Silhouette-to-Contour Registration: Aligning Intraoral Scan Models with Cephalometric Radiographs

Yiyi Miao[1,2*], Taoyu Wu[3,4*], Ji Jiang[5], Tong Chen[1,2], Zhe Tang[6],
Zhengyong Jiang[1], Angelos Stefanidis[1], Limin Yu[3‡], Jionglong Su[1‡]

[1]School of AI and Advanced Computing, Xi'an Jiaotong-Liverpool University, China
[2]School of Electrical Engineering, Electronics and Computer Science, University of Liverpool, United Kingdom
[3]School of Advanced Technology, Xi'an Jiaotong-Liverpool University, China
[4]School of Physical Sciences, University of Liverpool, Liverpool, United Kingdom
[5]School of Mathematics and Physics, Xi'an Jiaotong-Liverpool University, China
[6]Institute of Artificial Intelligence Innovation, Zhejiang University of Technology, China

Abstract—Reliable 3D–2D alignment between intraoral scan (IOS) models and lateral cephalometric radiographs is critical for orthodontic diagnosis, yet conventional intensity-driven registration methods struggle under real clinical conditions, where cephalograms exhibit projective magnification, geometric distortion, low-contrast dental crowns, and acquisition-dependent variation. These factors hinder the stability of appearance-based similarity metrics and often lead to convergence failures or anatomically implausible alignments. To address these limitations, we propose DentalSCR, a pose-stable, contour-guided framework for accurate and interpretable silhouette-to-contour registration. Our method first constructs a U-Midline Dental Axis (UMDA) to establish a unified cross-arch anatomical coordinate system, thereby stabilizing initialization and standardizing projection geometry across cases. Using this reference frame, we generate radiograph-like projections via a surface-based DRR formulation with coronal-axis perspective and Gaussian splatting, which preserves clinical source–object–detector magnification and emphasizes external silhouettes. Registration is then formulated as a 2D similarity transform optimized with a symmetric bidirectional Chamfer distance under a hierarchical coarse-to-fine schedule, enabling both large capture range and subpixel-level contour agreement. We evaluate DentalSCR on 34 expert-annotated clinical cases. Experimental results demonstrate substantial reductions in landmark error—particularly at posterior teeth—tighter dispersion on the lower jaw, and low Chamfer and controlled Hausdorff distances at the curve level. These findings indicate that DentalSCR robustly handles real-world cephalograms and delivers high-fidelity, clinically inspectable 3D–2D alignment, outperforming conventional baselines.

Index Terms—Intraoral scan, Cephalometric Radiograph, 3D–2D registration

I. INTRODUCTION

The lateral cephalometric radiograph (CR) is pivotal in orthodontic diagnosis and treatment planning. The integration of 3D dental models into the cephalometric coordinate system is increasingly required for reproducible measurements and consistent 3D–2D visualization [1] [2]. However, clinical X-rays exhibit projective magnification and geometric distortion

from the point source, making direct orthogonal projection inaccurate. Furthermore, dental crown contours in lateral projections are thin and low-contrast, susceptible to artifacts and occlusion, which compromises registration robustness. Conventional intensity-driven 2D–3D registration methods depend on accurate forward imaging models and similarity metrics like mutual information, but they typically have limited convergence range under exposure variations and scatter [3]. This creates a clear clinical need for a geometry-based approach that is interpretable, robust to initialization, and suitable for diagnostic use [4].

Existing methodologies comprise three main categories. Intensity-driven methods synthesize DRR and maximize similarity with acquired X-rays [5], yet they remain sensitive to real-world acquisition variations with limited capture range [6]. Geometry-driven methods utilize contours or edges as anchors via distance fields or point set registration [7], but face instability in thin, low-contrast structures. Hybrid and differentiable rendering techniques unify appearance and geometry [8], albeit with increased complexity. While X-ray imaging physics provides reliable geometric constraints [9], a practical solution incorporating structural priors with a stable, interpretable objective function remains lacking [10].

To address this, we propose *DentalSCR*, a 3D–2D dental model registration framework grounded in an "analysis-by-synthesis" paradigm with a contour-distance field at its core. The method begins by constructing a U-midline Dental Axis (UMDA) to establish a shared and reproducible anatomical coordinate system between the upper and lower jaws, standardizing pose initialization and midline consistency while mitigating the impact of inter-individual variation on optimization. It subsequently employs perspective projection and cephalometry-like Gaussian splatting rendering to explicitly preserve the magnification and distortion inherent to the source-object-detector geometry, thereby ensuring geometric alignment between synthesized projections and clinical cephalograms; the geometric correctness can be cross-

---

validated using open-source Digitally Reconstructed Radiography (DRR) tools when necessary [11]. The registration process itself is formulated as the geometric matching of projected STL contours to expert-annotated cephalogram contours, minimizing a symmetric Chamfer distance field loss under 2D similarity transformation through a three-stage coarse-to-fine schedule designed to expand the capture range and achieve sub-pixel refinement within the convergence neighborhood. This is complemented by Hausdorff distance metrics to diagnose residual outliers and local worst-case misalignments [12]. This design maintains geometric interpretability while circumventing any reliance on complex modeling of illumination and contrast, and remains compatible with—and allows for the substitution of—common point-set optimization tools such as Iterative Closest Point (ICP) [13] and Coherent Point Drift (CPD) [14].

Our contributions are threefold. First, we introduce the *U-Midline Dental Axis* (UMDA), a cross-arch reference frame that standardizes initialization and evaluation coordinates, thereby stabilizing geometric optimization. Second, we model clinical acquisition with a coronal-axis perspective and a surface-based DRR with radiograph-like splatting, narrowing the geometric gap between synthesized and acquired images within an analysis-by-synthesis framework. *Third*, we cast silhouette-to-contour registration as a symmetric distance-field objective based on bidirectional Chamfer distance, optimized with a hierarchical coarse-to-fine schedule that balances capture range and subpixel accuracy. We validate DentalSCR on 34 expert-annotated clinical cases: upper-jaw landmark mean error and RMSE decrease substantially, dispersion on the lower jaw tightens, and the largest gains appear in posterior teeth. Curve-level metrics further show low bidirectional Chamfer distances and controlled Hausdorff tails, with residuals concentrated in thin, high-curvature regions. Consistent with these findings, recent systematic reviews and multi-center evidence on automated cephalometric landmarking support the clinical plausibility and translational potential of our approach (e.g., Serafin et al. [15] and Hendrickx et al. [16]), while studies on arch shape and anatomical priors provide additional methodological context [17].

