

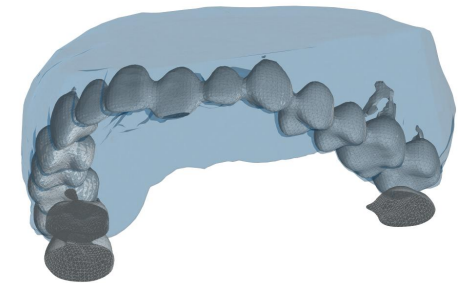


Registration Teaches Registration

Transform-Derived Crown Guidance for
Semi-Supervised CBCT-IOS Alignment

Yi Zhu

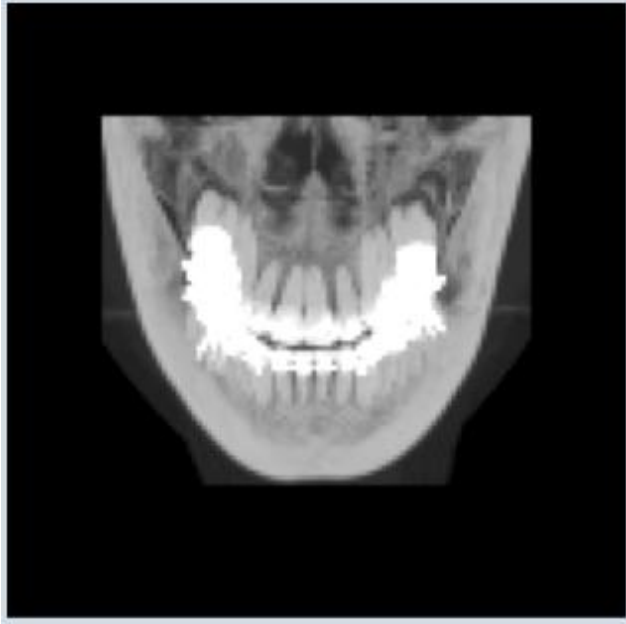
Razmig Kéchichian / Raphaël Richert / Sébastien Valette



STSR 2026 Task 2 / 1st place

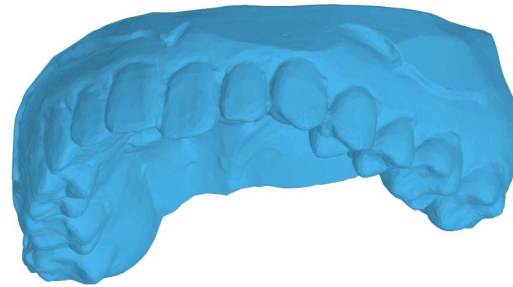
Which anatomy actually overlaps?

CBCT



Crowns, roots and bone

IOS / one jaw



Crowns and gingiva

Partial overlap

Bone distractors

Repeated tooth shapes

Incomplete surfaces

Crown overlap must be identified before a local fit can be trusted.

Learning the overlap region

Segmentation first

Extract anatomy,
then align surfaces

Requires suitable
segmentation supervision

Direct pose learning

Learn the transform
from paired observations

A large pose space
with few labeled pairs

Our approach

Learn where crowns fit.
Use geometry to align.

**Known poses become
weak spatial targets**

Overlap prediction + rigid hypotheses + learned selection

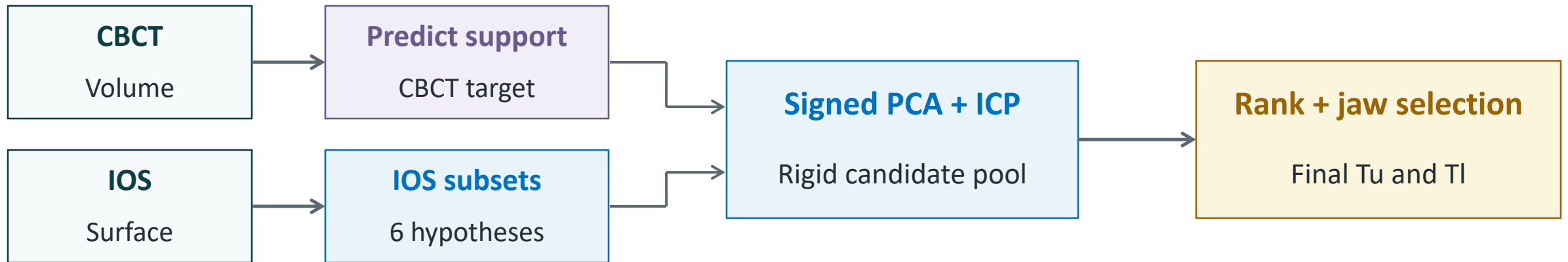
A transform provides spatial supervision for the next registration.

