Robust Teeth Detection in 3D Dental Scans by Automated Multi-View Landmarking

Tibor Kubík[1] and Michal Španěl[1,2]

[1]Department of Computer Graphics and Multimedia
Faculty of Information Technology, Brno University of Technology

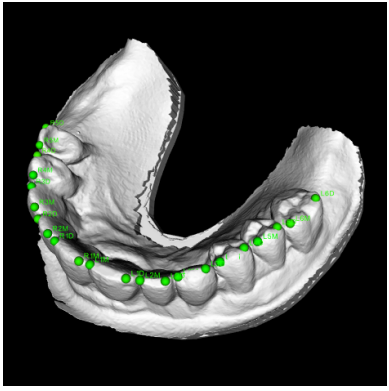
[2]TESCAN 3DIM, Brno, Czech Republic

February 3, 2022

Motivation

Automation of time-consuming step in digital orthodontics.

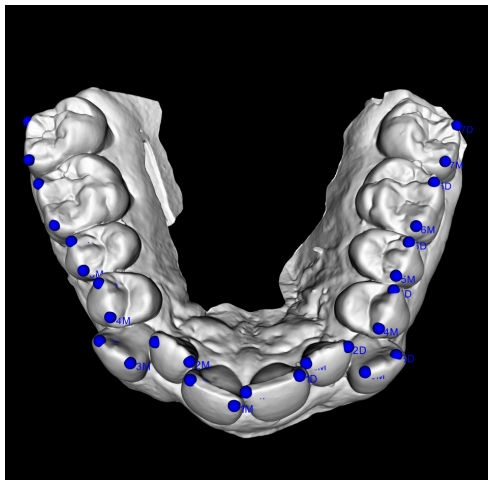
Unannotated dental cast



Arch with manual annotations

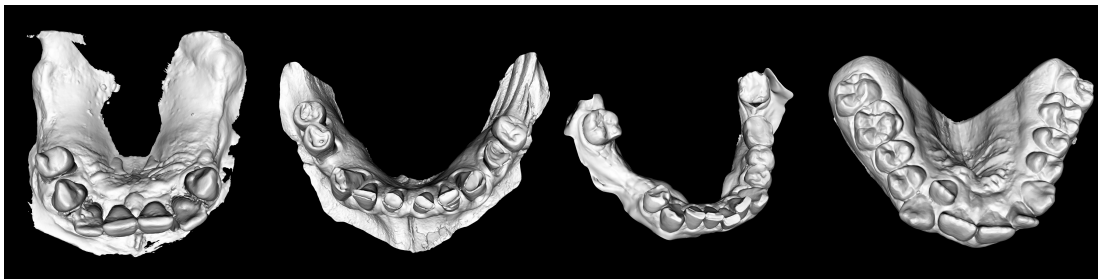
Task Definition

- Automatic DL-based teeth detection in intraoral scans.
- Detection by landmark localization.
- Low computational time (seconds) on CPU.
- Robustness to missing teeth.

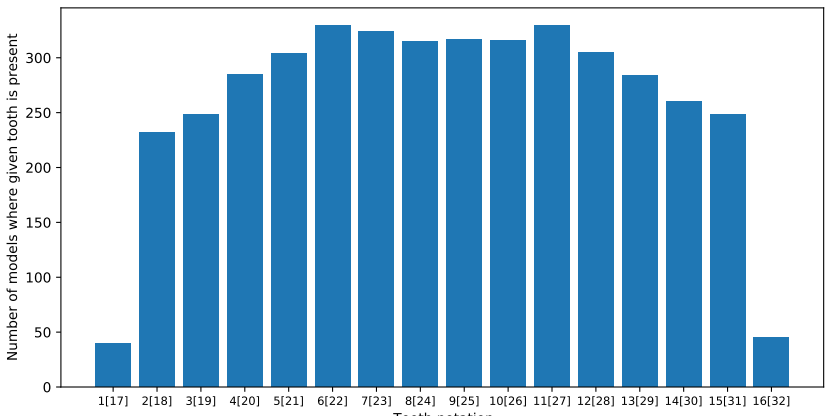


Challenging Real-world Dataset

337 3D dental scans of out of which **90** scans were used for testing.
Each of the used scans is an orthodontic case of real patient – patients with anomalies need the treatment.