II. RELATED WORK

A. Paradigms of 2D–3D Registration in Medical Imaging

In medical imaging, 2D–3D registration is broadly categorized into three paradigms: intensity-based, feature/geometry-based, and hybrid methods. Intensity-based approaches optimize a similarity metric objective function between the acquired X-ray image and a Digitally Reconstructed Radiograph (DRR) synthesized from a 3D volume or surface model [1]. Metrics such as Mutual Information and Normalized Cross-Correlation are most prevalent, valued for their robustness in cross-modal scenarios or under varying contrast conditions [2]. This paradigm is contingent upon a physically plausible forward imaging model and computationally efficient DRR generation strategies, and is typically coupled with multi-resolution or coarse-to-fine optimization schemes to enlarge

the capture range [4]. Review articles have systematically consolidated this trajectory and its subsequent variants.

Feature or geometry-driven methods use inter-modally stable structures as anchors (landmarks, curves, contours, edges, or silhouettes) and achieve registration through point-to-set or set-to-set geometric distances. Recent advances have progressed in three key aspects: First, for curve or boundary alignment, Chamfer distance transform metrics [18] continue to be adopted, while enhancements in distance field construction and mask design improve numerical stability and sub-pixel resolution. Second, for explicit correspondence-based rigid/near-rigid alignment, various robust or adaptively weighted ICP (Iterative Closest Point) [19] variants have been proposed to handle outliers, low overlap, and noise. Third, in learned registration and neural field frameworks, differentiable geometric losses are introduced as supervision to enhance cross-modal generalization and optimization differentiability. These advancements make geometry-based methods more robust under varying imaging devices, artifacts, and exposure changes, particularly suitable for contour/silhouette alignment in X-ray scenarios.

Hybrid methods integrate appearance evidence with geometric constraints within a unified objective, or alternate between contour-driven pose refinement and intensity-driven fine-tuning. Recent trends involve leveraging differentiable X-ray rendering/physical imaging models to incorporate both DRR intensity similarity and geometric terms (e.g., silhouettes, boundaries, keypoints, or regularizations) into end-to-end or two-stage optimization frameworks [20]. Simultaneously, techniques such as coarse-to-fine parameter scheduling, robust penalty functions, and normalized initialization are employed to enlarge the convergence domain. Emerging research from computer vision and medical imaging demonstrates that such approaches significantly enhance the stability and accuracy of 2D-3D registration under multi-view, cross-patient generalization, and weakly-supervised conditions [21].

B. Structural Priors for 3D Dental Geometry

3D dental models come with strong, actionable priors that can be encoded for registration and reconstruction [22]. First, the dental arch on the occlusal plane follows stable shape families—square, ovoid, and tapered—so spline/template curves can serve as global initializers or population priors, with recent AI systems automating arch-shape classification [23]. Second, craniofacial structures exhibit approximate bilateral symmetry; combining a mid-sagittal reference plane (MSP) with the arch midline constrains pose and lateral deviation, and recent studies review MSP construction and reliability for asymmetry analysis [24]. Third, functional cusps and occlusal contact regions provide robust cues to estimate the occlusal plane [25], unifying upper/lower jaws and improving reproducibility; deep-learning pipelines on CBCT now automate landmark/plane extraction and validate measurement fidelity. Beyond these, tooth count and ordering (UR1–UR7/LR1–LR7), crown–gingiva boundary continuity, and adjacency topology form semantic–topological priors

that support segmentation, instance labeling, alignment, and completion; statistical shape models (SSMs) [26]of whole dentitions or mandibles operationalize such priors as a "shape space". Recent related work leverages these priors in three directions: (i) boundary-preserving 3D dental segmentation, which maintains crown–gingiva junction geometry via mesh-aware sampling and multi-view cues [**?**]; (ii) task-driven applications such as forensic 3D tooth identification and robust segmentation/labeling on partial-arch intraoral scans under limited coverage [27]; and (iii) 3D cephalometric evaluation and normative datasets [28], which provide reusable population standards to anchor geometric alignment and quantitative assessment.

C. Automated Analysis of Lateral Cephalometric Radiography

In lateral cephalograms, the core computer-vision task is automated anatomical landmark detection and quantitative analysis. Recent systematic reviews and meta-analyses report that deep learning achieves accuracy comparable to that of expert tracing on 2D cephalograms, supporting standardized pipelines for routine measurements [15]. Methodologically, approaches span two-stage regressors and multi-scale candidate refinement to end-to-end systems, with stable errors reported on public datasets and growing evidence from prospective or clinical evaluations [29]. Beyond landmarking, cross-imaging 2D alignment and superimposition is increasingly integrated into workflows—for example, automatic registration between a lateral cephalogram and a facial profile photograph to jointly assess soft- and hard-tissue relations—using hierarchical contour matching or learned fusion with validated error analyses [30]. Complementary tasks such as structure segmentation and target detection (e.g., mandibular boundary, lesions, or third-molar status) provide additional channels for automated measurement and quality control, with recent surveys highlighting robustness under low contrast and metal artifacts [31]. In parallel, benchmarks and protocols are being updated: new large-scale cephalometric datasets and overviews place stronger emphasis on annotation consistency, metric definitions, and stratified error reporting, improving comparability and external validity [32]. Finally, to reduce radiation exposure, emerging studies explore landmark inference without X-rays, estimating standard cephalometric points from profile photographs alone, opening avenues for multimodal fusion and streamlined care pathways [33].

III. METHODOLOGY

A. U-Midline Dental Axis Construction

We establish a unified, reproducible, and anatomically consistent 3D reference frame $\{X, Y, Z\}$ for the upper and lower dental meshes. The frame encodes explicit anatomical priors:

- **Vertical axis** (Z) aligns with the dominant normal of the occlusal/crown region;
- **Anterior axis** (X) follows the U-shaped dental arch within the occlusal plane toward the incisors;
- **Lateral axis** (Y) is shared across jaws to avoid subtle yaw discrepancies from independent estimation.