Training and inference in one framework

TRAIN

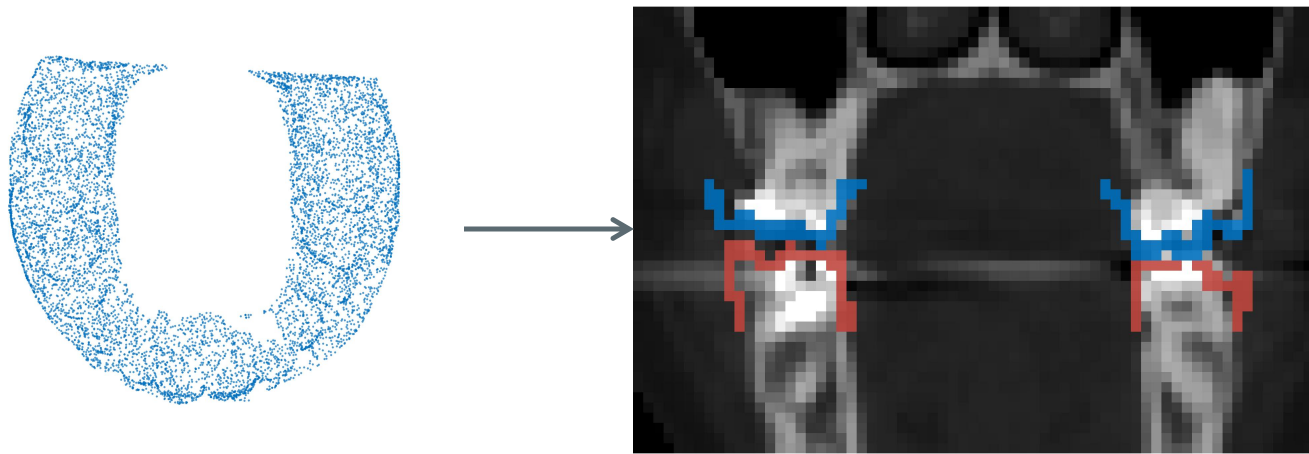


INFER



Training learns crown support; inference searches and selects rigid poses.

A known pose reveals crown support



Aligned upper crown

Weak support in CBCT

Blue: upper / coral: lower

Background / upper / lower
Closest surface resolves overlap

Two 35% PCA tails; CBCT density selects the crown-side target.

The target marks the shared surface region needed by registration.