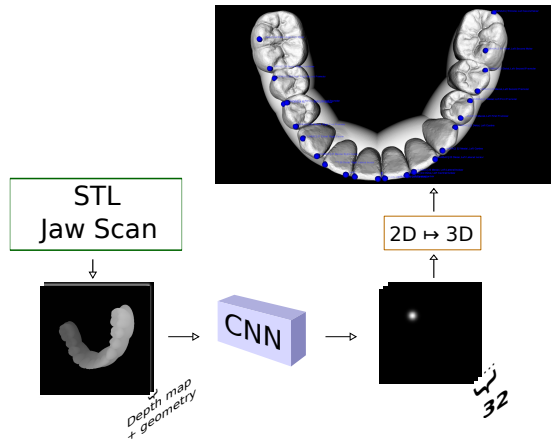


Challenging Real-world Dataset



Proposed Solution – Baseline Method

- U-Net [1], Attention U-Net [2], or Nested U-Net [3] as the architecture of the CNN.
- Information propagation to 3D via ray-casting.

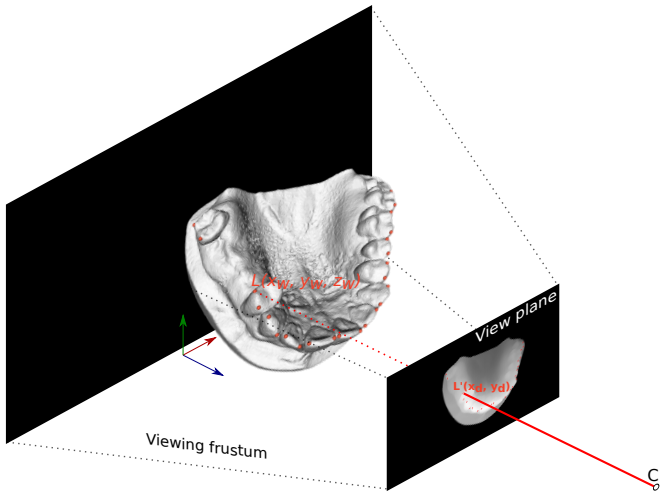


[1] Ronneberger O., Fischer P., Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. MICCAI 2015 (2015).

[2] Oktay, O., Schlemper, J. et al. Attention U-Net: Learning Where to Look for the Pancreas. CoRR. 2018 (2018).

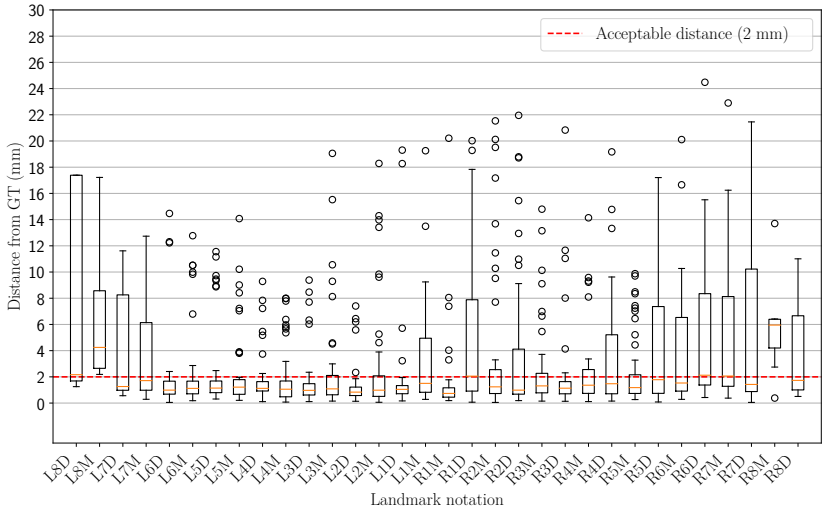
[3] Zhou, Z., Siddiquee, M. M. R., Tajbakhsh, N. and Liang, J. UNet++: A Nested U-Net Architecture for Medical Image Segmentation. MICCAI 2018