This UMDA provides a stable global reference for cross-jaw alignment and statistics, a reproducible orthogonal basis for 2D projections, and a common coordinate system for quantifying occlusion, arch expansion, and inclination. Embedding the axes into each IOS model also facilitates visual inspection and batch quality control.

Let the mesh vertices be $\{v_i\}$ with normals $\{n_i\}$ and centroid

$$c = \frac{1}{N} \sum_i v_i, \quad \text{where unit}(\cdot) \text{ denotes normalization.} \quad (1)$$

The UMDA is determined sequentially by estimating the vertical axis Z, anterior axis X, and shared lateral axis Y.

1) Vertical Axis Z via Crown Localization: For each vertex, the angle between its surface normal and the occlusal/crown reference is computed. The top $p\%$ of vertices with highest alignment are selected as crown candidates. PCA is then applied, and the eigenvector corresponding to the smallest eigenvalue defines

$$Z = \text{unit}\big(u_{\min}(C)\big), \quad (2)$$

where $u_{\min}(C)$ is the eigenvector of the covariance C with minimal eigenvalue. Orientation is unified by flipping $Z \leftarrow -Z$ if necessary.

2) Anterior Axis X and Initial Lateral Axis Y_0 via U-Shape Fitting: Vertices are projected onto the plane orthogonal to Z; the dental arch is fitted using the baseline b between left and right molars, with the incisal edge f as an anterior reference. The midpoint m on b defines the 2D anterior direction:

$$x_{2D} = \text{unit}\big((f - m) - ((f - m)^\top b)b\big). \quad (3)$$

This vector is back-projected into 3D and orthogonalized with respect to Z:

$$X_0 = \text{unit}(Ux_{2D}), \quad Y_0 = \text{unit}(Z \times X_0). \quad (4)$$

Finally, X_0 is re-normalized via $(Y_0 \times Z)$ to ensure orthogonality and stability.

3) Shared Lateral Axis Y via Cross-Jaw Consistency: To suppress minor asymmetries, the lateral directions from both jaws are averaged and normalized:

$$Y_{\text{sh}} = \text{unit}\big(Y_0^{(\text{up})} + Y_0^{(\text{lo})}\big). \quad (5)$$

With fixed Z, the orthogonal basis is completed while maintaining X consistency:

$$X = \text{unit}(Y_{\text{sh}} \times Z), \quad Y = Y_{\text{sh}}. \quad (6)$$

The final UMDA frame is $R = [X, Y, Z]$ with origin $o = c$.

The UMDA induces three clinically interpretable anatomical planes: the X–Y plane (occlusal plane), which is orthogonal to the Z-axis and serves as a stable basis for 2D projections; the X–Z plane (U-midline plane), which encodes midline symmetry to assess anterior–posterior and vertical deviations; and the Y–Z plane (coronal plane), which captures left–right symmetry and is useful for evaluating lateral tilt and arch inclination.
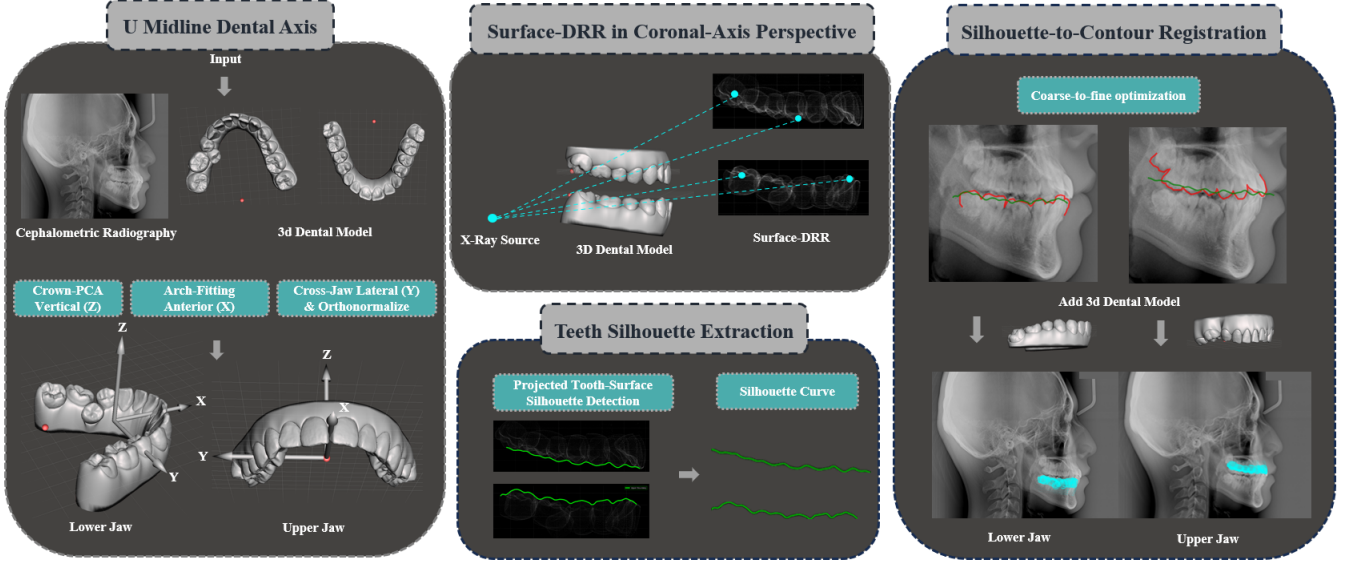
Fig. 1. **Overall of the proposed framework.** The system first constructs a unified U-Midline Dental Axis (UMDA) to establish a reproducible anatomical coordinate frame shared by both arches. Using this reference, intraoral scan (IOS) meshes are projected into radiograph-like views through a coronal-axis surface-based DRR with Gaussian splatting, preserving clinical magnification and geometric distortion. The resulting silhouette is then aligned to expert-annotated cephalometric contours via a 2D similarity transform optimized with a symmetric bidirectional Chamfer distance under a hierarchical coarse-to-fine schedule. This pipeline yields stable initialization, radiographically consistent projections, and accurate silhouette-to-contour registration.