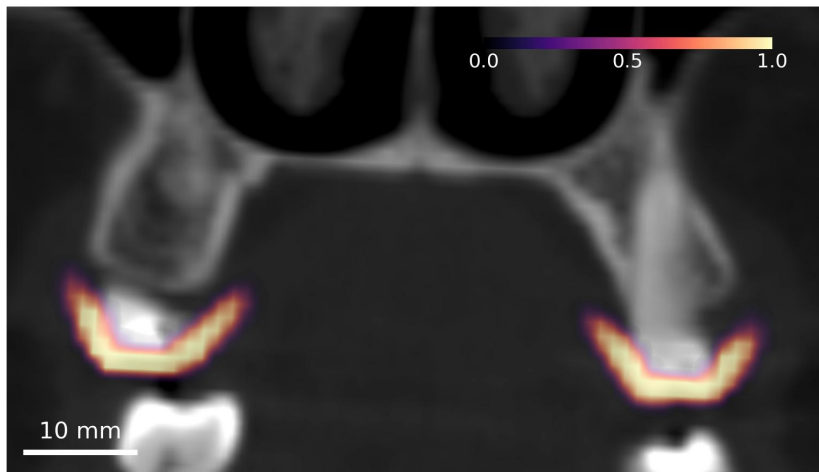
$$C_j = T_j^* X_j^{\text{crown}}$$

$$Y_j = \text{Rasterize}(C_j; A, \delta)$$

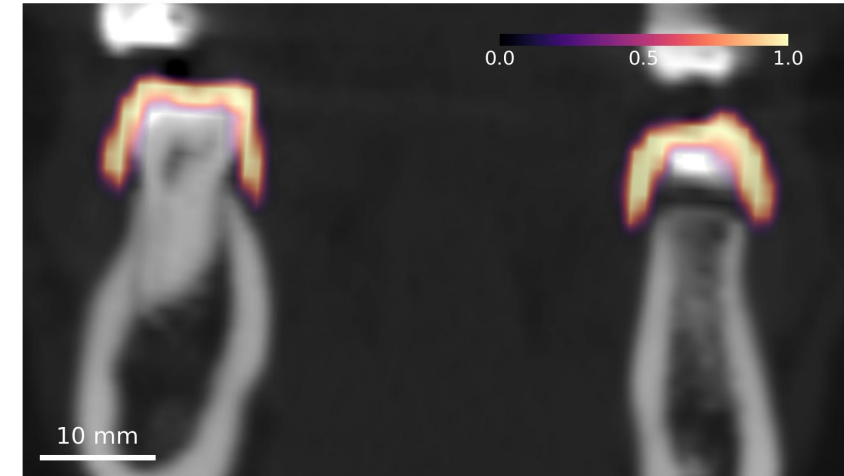
$$\delta = 0.7 \text{ mm} / h = 1.25 \text{ mm}$$

What the network predicts

Upper support probability



Lower support probability



3D U-Net

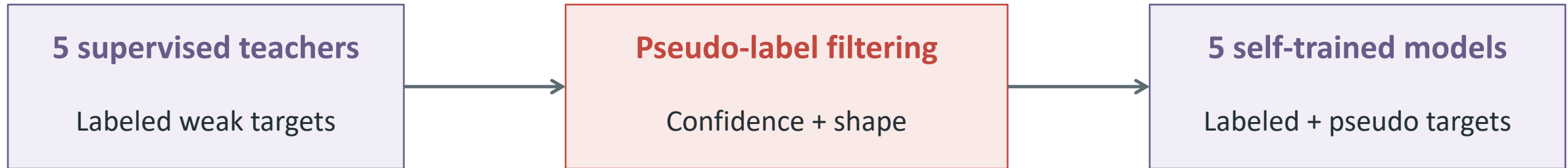
128³ grid / 1.25 mm spacing

$p \geq 0.25$

Case 145 / actual deployed ensemble / grayscale CBCT + probability overlay

A probability field becomes the target for geometric registration.