Proposed Solution – Baseline Method

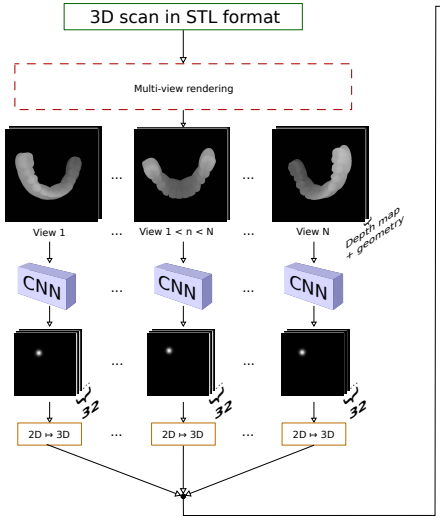


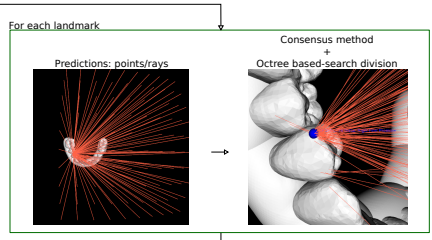
Proposed Solution – Accuracy of Baseline Method

- Accuracy of ***2.94 ± 4.62 mm***
- ***71.94 %*** predictions with the error less than 2 mm.

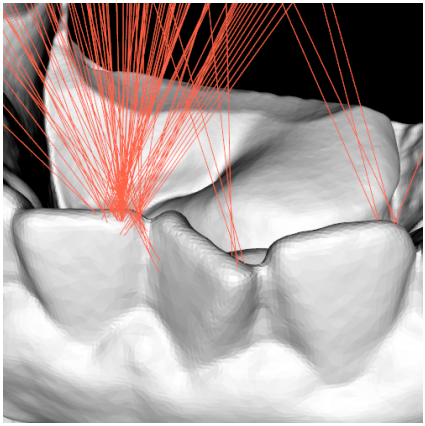


Proposed Solution – The Multi-view Approach





Proposed Solution – Post-processing



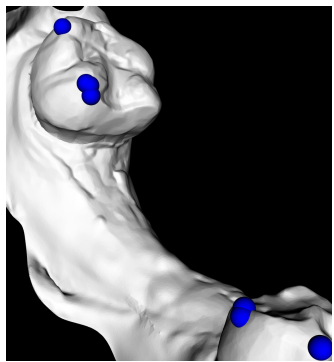
Predictions from each viewpoint are post-processed using **RANSAC method** [4]. With the least-squares fit, individual predictions are classified as **inliers** and **outliers**.

[4] Paulsen, R. R., Juhl, K. A., Haspang, T. M., Hansen, T. F. et al. Multi-view Consensus CNN for 3D Facial Landmark Placement. ACCV 2018 (2018).

Proposed Solution – Analysing Teeth Presence

Complete dentition is rare so we would like to get rid of *false positive* predictions.

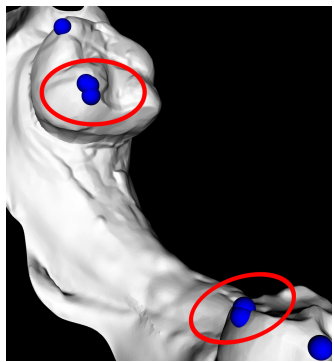
- Binary classification task.
- Post-processing that classifies the teeth presence by *self-evaluation*.
- Based on two uncertainty measures:
  - **Maximum Heatmap Activation Confidence** and
  - **Multi-view Confidence**.