B. Surface-DRR in Coronal-Axis Perspective

After establishing the UMDA coordinate system, we further simulate the true radiographic geometry to generate lateral X-ray–like projections consistent with clinical cephalometric imaging. Traditional DRR are typically based on volumetric data such as CT or CBCT, where the X-ray intensity at each detector pixel (u, v) is computed by integrating the attenuation coefficient $\mu(\mathbf{x})$ along the corresponding ray path $\mathcal{R}(u, v)$ according to the Beer–Lambert law:

$$I(u,v) = I_0 \exp\Big(-\int_{\mathcal{R}(u,v)} \mu(\mathbf{x}) \, \mathrm{d}l \Big). \quad (7)$$

However, intraoral scan (IOS) models represent only the tooth surfaces as triangular meshes without internal density information. Since our registration objective focuses on the external *silhouette* rather than internal attenuation, we adopt a **surface-based DRR** formulation. The mesh is treated as a hollow shell of uniform density ρ_0, and the projected contribution of each surface element is accumulated on the detector plane to form an equivalent thickness map $L(u, v)$:

$$L(u,v) = \sum_{i=1}^{N} \rho_0 \, a_i \, K_\sigma(u - y_i', \, v - z_i'), \quad (8)$$

where a_i is the area weight of the i-th surface element and K_σ is a kernel centered at (y_i', z_i') (typically Gaussian). The final intensity is then derived by substituting $L(u, v)$ into the Beer–Lambert model:

$$I(u,v) = I_0 \exp(-\mu_0 \, L(u,v)), \quad (9)$$

where μ_0 is a uniform effective attenuation coefficient. This surface-based DRR preserves the physical interpretability of radiographic formation while avoiding fictitious volumetric assumptions, emphasizing geometrically accurate outer contours.

In terms of imaging geometry, we employ a **coronal-axis perspective model** consistent with clinical lateral cephalograms. In the UMDA coordinate frame, X denotes the left–right direction, Y denotes anterior–posterior, and Z denotes superior–inferior. The X-ray source is placed at $S = (x_s, 0, 0)$ along the $+X$ direction, and the detector plane is defined as $X = X_{\mathrm{DET}}$, such that rays propagate along the coronal axis through the dentition model. For a vertex $V = (x, y, z)$, the perspective projection onto the detector plane follows:

$$P(t) = S + t(V - S), \qquad t = \frac{X_{\mathrm{DET}} - x_s}{x - x_s}, \quad (10)$$

and the corresponding detector coordinates are

$$(y', z') = (t \, y, \, t \, z). \quad (11)$$

This configuration explicitly models the depth-dependent magnification and geometric distortion determined by the source–object–detector (SOD/SID) geometry, yielding a radiographic appearance and spatial proportion consistent with real CR.

a) Gaussian splatting approximation.: Because the intraoral mesh is a sparse surface point set, direct projection produces discontinuous and aliased silhouettes. We therefore employ a **Gaussian splatting** strategy to approximate a continuous thickness distribution:

$$L(u,v) = \sum_{i=1}^{N} w_i \exp\Big(-\frac{(u - y_i')^2 + (v - z_i')^2}{2\sigma_i^2} \Big), \quad (12)$$

where $w_i = \rho_0 a_i$ is the effective weight and σ_i controls spatial smoothness. When used within the exponential attenuation model $I(u,v) = I_0 \exp(-\mu_0 L(u,v))$, this formulation produces radiograph-like intensity with soft gradients and realistic contours. Alternatively, when only geometric alignment is required, $L(u,v)$ can be directly used as a silhouette map without radiometric scaling.

C. Silhouette-to-Contour Registration

After obtaining the lateral X-ray projection of the dental IOS model meshes, we align the *projected IOS model silhouette* (outer boundary of the 3D projection) to the *annotated CR contour* (polyline on the radiograph) to establish geometric correspondence between the synthesized projection and the clinical CR. We formulate this as a 2D similarity transform that maps the projected IOS model silhouette/contour $\mathcal{C}_{\text{IOS model}}$ onto the CR contour $\mathcal{C}_{\text{CR}}$. For a projected vertex $\mathbf{p} = [y, z]^\top$, the transform is

$$\mathbf{p}' = s\, R(\theta) \begin{bmatrix} y \\ -z \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix}, \qquad (13)$$

where s is the isotropic scale, (t_x, t_y) are in-plane translations, θ is the in-plane rotation, and the sign inversion on z follows the radiographic convention that image y increases downward.

To quantify alignment, we use the symmetric Chamfer distance between the two sets:

$$
\begin{aligned}
\mathcal{L}_{\text{Chamfer}} = {} & \frac{1}{|\mathcal{C}_{\text{IOS model}}|} \sum_{\mathbf{p} \in \mathcal{C}_{\text{IOS model}}} \min_{\mathbf{q} \in \mathcal{C}_{\text{CR}}} \|\mathbf{p} - \mathbf{q}\|_2 \\
& + \frac{1}{|\mathcal{C}_{\text{CR}}|} \sum_{\mathbf{q} \in \mathcal{C}_{\text{CR}}} \min_{\mathbf{p} \in \mathcal{C}_{\text{IOS model}}} \|\mathbf{q} - \mathbf{p}\|_2.
\end{aligned}
\qquad (14)
$$

For robustness and accuracy, we adopt a hierarchical coarse-to-fine optimization. We first initialize the similarity parameters by matching bounding boxes and centroids of $\mathcal{C}_{\text{IOS model}}$ and $\mathcal{C}_{\text{CR}}$ to obtain a scale-consistent, well-centered starting point. A *coarse* stage then explores a broad parameter range to reject spurious hypotheses and accommodate inter-case variability (including jaw differences). Next, a *fine* stage shrinks the bounds around the current estimate to sharpen alignment to the anatomical outline. Finally, a *super-fine* stage performs sub-pixel refinement within the local basin to further reduce residuals. This multi-stage procedure balances global stability with local precision, yielding reliable, high-accuracy *silhouette-to-contour* registration across cases.

