Towards the Digital Dental Models: High-Fidelity CBCT-to-IOS Fusion via Anatomy-Aware Geometric Transformers

Hanqing Zhou[1†], Xiner Ning[2†], Chunxia Ren[1], Hong Su[2✉], and Li Xiao[1✉]

[1] Beijing University of Posts and Telecommunications, Beijing, China
[2] Peking University School and Hospital of Stomatology, Beijing, China
andrewxiao@bupt.edu.cn, suhong@pkuss.bjmu.edu.cn

Abstract. The precise fusion of Cone Beam Computed Tomography (CBCT) and 3D Intraoral Scans (IOS) is essential for constructing digital dental models in stomatology. Despite its importance, automated alignment remains challenging due to the scarcity and limited accessibility of open-source datasets, discrepancies in cross-source resolution, and the complex topology of dental structures. To address these limitations, we introduce a newly curated multimodal dataset comprising 425 patients, along with the Anatomy-Aware Geometric Transformers (AAGT)—a novel registration framework specifically designed for dental imaging fusion. Unlike conventional models, AAGT incorporates anatomy-aware priors to effectively capture high-frequency anatomical features. By integrating surface topologies and spatial attributes through a topology-enhanced attention mechanism, our framework emphasizes discriminative landmarks and robustly mitigates challenges posed by metal artifacts and resolution inconsistencies. Extensive experiments demonstrate that AAGT achieves state-of-the-art performance, and external validation confirms its effectiveness in preserving distinct dental anatomy, enabling reliable and fully automated clinical workflows. The data and code are publicly available at https://github.com/AI4MyBUPT/DIgital-dental.

Keywords: Dental Registration · Multi-Modality Fusion · Geometric Deep Learning.

1 Introduction

The construction of Digital Dental Models—entailing the precise reconstruction of patient anatomy—is fundamental to modern stomatology [14,19,20]. This process relies on fusing multi-modal data: Cone Beam Computed Tomography (CBCT) for internal volumetric structures (e.g., tooth roots) and 3D Intraoral Scans (IOS) for high-precision surface details. Accurate alignment of these geometrically distinct datasets is a prerequisite for clinical safety; for instance, in

† These authors contributed equally to this work.
✉ Corresponding authors: andrewxiao@bupt.edu.cn, suhong@pkuss.bjmu.edu.cn

implant surgery, registration errors can compromise surgical guide precision and risk nerve damage [8]. However, current clinical workflows [15] are highly dependent on manual operation or traditional algorithms, which suffer from inefficiency, operator variability, and sensitivity to initialization. Furthermore, current methods [4,25] often fail to robustly handle the significant density discrepancies and metal artifacts inherent in dental data, highlighting the urgent need for a robust, automated solution capable of leveraging anatomical geometric priors.

Clinically, registration remains dominated by manual annotation and traditional geometric algorithms [1,5,23] like the Iterative Closest Point (ICP) [3]. However, ICP's non-convex nature makes it highly sensitive to initial misalignments and prone to local optima. Moreover, ICP lacks robustness against occlusions, density variations, and topological artifacts common in IOS-CBCT matching. Although Random Sample Consensus (RANSAC) [7] improves noise tolerance, its computational cost is prohibitive for high-dimensional dental point clouds. Recent AI-based methods [16,22] have attempted to automate this process but often retain traditional algorithmic components for the core alignment task, failing to fully resolve the bottlenecks of limited accuracy and efficiency.

To overcome these limitations, formulating registration as a point cloud problem allows leveraging deep learning advances. State-of-the-art methods [9,18,24,28] have demonstrated success in generic scenes by modeling correspondences or encoding geometric structures. However, directly applying these generic models to dentistry is suboptimal. Teeth exhibit unique topological structures characterized by specific curvature features and normal distributions. General-purpose geometric encoders tend to prioritize low-frequency semantic information, inadvertently smoothing these critical high-frequency anatomical details. This underutilization of fine-grained geometry leads to insufficient accuracy in complex occlusal scenarios, failing to meet the rigorous standards of clinical precision.