Proposed Solution – Analysing Teeth Presence

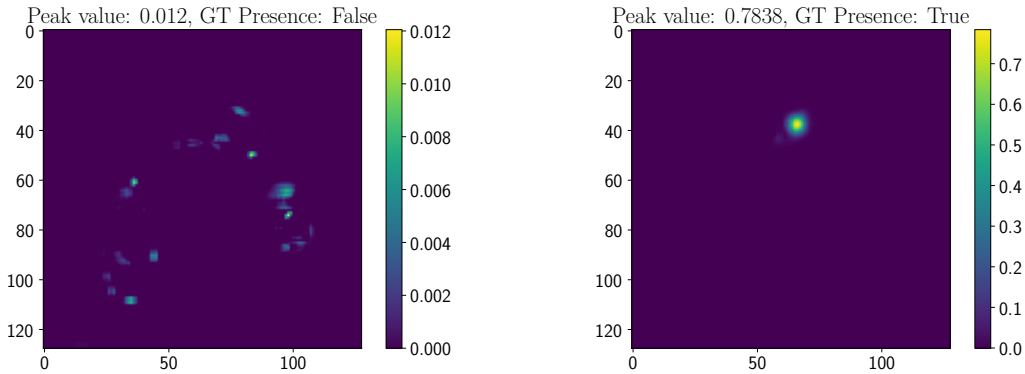
Complete dentition is rare so we would like to get rid of *false positive* predictions.

- Binary classification task.
- Post-processing that classifies the teeth presence by *self-evaluation*.
- Based on two uncertainty measures:
  - **Maximum Heatmap Activation Confidence** and
  - **Multi-view Confidence**.



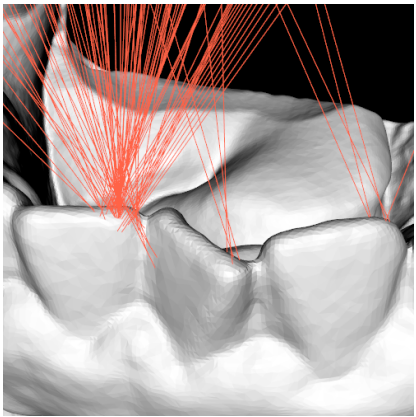
Proposed Solution – Analysing Teeth Presence

Maximum Heatmap Activation Confidence



Proposed Solution – Analysing Teeth Presence

Multi-view Confidence



For both confidentes, the optimal threshold value is determined by ROC curve.

Both confidences are combined to bring a decision with highest possible accuracy.

Experiments and Results

We experimentally investigated and compared several part of the method:

- architecture designs (U-Net, Attention U-Net, and Nested U-Net),
- consensus methods (RANSAC vs. centroid),
- viewpoint numbers (1, 9, 25, and 100) and
- CNN inputs (depth map, geometry rendering and their combination).

Training Procedure

- Inputs, ground truth maps as well as outputs were of size 128×128.
- **Adam** optimizer with the weight decay set to 10^{-3}.
- Batch size of **32**, learning rate of 10^{-3}, **RMSE loss** function.
- Early stopping monitored on validation loss with the patience set to **20**.