IV. EXPERIMENTS

A. Dataset and Annotation Policy

We collected 34 cases; each case includes upper and lower 3D dental models and the corresponding CR. All contours and landmarks were independently annotated and reconciled by two experienced orthodontists. Tooth codes follow standard orthodontic notation: **UR = Upper Right**, **LR = Lower Right**; digits 1–7 denote central incisor, lateral incisor, canine, first premolar, second premolar, first molar, and second molar, respectively. To avoid target leakage, the training signal comes from contours only, whereas quantitative evaluation uses landmarks only.

TABLE I
TOOTH CODE MAPPING USED FOR CONTOUR ANNOTATION (UR/LR COMBINED).

Upper Right (UR)		Lower Right (LR)	
Code	**Tooth (EN)**	**Code**	**Tooth (EN)**
UR1	Central incisor	LR1	Central incisor
UR2	Lateral incisor	LR2	Lateral incisor
UR3	Canine	LR3	Canine
UR4	First premolar	LR4	First premolar
UR5	Second premolar	LR5	Second premolar
UR6	First molar	LR6	First molar
UR7	Second molar	LR7	Second molar

1) Contours for Training: On the CR, the visible labial/buccal boundary of each tooth is traced as segmented polylines and concatenated, in anatomical order, into a continuous arch per jaw. The resulting upper and lower arches—illustrated in Fig. 2—are used to construct the distance field and to drive contour-to-contour alignment. Table I lists the tooth-code mapping used during annotation.
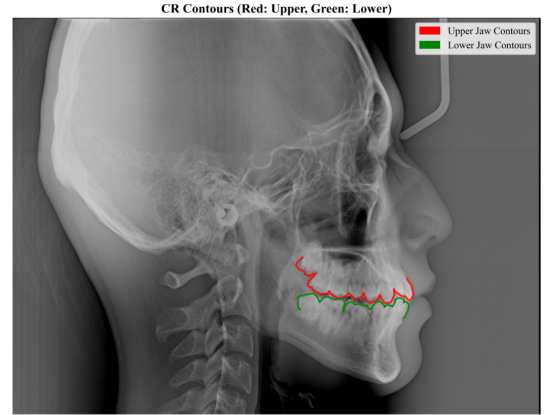

Fig. 2. Contours on the cephalometric radiograph. Upper arch shown in red and lower arch in green; each arch is formed by concatenating tooth-level labial/buccal polylines in anatomical order.

2) Landmarks for Evaluation: Figures 3 and Figures 4 illustrate the evaluation landmarks in 3D and 2D, respectively. In 3D, the upper (UR, blue) and lower (LR, red) IOS model surfaces are overlaid with landmark points spanning anterior and posterior segments. In 2D, the same set is placed on the lateral CR. Landmarks are selected to be clinically meaningful, visually distinctive, and consistently visible across IOS model and CR. The corresponding codes and anatomical definitions are summarized in Table II.

Selection rationale: to cover both jaws and both anterior/posterior segments with points that are clinically meaningful, visually distinctive, and cross-modality visible. Incisal edges provide midline and vertical references; canine cusps act as robust occlusal anchors; first-molar buccal cusps serve as standard occlusal benchmarks; when visible, second-molar mesiobuccal cusps extend posterior coverage. All points must
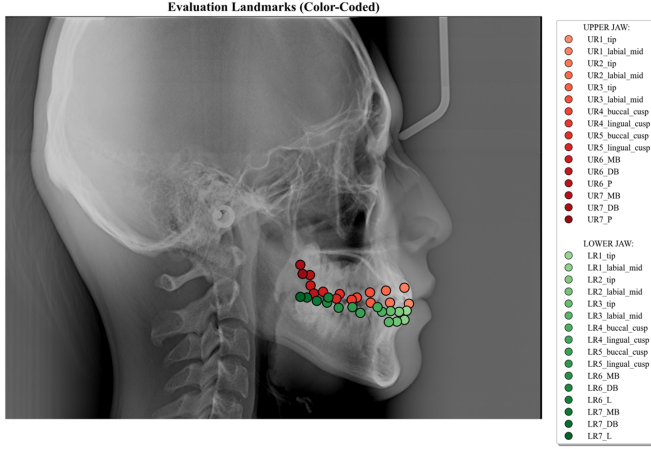
Fig. 3. Two-dimensional view displaying the corresponding anatomical landmarks on a lateral cephalometric radiograph. Landmark definitions and their corresponding codes are detailed in Table II.
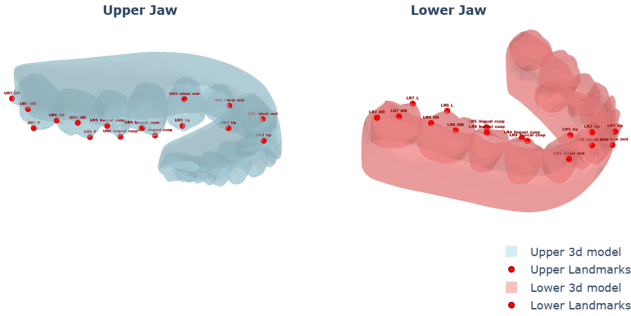

Fig. 4. Three-dimensional visualization of the upper jaw (UR, blue) and lower jaw (LR, red) dental models with overlaid anatomical landmarks (red points). The anatomical significance of the corresponding landmarks is detailed in Table II.

be clearly identifiable on both IOS model and CR. Table II summarizes the UR/LR pairs and their anatomical definitions.

3) Per-case Data Size and Registration Parameters: For completeness, the data dimensions used in the silhouette-guided optimization are reported. The CR image size was 703×938 pixels. The number of contour points extracted was 79 for the upper jaw and 75 for the lower jaw on the CR, which were matched against 441 and 439 points from the IOS model contours for the upper and lower jaw, respectively.

