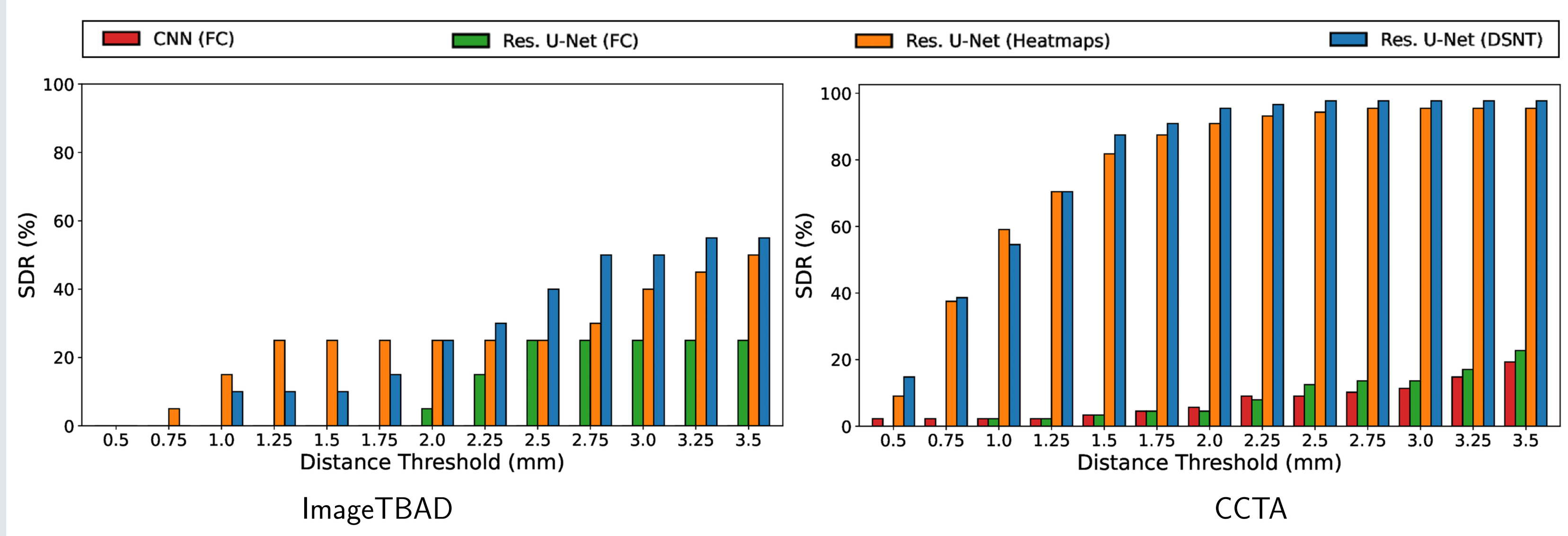
where $D_{JS}(\cdot||\cdot)$ is the Jensen-Shannon divergence.

Quantitative results

Comparison of the Euclidean distance error (mm) between proposed method and other approaches on test volumes. Best results are bolded.