Addressing these gaps, we propose **Anatomy-Aware Geometric Transformers (AAGT)**, a fully automatic framework tailored for cross-source dental registration. Diverging from methods [18,24,28] dependent on ambiguous low-frequency semantics, AAGT explicitly exploits the rich surface topology. Specifically, the lightweight **Equivariant Saliency Weighting Module (ESWM)** encodes surface information to derive saliency weights. Subsequently, the **Anatomy-Aware Embedding (AAEmbedding)** captures topological priors, while the **Anatomy-Aware Transformer (AATransformer)** deeply fuses these with spatial contexts. This anatomically guided strategy directs focus toward highly discriminative anatomical features, ensuring robust alignment despite metal artifacts and resolution discrepancies. Our contributions are as follows:

1. We present the first end-to-end registration network for 3D IOS and CBCT images that explicitly incorporates surface priors into the deep learning pipeline.
2. We design **ESWM**, **AAEmbedding** and **AATransformer** to jointly learn semantic and structural features. This mechanism significantly enhances robustness against metal artifacts and resolution discrepancies, outperforming existing methods in accuracy and efficiency.

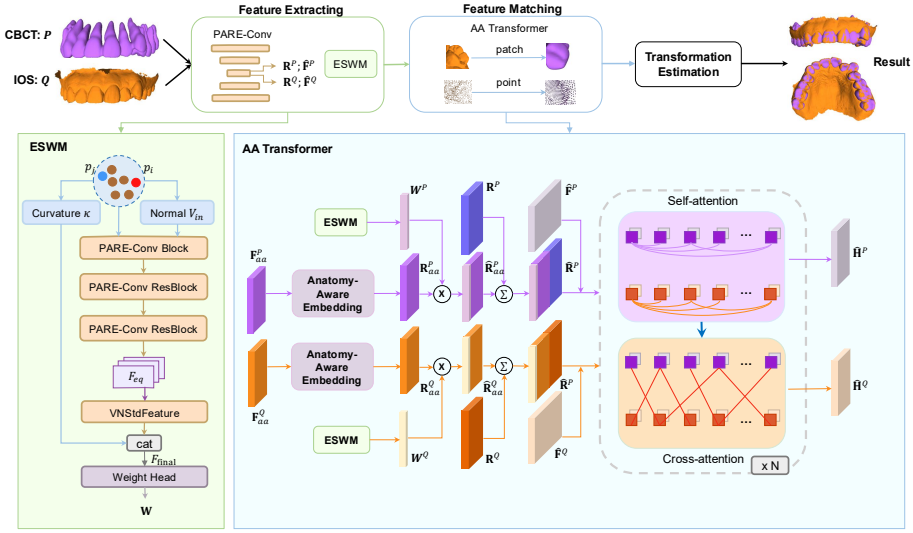

Fig. 1. The architecture of AAGT. It consists of three parts: Feature Extracting, Feature Matching, and Transformation Estimation. The lower left shows the ESWM part of the dual-stream strategy in Feature Extracting, and the lower right shows the AA-Transformer part in Feature Matching.

3. We release a challenging multi-modal clinical dataset. This addresses the critical scarcity and inaccessibility of open-source datasets caused by the prohibitive cost of acquiring strictly paired IOS-CBCT data and labor-intensive curation, with code and data publicly available.

2 Method

To ensure valid overlap, we isolate the dentition from CBCT via nnU-Net [11], eliminating non-corresponding anatomy (e.g., jaws, soft tissue). The source point cloud $P = \{p_i \in R^3\}_{i=1}^{N}$ is then reconstructed using Marching Cubes [13]. We aim to estimate a rigid transformation $T = \{R, t\}$ $(R \in SO(3), t \in R^3)$ aligning P to the IOS target $Q = \{q_i \in R^3\}_{i=1}^{M}$. The transformation can be solved by:

$$\min_{R,t} \sum_{(p^*_{x_i}, q^*_{y_i}) \in \mathcal{C}^*} \| R \cdot p^*_{x_i} + t - q^*_{y_i} \|_2^2. \tag{1}$$