B. Evaluation metrics

Let $\mathcal{C}_{3D} = \{\, \mathbf{p}_i \in \mathbb{R}^2 \,\}_{i=1}^N$ be the projected *IOS model* silhouette points and $\mathcal{C}_{CR} = \{\, \mathbf{q}_j \in \mathbb{R}^2 \,\}_{j=1}^M$ the annotated CR contour points. Distances are Euclidean in pixels and lower is better. For a point set $\mathcal{A}$ and $\mathcal{B}$, define the point-to-set distance

$$d(\mathbf{x}, \mathcal{B}) \;=\; \min_{\mathbf{y} \in \mathcal{B}} \|\mathbf{x} - \mathbf{y}\|_2. \qquad (15)$$

Code (UR)	Code (LR)	Anatomical definition
UR1_tip	LR1_tip	Central incisor incisal edge
UR2_tip	LR2_tip	Lateral incisor incisal edge
UR3_cusp	LR3_cusp	Canine cusp tip
UR6_MB	LR6_MB	1st molar mesiobuccal cusp
UR6_DB	LR6_DB	1st molar distobuccal cusp
UR7_MB	LR7_MB	2nd molar mesiobuccal cusp

1) One-sided mean distances: We report the two directed means to diagnose asymmetric coverage:

$$\overline{D}\big(\mathcal{C}_{3D} \!\rightarrow\! \mathcal{C}_{CR}\big) = \frac{1}{N} \sum_{i=1}^N d(\mathbf{p}_i, \mathcal{C}_{CR}),$$
$$\overline{D}\big(\mathcal{C}_{CR} \!\rightarrow\! \mathcal{C}_{3D}\big) = \frac{1}{M} \sum_{j=1}^M d(\mathbf{q}_j, \mathcal{C}_{3D}). \qquad (16)$$

2) Chamfer distance (bi-directional mean): We use the symmetric (averaged) Chamfer distance:

$$\mathrm{CD}(\mathcal{C}_{3D}, \mathcal{C}_{CR}) = \frac{1}{2}\Big[\overline{D}\big(\mathcal{C}_{3D} \!\rightarrow\! \mathcal{C}_{CR}\big) + \overline{D}\big(\mathcal{C}_{CR} \!\rightarrow\! \mathcal{C}_{3D}\big)\Big]. \quad (17)$$

(Equivalently, some works sum the two directed means; we report their average.)

3) Hausdorff distance: The directed Hausdorff distance is

$$h\big(\mathcal{A} \!\rightarrow\! \mathcal{B}\big) = \max_{\mathbf{x} \in \mathcal{A}} d(\mathbf{x}, \mathcal{B}), \qquad (18)$$

and the symmetric Hausdorff distance is

$$\mathrm{H}(\mathcal{C}_{3D}, \mathcal{C}_{CR}) = \max\Big\{ h\big(\mathcal{C}_{3D} \!\rightarrow\! \mathcal{C}_{CR}\big),\; h\big(\mathcal{C}_{CR} \!\rightarrow\! \mathcal{C}_{3D}\big) \Big\}. \quad (19)$$

V. QUANTITATIVE AND QUALITATIVE RESULTS

We compare three settings under the unified UMDA frame using landmarks as the primary metric: Bounding Box (initial alignment only), Single-Stage (80 iterations), and DentalSCR with a three-stage schedule. As summarized in Table III, DentalSCR consistently lowers landmark errors across jaws. On the upper jaw, the mean decreases from 43.508 (Bounding Box) and 54.567 (Single-Stage) to 17.546, with RMSE dropping to 20.339 and dispersion narrowing (std 10.288). On the lower jaw, the mean reduces from 44.887/49.625 to 40.617 with RMSE 42.853. Stratifying by tooth type (Inc.&Can.: Incisors and Canines; PreM: Premolars) shows that posterior structures benefit the most on the upper jaw (molars mean from 70.539 to 14.513), while the lower molars also improve to 24.652; by contrast, the lower anterior segment remains relatively harder (incisors & canines mean 49.644), reflecting thinner silhouettes and stronger curvature on lateral projections.

To contextualize the final fit quality at the curve level, Table IV reports contour-distance metrics for DentalSCR only. The bi-directional Chamfer is 8.397 for the upper jaw and 4.055 for the lower jaw, indicating tight two-way coverage
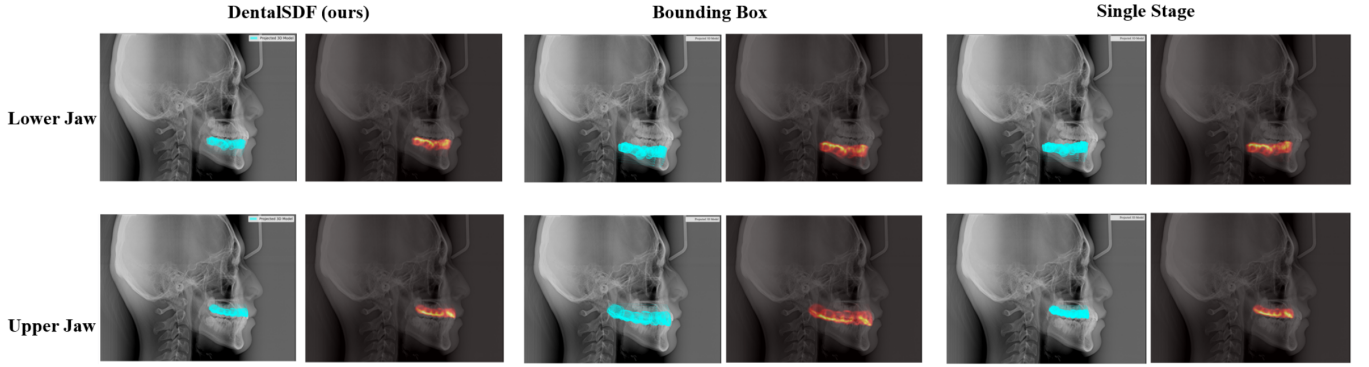
Fig. 5. Qualitative comparison on representative cases. Columns: DentalSCR (ours), Bounding Box, Single-Stage. DentalSCR produces tighter IOS model→CR overlays with fewer gaps, especially at incisors/canines and posterior cusps.

TABLE III
BASELINE COMPARISON WITH *landmark-based* EVALUATION (PIXELS). VALUES ARE ROUNDED TO THREE DECIMALS.

Upper Jaw						
Method	Mean	RMSE	Std	Inc.&Can.	PreM	Molars
Bounding Box	43.508	50.734	26.094	18.405	40.618	70.539
Single-Stage	54.567	58.666	21.546	35.163	51.224	76.199
DentalSCR	17.546	20.339	10.288	23.472	13.205	14.513
Lower Jaw						
Method	Mean	RMSE	Std	Inc.&Can.	PreM	Molars
Bounding Box	44.887	46.859	13.451	29.939	54.100	53.693
Single-Stage	49.625	50.662	10.199	38.577	58.672	54.640
DentalSCR	40.617	42.853	13.660	49.644	51.025	24.652

TABLE IV
CONTOUR-DISTANCE METRICS FOR THE FINAL DENTALSCR ALIGNMENT (PIXELS). VALUES ARE ROUNDED TO THREE DECIMALS.