Unlabeled scans extend the spatial teacher



Unlabeled CBCT

289 complete cases screened
Up to 80 retained per teacher

Training balance

Pseudo sampling weight: 0.2
Labeled targets remain in every epoch

$$p = 0.5 p(\text{supervised}) + 0.5 p(\text{self-trained})$$

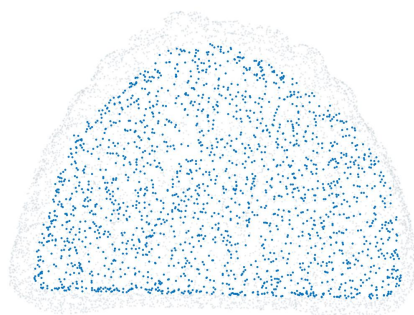
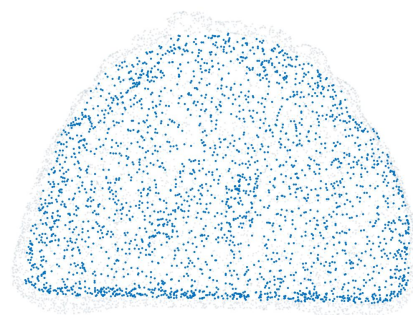
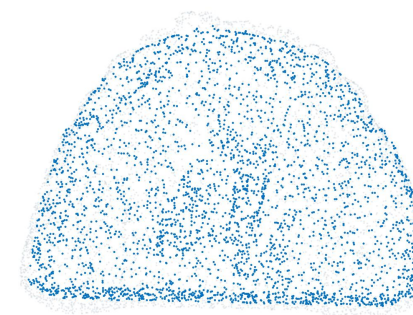
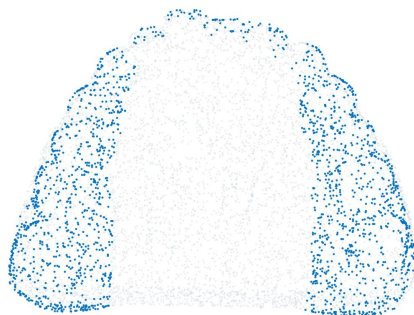
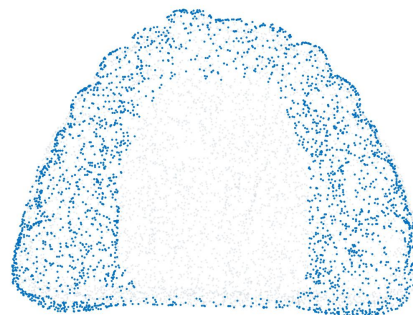
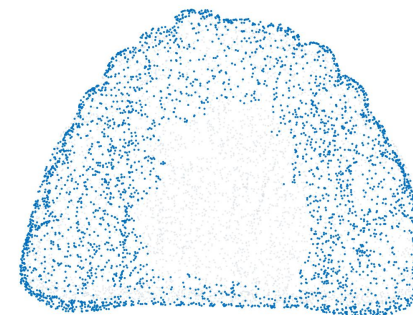
Filtered pseudo targets expand support learning beyond labeled transforms.

Six crown-side hypotheses

25%

35%

45%

Lower
tailUpper
tail

Gray: full IOS / blue: retained points / one jaw, fixed view and scale

Retaining alternatives avoids an irreversible crown-side decision.

Trimmed ICP refines coarse hypotheses

Signed PCA initialization



After multiscale ICP



$$\min_{\mathbf{R}, \mathbf{t}} \sum_{i \in \Omega} \| \mathbf{R} \mathbf{x}_i + \mathbf{t} - \text{NN}_{\mathcal{Y}}(\mathbf{R} \mathbf{x}_i + \mathbf{t}) \|^2$$

Ω : best 70% of valid matches

Matching radius: 8 to 1.25 mm

Gray: predicted CBCT target / blue: IOS subset

Partial overlap is handled by fitting the strongest valid correspondences.

Coordinate parity is part of the pose

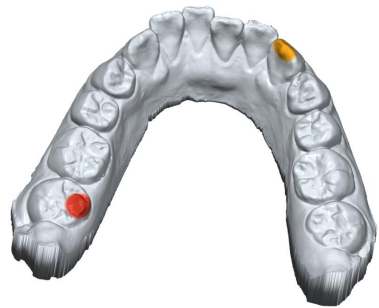
$$T = [R \ t; 0 \ 1]$$

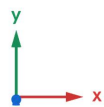
$$R^T R = I$$

$$\det(R) \in \{+1, -1\}$$

Signed PCA initialization
Parity-preserving ICP updates

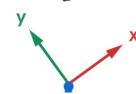
REFERENCE



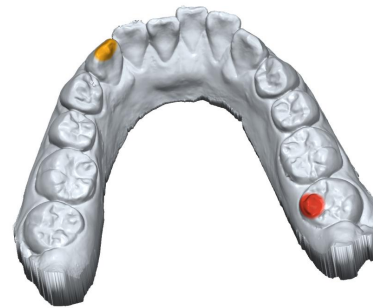

 $\det(R) = +1$
Original handedness

ROTATION




 $\det(R) = +1$
Handedness preserved

REFLECTION




 $\det(R) = -1$
Handedness reversed

Case 280

The optimizer preserves the transform family required by the input convention.