Experiments and Results – Overall Results

Architecture & consensus method		Single-view		Multi-view					
				$N = 9$		$N = 25$		$N = 100$	
		$\overline{R}$	SD	$\overline{R}$	SD	$\overline{R}$	SD	$\overline{R}$	SD
U-Net (Depth)	Centroid	2.24	4.02	2.00	2.37	1.74	2.33	1.80	1.96
	RANSAC			1.24	2.86	1.02	3.75	1.01	4.28
U-Net (Geom)	Centroid	2.13	4.41	2.03	3.14	1.69	2.21	1.67	2.41
	RANSAC			1.20	3.01	1.17	2.16	1.06	2.22
U-Net (Depth & Geom)	Centroid	2.02	4.10	1.90	2.12	1.82	2.48	1.85	3.23
	RANSAC			1.01	3.77	0.84	2.05	0.77	1.94
Attention U-Net (Depth)	Centroid	1.73	3.48	2.37	3.37	2.02	2.87	2.01	1.99
	RANSAC			1.18	1.88	1.10	2.05	0.95	1.62
Attention U-Net (Geom)	Centroid	1.72	3.62	2.31	2.68	1.98	2.09	1.96	2.38
	RANSAC			1.14	1.51	1.02	3.75	0.91	1.11
Attention U-Net (Depth & Geom)	Centroid	1.67	3.06	2.00	2.37	1.74	2.33	1.80	1.96
	RANSAC			0.93	1.03	0.79	1.01	**0.75**	**0.96**
Nested U-Net (Depth)	Centroid	1.77	3.32	2.29	2.12	2.32	1.99	2.12	3.04
	RANSAC			1.09	2.60	1.00	1.85	0.95	2.82
Nested U-Net (Geom)	Centroid	1.77	3.00	2.44	1.98	2.30	3.01	2.23	2.58
	RANSAC			1.11	1.83	0.93	1.67	0.93	1.99
Nested U-Net (Depth & Geom)	Centroid	1.69	2.62	2.30	3.18	2.31	2.72	2.16	2.55
	RANSAC			0.98	2.09	0.83	2.12	0.80	1.45

All measured values are in **millimeters** (mm). $\overline{R}$ stands for the average radial error, *SD* stands for Standard Deviation.

Experiments and Results – Overall Results

Architecture & consensus method		Single-view		Multi-view					
				$N = 9$		$N = 25$		$N = 100$	
		$\overline{R}$	SD	$\overline{R}$	SD	$\overline{R}$	SD	$\overline{R}$	SD
U-Net (Depth)	Centroid	2.24	4.02	2.00	2.37	1.74	2.33	1.80	1.96
	RANSAC			1.24	2.86	1.02	3.75	1.01	4.28
U-Net (Geom)	Centroid	2.13	4.41	2.03	3.14	1.69	2.21	1.67	2.41
	RANSAC			1.20	3.01	1.17	2.16	1.06	2.22
U-Net (Depth & Geom)	Centroid	2.02	4.10	1.90	2.12	1.82	2.48	1.85	3.23
	RANSAC			1.01	3.77	0.84	2.05	0.77	1.94
Attention U-Net (Depth)	Centroid	1.73	3.48	2.37	3.37	2.02	2.87	2.01	1.99
	RANSAC			1.18	1.88	1.10	2.05	0.95	1.62
Attention U-Net (Geom)	Centroid	1.72	3.62	2.31	2.68	1.98	2.09	1.96	2.38
	RANSAC			1.14	1.51	1.02	3.75	0.91	1.11
Attention U-Net (Depth & Geom)	Centroid	1.67	3.06	2.00	2.37	1.74	2.33	1.80	1.96
	RANSAC			0.93	1.03	0.79	1.01	**0.75**	**0.96**
Nested U-Net (Depth)	Centroid	1.77	3.32	2.29	2.12	2.32	1.99	2.12	3.04
	RANSAC			1.09	2.60	1.00	1.85	0.95	2.82
Nested U-Net (Geom)	Centroid	1.77	3.00	2.44	1.98	2.30	3.01	2.23	2.58
	RANSAC			1.11	1.83	0.93	1.67	0.93	1.99
Nested U-Net (Depth & Geom)	Centroid	1.69	2.62	2.30	3.18	2.31	2.72	2.16	2.55
	RANSAC			0.98	2.09	0.83	2.12	0.80	1.45

All measured values are in **millimeters** (mm). $\overline{R}$ stands for the average radial error, *SD* stands for Standard Deviation.