Metric	Upper Jaw	Lower Jaw
Chamfer distance (bi-directional mean)	8.397	4.055
Hausdorff distance	45.714	22.045
IOS → CR mean distance	6.082	3.585
CR → IOS mean distance	10.713	4.526

of the CR silhouettes. The Hausdorff distances (45.714 upper; 22.045 lower) suggest residual worst-case discrepancies primarily in thin or highly curved regions. The single-sided means (6.082 IOS model→CR vs. 10.713 CR→IOS model on the upper jaw) further indicate mild under-coverage of certain annotated edges by the projected IOS model, which aligns with the larger upper-jaw Hausdorff. Qualitatively, visual inspection (not shown) confirms that Single-Stage aligns the bulk but leaves small gaps around incisal edges and canine tips, whereas the three-stage DentalSCR reduces these residuals and yields cleaner overlap at posterior cusps.

TABLE V
ABLATION RESULTS WITH *landmark-based* EVALUATION (PIXELS). UPPER AND LOWER JAWS ARE SHOWN SIDE-BY-SIDE; VALUES ARE ROUNDED TO THREE DECIMALS.

Config	Upper Jaw (landmarks)			Lower Jaw (landmarks)		
	Mean	RMSE	Std	Mean	RMSE	Std
DentalSCR (Ours)	17.546	20.339	10.288	40.617	42.853	13.660
w/o UMDA	49.472	52.328	17.050	80.423	88.578	37.125
w/o Bi-Chamfer	42.089	46.296	19.285	48.871	50.004	10.588
w/o Coarse-to-Fine	54.567	58.666	21.546	49.625	50.662	10.199

A. Ablation studies

We evaluate each component's contribution via landmark-based evaluation on both jaws (Table V). Removing the UMDA coordinate system causes the most significant degradation, with mean errors reaching 49.472 (upper) and 80.423 (lower), and substantially increased dispersion on the lower jaw (std 37.125), highlighting the necessity of a stable cross-case, cross-jaw frame. Eliminating the coarse-to-fine schedule also reduces accuracy (upper 54.567; lower 49.625), indicating that single-stage optimization tends to stall with residual misalignment even when variability appears moderate. Replacing symmetric Chamfer distance with a one-sided variant raises the means to 42.089 (upper) and 48.871 (lower), demonstrating that balanced bidirectional coverage is essential to prevent systematic silhouette under- or over-coverage. Among all variants, DentalSCR achieves the lowest means (17.546 upper; 40.617 lower) and RMSEs (20.339; 42.853), confirming that integrating UMDA, symmetric Chamfer, and the three-stage schedule is crucial for consistent landmark alignment.

VI. CONCLUSION AND FUTURE WORK

We presented *DentalSCR*, a silhouette–distance field framework for registering 3D dental models to lateral cephalograms. It integrates a reproducible U-midline Dental Axis (UMDA), perspective-consistent rendering, and a symmetric Chamfer objective within a coarse-to-fine schedule. On expert-annotated cases, our method reduced landmark errors, improved posterior

alignment, and achieved interpretable curve-level agreement. Future work will enhance robustness for thin structures with orientation/curvature cues, extend to multi-view settings for better depth constraints, and incorporate uncertainty-aware weighting for noisy contours, validated through multi-center studies.

VII. Acknowledgments

References

[1] J. A. Maintz and M. A. Viergever, "A survey of medical image registration," *Medical image analysis*, vol. 2, no. 1, pp. 1–36, 1998.

[2] B. Zitova and J. Flusser, "Image registration methods: a survey," *Image and vision computing*, vol. 21, no. 11, pp. 977–1000, 2003.

[3] W. M. Wells III, P. Viola, H. Atsumi, S. Nakajima, and R. Kikinis, "Multi-modal volume registration by maximization of mutual information," *Medical image analysis*, vol. 1, no. 1, pp. 35–51, 1996.

[4] G. P. Penney, J. Weese, J. A. Little, P. Desmedt, D. L. Hill *et al.*, "A comparison of similarity measures for use in 2-d-3-d medical image registration," *IEEE transactions on medical imaging*, vol. 17, no. 4, pp. 586–595, 1998.

[5] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens, "Multimodality image registration by maximization of mutual information," *IEEE transactions on Medical Imaging*, vol. 16, no. 2, pp. 187–198, 2002.

[6] P. Viola and W. M. Wells III, "Alignment by maximization of mutual information," *International journal of computer vision*, vol. 24, no. 2, pp. 137–154, 1997.

[7] H. Jiang, K. S. Holton, and R. A. Robb, "Image registration of multimodality 3d medical images by chamfer matching," in *Biomedical Image Processing and Three-Dimensional Microscopy*, vol. 1660. SPIE, 1992, pp. 356–366.

[8] D. F. Swinehart, "The beer-lambert law," *Journal of chemical education*, vol. 39, no. 7, p. 333, 1962.

[9] J. Rino Neto, J. B. d. Paiva, G. V. Queiroz, M. F. Attizzani, and H. Miasiro Junior, "Evaluation of radiographic magnification in lateral cephalograms obtained with different x-ray devices: experimental study in human dry skull," *Dental Press Journal of Orthodontics*, vol. 18, pp. 17e1–17e7, 2013.

[10] L. Levine and M. Levine, "Drrgenerator: a three-dimensional slicer extension for the rapid and easy development of digitally reconstructed radiographs," *Journal of Clinical Imaging Science*, vol. 10, p. 69, 2020.

[11] M. Unberath, J.-N. Zaech, S. C. Lee, B. Bier, J. Fotouhi, M. Armand, and N. Navab, "Deepdrr–a catalyst for machine learning in fluoroscopy-guided procedures," in *International conference on medical image computing and computer-assisted intervention*. Springer, 2018, pp. 98–106.