Two complementary views of candidate quality

$\mathbf{z}_i \in \mathbb{R}^{97}$

Distances, coverage, normals, whole-arch fit and candidate context

Regression / 7 forests

$$\hat{e}_i^{(s)} = \exp(f_s(\mathbf{z}_i)) - 1$$

400 trees each; target: $\log(1 + \text{TRE})$
Aggregate: votes for the best candidate

Pairwise / 7 forests

$$q_i^{(s)} = \text{mean}_k P_s(i \succ k \mid \mathbf{z}_i - \mathbf{z}_k)$$

250 trees each; compare opponents
Aggregate: median pairwise cost

$$s_i = 0.575 r(a_i) + 0.425 r(b_i)$$

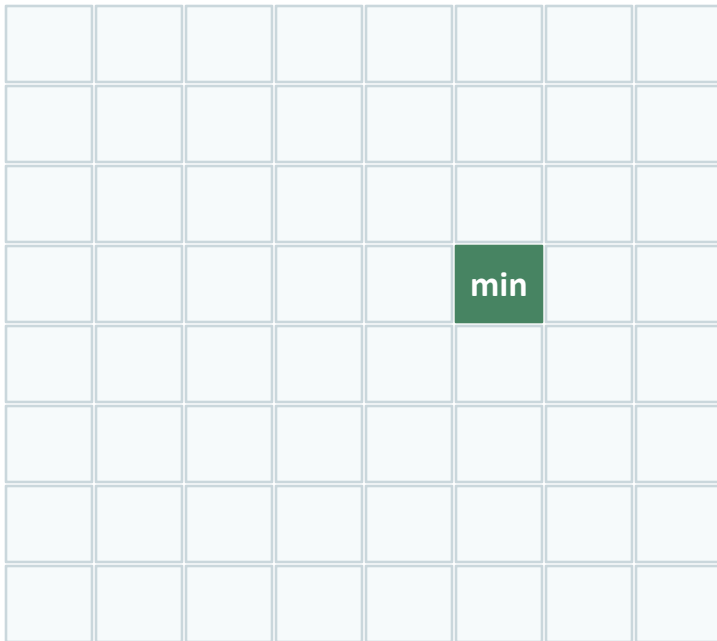
a: regression-vote cost b: negative pairwise win score r: fractional rank

Fractional-rank fusion compares candidate plausibility beyond local ICP fit.

Choose the jaws together

8 upper × 8 lower candidates

Rows: upper candidates i



					min		

Columns: lower candidates j

Conceptual cost grid

$$(i^*, j^*) = \operatorname{argmin}_{i,j} C_{ij}$$

$$C_{ij} = s_{u,i} + s_{l,j} + 0.01 \Delta\theta_{ij} + 0.075 \|\Delta t_{ij}\|_2$$

$\Delta\theta$: relative rotation deviation (degrees)

Δt : translation-difference deviation (mm)

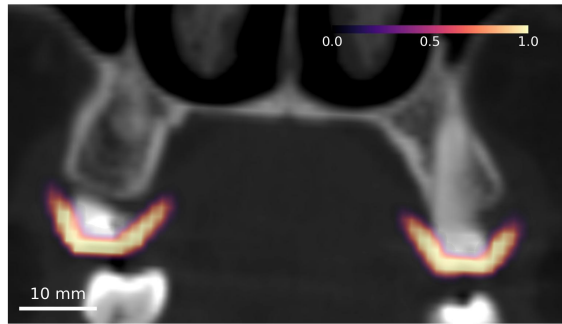
Prior from labeled relative poses

Different-parity pairs are excluded

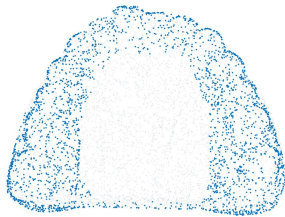
Joint selection keeps separate transforms while favoring a coherent relative pose.

What passes between the modules?