Experiments and Results – Overall Results

Architecture & consensus method		Single-view		Multi-view					
				$N = 9$		$N = 25$		$N = 100$	
		$\overline{R}$	SD	$\overline{R}$	SD	$\overline{R}$	SD	$\overline{R}$	SD
U-Net (Depth)	Centroid	2.24	4.02	2.00	2.37	1.74	2.33	1.80	1.96
	RANSAC			1.24	2.86	1.02	3.75	1.01	4.28
U-Net (Geom)	Centroid	2.13	4.41	2.03	3.14	1.69	2.21	1.67	2.41
	RANSAC			1.20	3.01	1.17	2.16	1.06	2.22
U-Net (Depth & Geom)	Centroid	2.02	4.10	1.90	2.12	1.82	2.48	1.85	3.23
	RANSAC			1.01	3.77	0.84	2.05	0.77	1.94
Attention U-Net (Depth)	Centroid	1.73	3.48	2.37	3.37	2.02	2.87	2.01	1.99
	RANSAC			1.18	1.88	1.10	2.05	0.95	1.62
Attention U-Net (Geom)	Centroid	1.72	3.62	2.31	2.68	1.98	2.09	1.96	2.38
	RANSAC			1.14	1.51	1.02	3.75	0.91	1.11
Attention U-Net (Depth & Geom)	Centroid	1.67	3.06	2.00	2.37	1.74	2.33	1.80	1.96
	RANSAC			0.93	1.03	0.79	1.01	**0.75**	**0.96**
Nested U-Net (Depth)	Centroid	1.77	3.32	2.29	2.12	2.32	1.99	2.12	3.04
	RANSAC			1.09	2.60	1.00	1.85	0.95	2.82
Nested U-Net (Geom)	Centroid	1.77	3.00	2.44	1.98	2.30	3.01	2.23	2.58
	RANSAC			1.11	1.83	0.93	1.67	0.93	1.99
Nested U-Net (Depth & Geom)	Centroid	1.69	2.62	2.30	3.18	2.31	2.72	2.16	2.55
	RANSAC			0.98	2.09	0.83	2.12	0.80	1.45

All measured values are in **millimeters** (mm). $\overline{R}$ stands for the average radial error, *SD* stands for Standard Deviation.

Experiments and Results – Overall Results

Architecture & consensus method		Single-view		Multi-view					
				$N = 9$		$N = 25$		$N = 100$	
		$\overline{R}$	SD	$\overline{R}$	SD	$\overline{R}$	SD	$\overline{R}$	SD
U-Net (Depth)	Centroid	2.24	4.02	2.00	2.37	1.74	2.33	1.80	1.96
	RANSAC			1.24	2.86	1.02	3.75	1.01	4.28
U-Net (Geom)	Centroid	2.13	4.41	2.03	3.14	1.69	2.21	1.67	2.41
	RANSAC			1.20	3.01	1.17	2.16	1.06	2.22
U-Net (Depth & Geom)	Centroid	2.02	4.10	1.90	2.12	1.82	2.48	1.85	3.23
	RANSAC			1.01	3.77	0.84	2.05	0.77	1.94
Attention U-Net (Depth)	Centroid	1.73	3.48	2.37	3.37	2.02	2.87	2.01	1.99
	RANSAC			1.18	1.88	1.10	2.05	0.95	1.62
Attention U-Net (Geom)	Centroid	1.72	3.62	2.31	2.68	1.98	2.09	1.96	2.38
	RANSAC			1.14	1.51	1.02	3.75	0.91	1.11
Attention U-Net (Depth & Geom)	Centroid	1.67	3.06	2.00	2.37	1.74	2.33	1.80	1.96
	RANSAC			0.93	1.03	0.79	1.01	**0.75**	**0.96**
Nested U-Net (Depth)	Centroid	1.77	3.32	2.29	2.12	2.32	1.99	2.12	3.04
	RANSAC			1.09	2.60	1.00	1.85	0.95	2.82
Nested U-Net (Geom)	Centroid	1.77	3.00	2.44	1.98	2.30	3.01	2.23	2.58
	RANSAC			1.11	1.83	0.93	1.67	0.93	1.99
Nested U-Net (Depth & Geom)	Centroid	1.69	2.62	2.30	3.18	2.31	2.72	2.16	2.55
	RANSAC			0.98	2.09	0.83	2.12	0.80	1.45

All measured values are in **millimeters** (mm). $\overline{R}$ stands for the average radial error, *SD* stands for Standard Deviation.