[12] D. P. Huttenlocher, G. A. Klanderman, and W. J. Rucklidge, "Comparing images using the hausdorff distance," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 15, no. 9, pp. 850–863, 2002.

[13] P. J. Besl and N. D. McKay, "Method for registration of 3-d shapes," in *Sensor fusion IV: control paradigms and data structures*, vol. 1611. Spie, 1992, pp. 586–606.

[14] A. Myronenko and X. Song, "Point set registration: Coherent point drift," *IEEE transactions on pattern analysis and machine intelligence*, vol. 32, no. 12, pp. 2262–2275, 2010.

[15] M. Serafin, B. Baldini, F. Cabitza, G. Carrafiello, G. Baselli, M. Del Fabbro, C. Sforza, A. Caprioglio, and G. M. Tartaglia, "Accuracy of automated 3d cephalometric landmarks by deep learning algorithms: systematic review and meta-analysis," *La radiologia medica*, vol. 128, no. 5, pp. 544–555, 2023.

[16] J. Hendrickx, R. S. Gracea, M. Vanheers, N. Winderickx, F. Preda, S. Shujaat, and R. Jacobs, "Can artificial intelligence-driven cephalometric analysis replace manual tracing? a systematic review and meta-analysis," *European Journal of Orthodontics*, vol. 46, no. 4, p. cjae029, 2024.

[17] Y. Lee, J.-H. Pyeon, S.-H. Han, N. J. Kim, W.-J. Park, and J.-B. Park, "A comparative study of deep learning and manual methods for identifying anatomical landmarks through cephalometry and cone-beam computed tomography: A systematic review and meta-analysis," *Applied Sciences*, vol. 14, no. 16, p. 7342, 2024.

[18] S. Sun, K. Han, C. You, H. Tang, D. Kong, J. Naushad, X. Yan, H. Ma, P. Khosravi, J. S. Duncan *et al.*, "Medical image registration via neural fields," *Medical Image Analysis*, vol. 97, p. 103249, 2024.

[19] J. Yang, M. Zhao, Y. Wu, and X. Jia, "Accurate and robust registration of low overlapping point clouds," *Computers & Graphics*, vol. 118, pp. 146–160, 2024.

[20] V. Gopalakrishnan, N. Dey, and P. Golland, "Intraoperative 2d/3d image registration via differentiable x-ray rendering," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2024, pp. 11 662–11 672.

[21] M. Leskovar, M. Heyland, A. Trepczynski, and S. Zachow, "Comparison of global and local optimization methods for intensity-based 2d–3d registration," *Computers in Biology and Medicine*, vol. 186, p. 109574, 2025.

[22] Y. Miao, T. Wu, T. Chen, S. Li, J. Jiang, J. Yang, A. Stefanidis, L. Yu, and J. Su, "Dentalsplat: Dental occlusion novel view synthesis from sparse intra-oral photographs," *arXiv preprint arXiv:2511.03099*, 2025.

[23] J. Tamayo-Quintero, J. Gómez-Mendoza, and S. Guevara-Pérez, "Dentalarch: Ai-based arch shape detection in orthodontics," *Applied Sciences*, vol. 14, no. 6, p. 2567, 2024.

[24] D. H. Ajmera, P. Singh, Y. Y. Leung, B. S. Khambay, and M. Gu, "Establishment of the mid-sagittal reference plane for three-dimensional assessment of facial asymmetry: a systematic review: establishment of the mid-sagittal reference plane: a systematic review," *Clinical Oral Investigations*, vol. 28, no. 4, p. 242, 2024.

[25] B. Du, K. Li, Z. Shen, Y. Cheng, J. Yu, Y. Pan, Z. Huang, F. Hu, X. Rausch-Fan, Y. Zhu *et al.*, "Feasibility of occlusal plane in predicting the changes in anteroposterior mandibular position: a comprehensive analysis using deep learning-based three-dimensional models," *BMC Oral Health*, vol. 25, no. 1, p. 42, 2025.

[26] H.-H. Kim, S. Choi, Y.-I. Chang, W.-J. Yi, and S.-J. Ahn, "Developing a three-dimensional statistical shape model of normal dentition using an automated algorithm and normal samples," *Clinical Oral Investigations*, vol. 27, no. 2, pp. 759–772, 2023.

[27] H. Mouncif, A. Kassimi, T. Bertin Gardelle, H. Tairi, and J. Riffi, "3d tooth identification for forensic dentistry using deep learning," *BMC Oral Health*, vol. 25, no. 1, p. 665, 2025.

[28] D.-H. Kim, Y. Li, and K. C. Lee, "Three-dimensional cephalometric evaluation of the craniofacial morphology in korean population utilizing cone-beam computed tomography," *Korean Journal of Orthodontics*, vol. 55, no. 4, pp. 254–265, 2025.

[29] S. Wang, H. Li, J. Li, Y. Zhang, and B. Zou, "Automatic analysis of lateral cephalograms based on multiresolution decision tree regression voting," *Journal of healthcare engineering*, vol. 2018, no. 1, p. 1797502, 2018.

[30] S. Wang, H. Li, B. Zou, and W. Zhang, "A novel contour-based registration of lateral cephalogram and profile photograph," *Computerized Medical Imaging and Graphics*, vol. 63, pp. 9–23, 2018.

[31] O. Khattak, A. S. Hashem, M. S. Alqarni, R. A. S. Almufarrij, A. Y. Siddiqui, R. Anis, S. Ahmad, M. A. Fareed, O. S. Alothmani, L. H. S. Alkhershawy *et al.*, "Deep learning applications in dental image-based diagnostics: A systematic review," in *Healthcare*, vol. 13, no. 12. MDPI, 2025, p. 1466.

[32] M. A. Khalid, K. Zulfiqar, U. Bashir, A. Shaheen, R. Iqbal, Z. Rizwan, G. Rizwan, and M. M. Fraz, "A benchmark dataset for automatic cephalometric landmark detection and cvm stage classification," *Scientific Data*, vol. 12, no. 1, p. 1336, 2025.

[33] K. Takahashi, Y. Shimamura, C. Tachiki, Y. Nishii, and M. Hagiwara, "Cephalometric landmark detection without x-rays combining coordinate regression and heatmap regression," *Scientific reports*, vol. 13, no. 1, p. 20011, 2023.