1 Probability field



$3 \times 128^3 \rightarrow N \times 3$ target points



+ 6 IOS source subsets

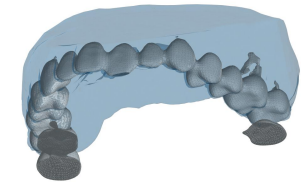
2 Geometric candidates

PCA + ICP
K transforms, each 4×4

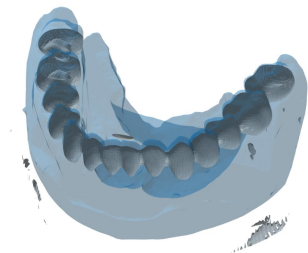
3 Learned selection

97 features per candidate
Top 8 per jaw $\rightarrow$ joint choice

4 Final jaw poses



Upper



Lower

The learned representation guides the pose; the IOS geometry stays unchanged.

Support quality and pose accuracy

Does the guide improve?

$$\text{Chamfer}(Y, \hat{Y}) / \text{Dice}(Y, \hat{Y})$$

Held-out transform-derived support
CBCT-identity-grouped development

Does the pose align?

$$e_t = \| t - t^* \|_2$$

$$e_r = \text{angle}(R^{*T} R)$$

Official reference transformations
Translation: mm / rotation: degrees

Each measurement answers a different scientific question.

Self-training improves the spatial target

Symmetric Chamfer (mm)



Foreground Dice

Model	Dice
Supervised	0.5238
Self-training	0.5321
Fusion	0.5339

-0.0364 mm

Fusion vs. supervised Chamfer

Both support metrics favor fusion, with a modest absolute improvement.

Official registration results

1st

Final dataset

STSR 2026 Task 2 / avalanchezy

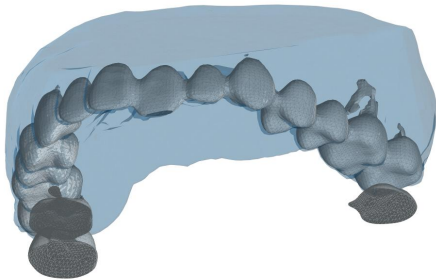
Evaluation	Translation	Rotation	Pairs
Final dataset	7.9025 mm	2.8625°	198 / 198
Online test	8.5435 mm	2.6129°	162 / 162

Different evaluation sets; not an ablation comparison.

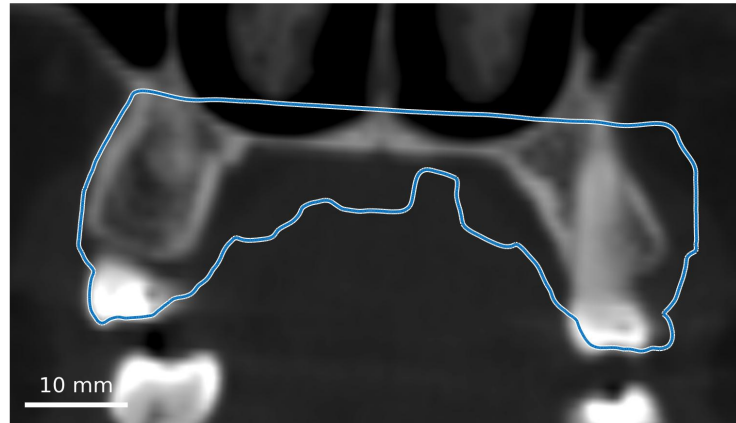
The scores evaluate the complete submitted container.