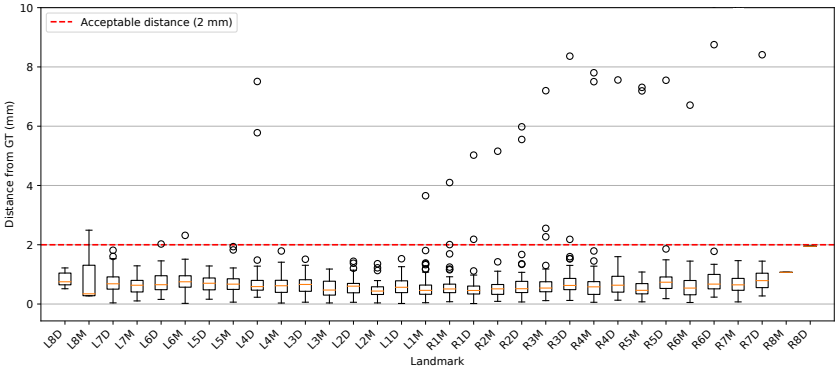
Experiments and Results – Overall Results

Architecture & consensus method		Single-view		Multi-view					
				$N = 9$		$N = 25$		**N = 100**	
		$\overline{R}$	SD	$\overline{R}$	SD	$\overline{R}$	SD	$\overline{R}$	SD
U-Net (Depth)	Centroid	2.24	4.02	2.00	2.37	1.74	2.33	1.80	1.96
	RANSAC			1.24	2.86	1.02	3.75	1.01	4.28
U-Net (Geom)	Centroid	2.13	4.41	2.03	3.14	1.69	2.21	1.67	2.41
	RANSAC			1.20	3.01	1.17	2.16	1.06	2.22
U-Net (Depth & Geom)	Centroid	2.02	4.10	1.90	2.12	1.82	2.48	1.85	3.23
	RANSAC			1.01	3.77	0.84	2.05	0.77	1.94
Attention U-Net (Depth)	Centroid	1.73	3.48	2.37	3.37	2.02	2.87	2.01	1.99
	RANSAC			1.18	1.88	1.10	2.05	0.95	1.62
Attention U-Net (Geom)	Centroid	1.72	3.62	2.31	2.68	1.98	2.09	1.96	2.38
	RANSAC			1.14	1.51	1.02	3.75	0.91	1.11
Attention U-Net (Depth & Geom)	Centroid	1.67	3.06	2.00	2.37	1.74	2.33	1.80	1.96
	RANSAC			0.93	1.03	0.79	1.01	**0.75**	**0.96**
Nested U-Net (Depth)	Centroid	1.77	3.32	2.29	2.12	2.32	1.99	2.12	3.04
	RANSAC			1.09	2.60	1.00	1.85	0.95	2.82
Nested U-Net (Geom)	Centroid	1.77	3.00	2.44	1.98	2.30	3.01	2.23	2.58
	RANSAC			1.11	1.83	0.93	1.67	0.93	1.99
Nested U-Net (Depth & Geom)	Centroid	1.69	2.62	2.30	3.18	2.31	2.72	2.16	2.55
	RANSAC			0.98	2.09	0.83	2.12	0.80	1.45

All measured values are in **millimeters** (mm). $\overline{R}$ stands for the average radial error, *SD* stands for Standard Deviation.