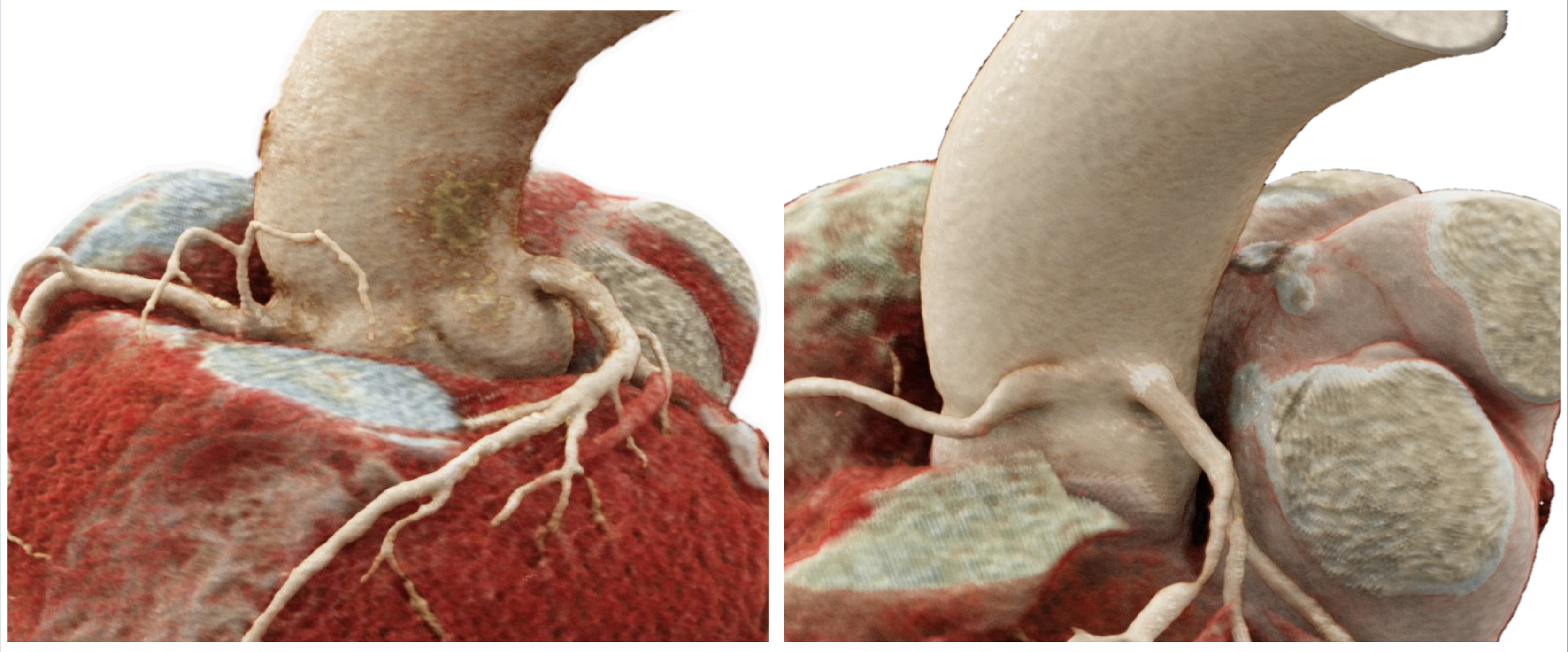
Model	Landmark	CCTA dataset				ImageTBAD dataset			
		Median	Mean	Std	IQR	Median	Mean	Std	IQR
CNN (FC)	Left CA ostium	6.70	6.84	2.98	3.82	8.38	9.43	4.20	4.98
	Right CA ostium	6.83	7.22	3.51	4.97	13.65	13.06	5.40	5.66
Res. U-Net (FC)	Left CA ostium	5.25	5.56	2.48	3.01	7.42	8.00	4.17	5.40
	Right CA ostium	5.60	6.08	2.99	4.43	11.91	12.12	6.10	8.45
Res. U-Net (Heatmaps)	Left CA ostium	1.21	1.45	1.23	0.76	3.63	3.85	1.44	2.00
	Right CA ostium	1.03	1.38	1.33	0.91	3.65	3.74	2.25	3.48
Res. U-Net (DSNT)	Left CA ostium	**1.14**	**1.18**	0.56	0.74	**3.48**	**3.49**	1.42	1.66
	Right CA ostium	**0.98**	**1.29**	1.16	0.82	**2.97**	**3.54**	1.88	2.71

Success detection rate



Success Detection Rate (SDR) is defined as the percentage of predicted landmarks which have the Euclidean distance to the reference landmark below the given threshold. For the clinically accepted SDR range of 3.5mm, our method outperforms other approaches on the CCTA dataset achieving 97.73% SDR at 3.5 mm. On the ImageTBAD dataset, which consist of more challenging volumes, all methods perform much worse with our method achieving 55% SDR on the accepted range of 3.5 mm.

Discussion



Standard anatomy

Anatomical outlier

One of the main challenges are anatomical differences between patients

tomasz.konopczynski@hemolens.eu; milosz.gajowczyk@hemolens.eu; patryk.rygiel@hemolens.eu