Upper jaw: surface alignment and slice agreement

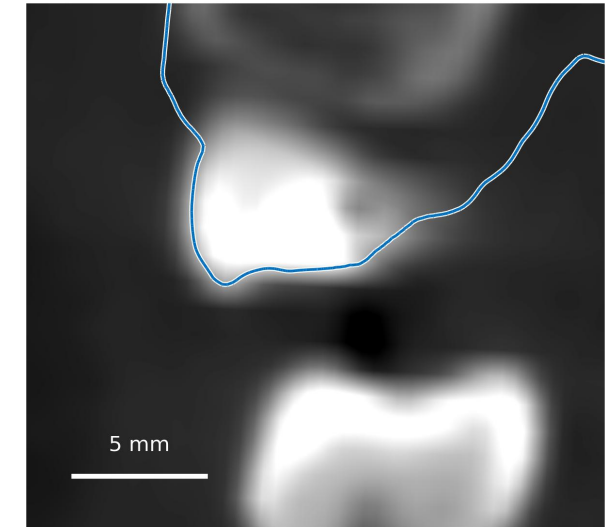
3D context



CBCT + IOS contour



Local detail



■ CBCT

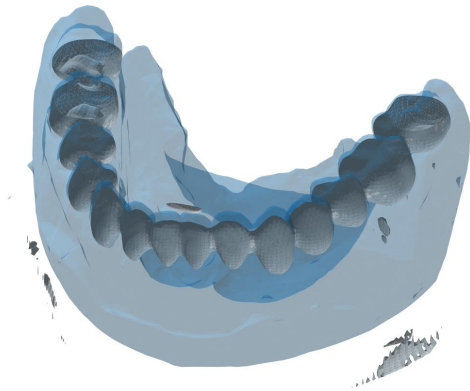
■ Transformed IOS

Case 145 / submitted pose unchanged

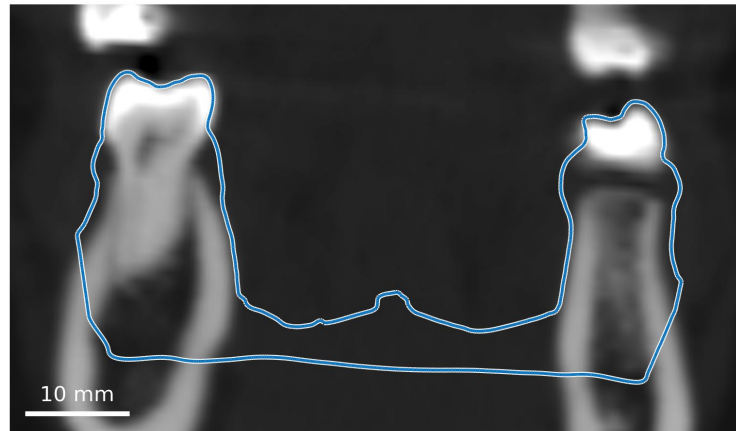
A CBCT section exposes local discrepancies that an opaque overlay can hide.

Lower jaw: surface alignment and slice agreement

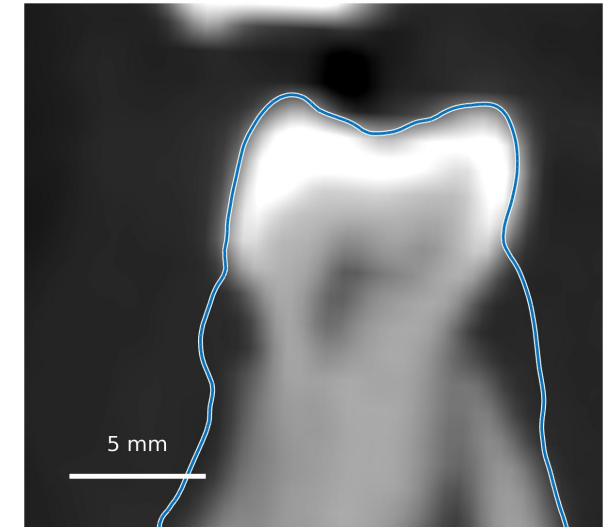
3D context



CBCT + IOS contour



Local detail



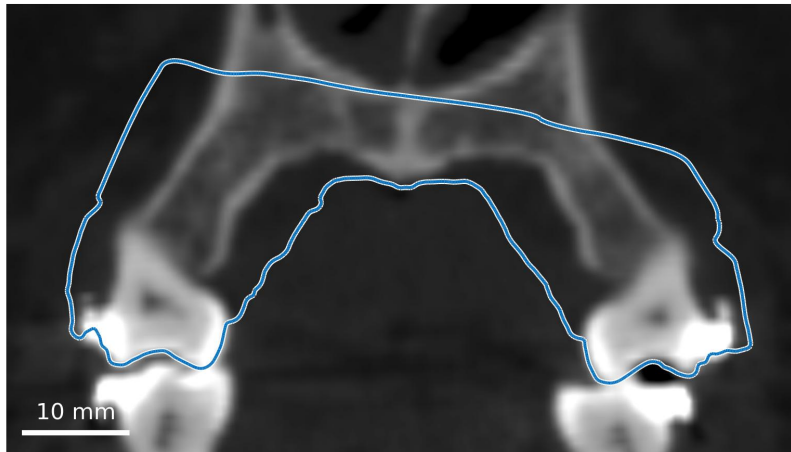
■ CBCT

■ Transformed IOS

Case 145 / submitted pose unchanged

A CBCT section exposes local discrepancies that an opaque overlay can hide.