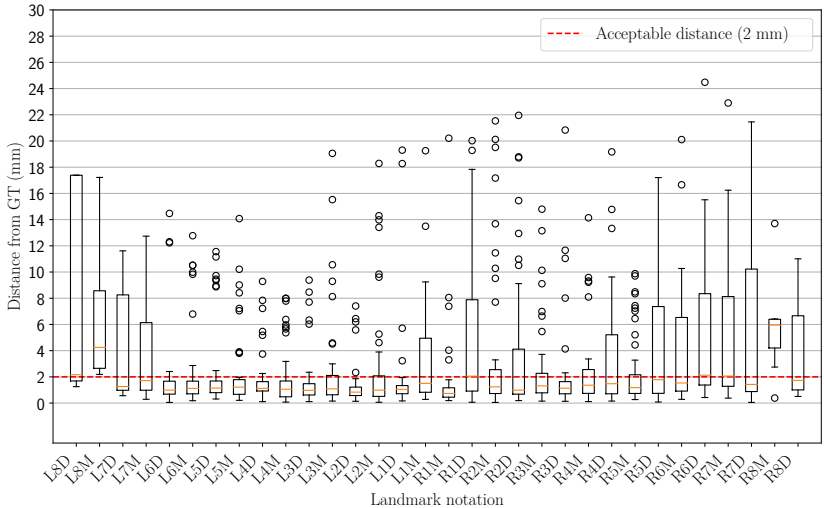
Experiments and Results – Multi-view

- Accuracy of ***0.75 ± 0.96 mm***
- ***98.07 %*** predictions with the error less than 2 mm.

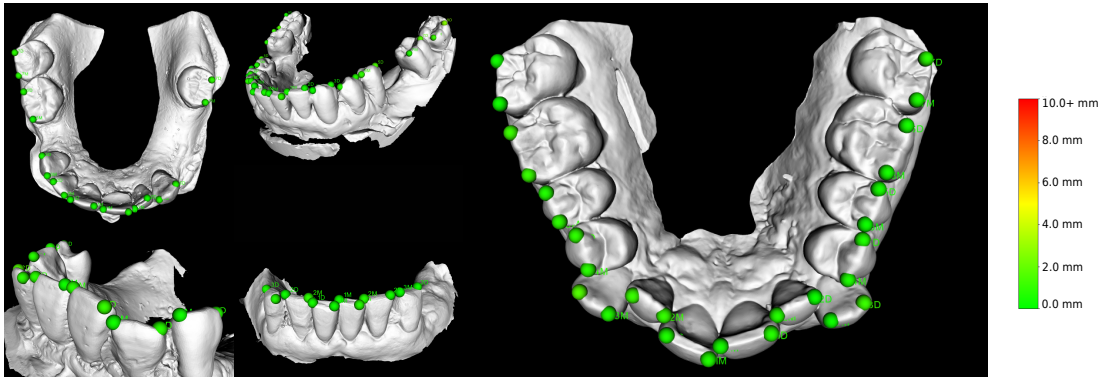


Experiments and Results – Baseline Method

- Accuracy of ***2.94 ± 4.62 mm***
- ***71.94 %*** predictions with the error less than 2 mm.

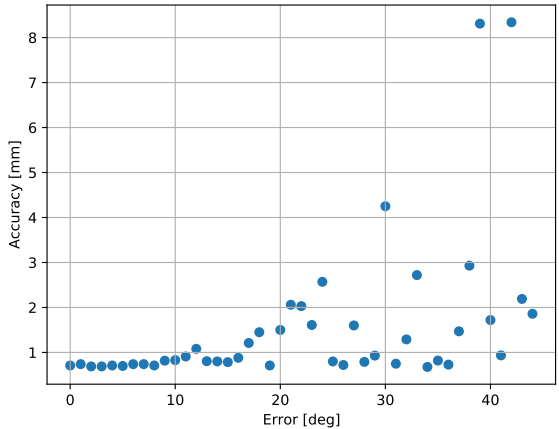


Experiments and Results – Example Outputs



Robustness to Model Rotations

- This approach requires the processed mesh to be specifically aligned – so the occlusal and incisal surfaces face the camera.
- We analysed how the method performs with increasing alignment error.



Summary

- Landmarking accuracy of ***0.75*** $\pm$ ***0.96 mm*** on the test set of challenging ortho cases (Attention U-Net, 100 viewpoints and RANSAC).
- By the analysis of heatmap and RANSAC uncertaintes, our method can correctly detect tooth presence in ***97.68%*** cases.
- Seconds to compute even on average CPUs (for example on laptop Intel Core i7-8750H CPU @ 2.20 GHz with 6 cores).
- Relies on properly aligned mesh.
- **Code and the validation dataset is available to public here**:

`https://tiborkubik.github.io/Robust-Teeth-Detection-in-3D-Dental-Scans/`