What remains uncertain?



Case 049 / upper jaw

■ CBCT ■ Transformed IOS

Ambiguous crown support

Isolate support quality in a matched end-to-end registration comparison.

Competing geometric poses

Test ranking and jaw coupling with the same candidate pool.

Generalization and runtime

Independent scans; faster initialization.

Independent acquisitions and matched ablations are the next scientific tests.

A pose label can teach where to register

Known transforms reveal overlap.

Learned support focuses geometry.

Joint selection resolves the pose pair.

Thank you

Yi Zhu / yi.zhu@creatis.insa-lyon.fr

Paper on HAL

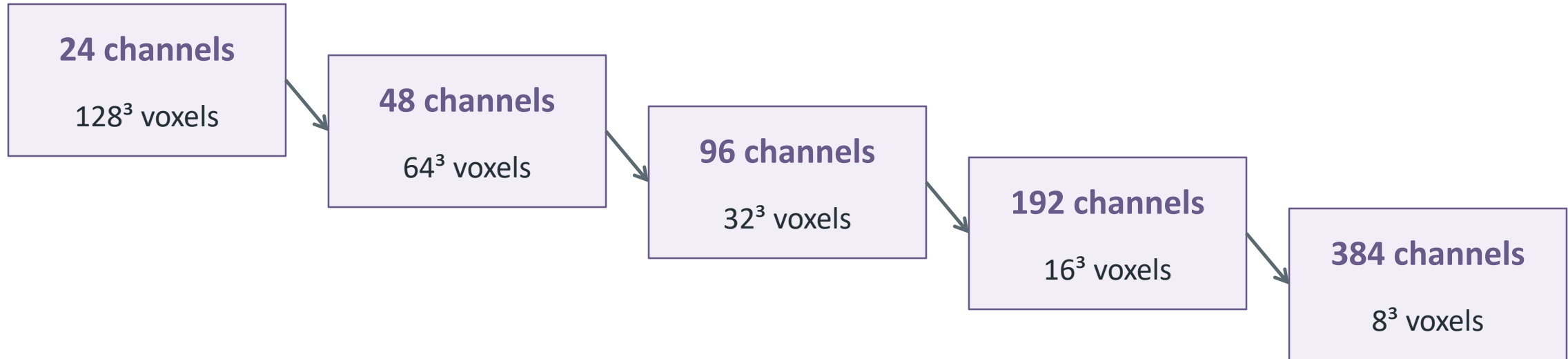


Source code



The spatial predictor

Input: $B \times 1 \times 128^3$



Decoder: transposed convolutions and skip connections

$$L = \text{weighted cross-entropy} + \text{foreground Dice loss}$$

The architecture serves the transform-derived supervision.

Pseudo-target selection

Criterion	Setting
Support voxels / jaw	500–6000
Foreground confidence	≥ 0.45
Foreground entropy	≤ 1.0
Retained pseudo cases	Up to 80 per teacher
Pseudo sampling weight	0.2

5 supervised + 5 self-trained members

Final fitting combines labeled supervision with selected pseudo targets.

Main method and complete submitted system

Main scientific method

Transform-derived support, semi-supervised learning, geometric hypotheses, learned ranking and joint jaw selection.

Additional submitted behavior

Verified reference correspondence for repeated acquisitions.
If verification fails, the general crown-guided method runs.

Official scores describe the full container; the main-method-only score is not isolated.

What the numbers measure

Measure	Compared objects
Translation / rotation error	Predicted and reference transforms
TRE	IOS points under both transforms
Support Chamfer / Dice	Predicted and transform-derived support
ICP target residual	IOS subset and predicted CBCT target
Evaluated pair count	Completed evaluation outputs

Geometric target fit and output completion do not establish reference-based accuracy.