Here $\mathcal{C}^*$ is the set of ground-truth correspondences composed of matched point pairs $(p^*_{x_i}, q^*_{y_i})$. As depicted in Fig. 1, we adopt a coarse-to-fine framework [18,26,27] where a dual-stream backbone PARE-Conv [24] and **ESWM** (Sec. 2.1) jointly encodes semantic and geometric features. To enhance discriminability, **AAEmbedding** (Sec. 2.2) explicitly integrates anatomy aware priors into the **AATransformer** (Sec. 2.3). Finally, the initial pose is generated via a Hypothesis Proposer and refined by optional ICP refinement.

2.1 Equivariant Saliency Weighting Module (ESWM)

To further suppress non-overlapping outliers and encode surface information, we design the lightweight **Equivariant Saliency Weighting Module (ESWM)**. ESWM employs siamese architecture leveraging the Vector Neuron (VN) paradigm [6] to ensure $SO(3)$ robustness. Given input points P and normals V_{in}, the rotation-equivariant embedding is computed via vector-space convolutions to preserve directionality:

$$F_{\text{eq}} = \Phi_{\text{Res}}\left(\Phi_{\text{Block}}(P, V_{in})\right), \tag{2}$$

where Φ_{Block} and Φ_{Res} denote PARE-Conv Block and ResBlocks [24]. Subsequently, F_{eq} is projected into a rotation-invariant space via a VN-Standardization layer (Ψ_{std}) to fully capitalize on the interplay between rotation equivariance and invariance, and then concatenated with the explicit surface curvature κ:

$$F_{\text{final}} = \text{Concat}\left(\Psi_{\text{std}}(F_{\text{eq}}), \kappa\right), \tag{3}$$

Finally, a shared MLP [17] outputs a probability weight w_i for each point:

$$w_i = \text{MLP}(F_{\text{final}}). \tag{4}$$

2.2 Anatomy-Aware Embedding (AAEmbedding)

Illustrated in Fig. 2, **Anatomy-Aware Embedding (AAEmbedding)** captures intrinsic dental topology by transcending the limitations of fixed-scale approaches [18,24]. Specifically, it employs learnable sensitivity parameters to adaptively fuse curvature and normal anatomically-aware details with spatial contexts, thereby resolving structural scale discrepancies.

Given a set of superpoints $\mathcal{P} = \{p_i\}_{i=1}^{N}$ with normals $\{n_i\}$ and curvature $\{c_i\}$, we compute three core primitives: Euclidean distance $d_{i,j} = \|p_i - p_j\|_2$,

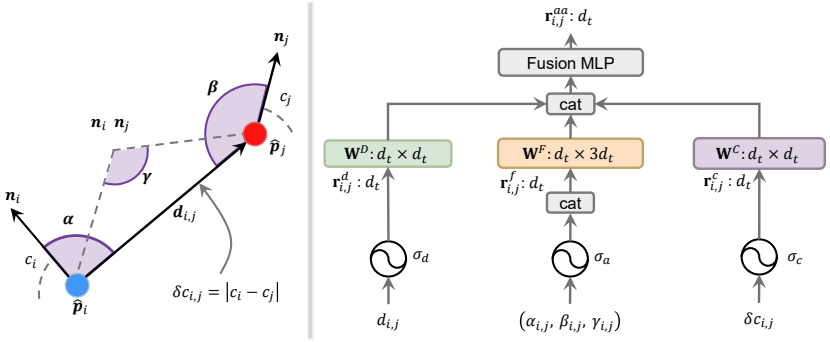

Fig. 2. Anatomy-Aware Embedding. Left is an illustration of the point pair feature angles and curvature difference. Right is the computation graph of the Anatomy-Aware Embedding.

angles $(\alpha_{i,j}, \beta_{i,j}, \gamma_{i,j})$, and curvature difference $\delta c_{i,j} = |c_i - c_j|$. Learnable log-scale parameters $\sigma = \{\sigma_d, \sigma_a, \sigma_c\}$ adaptively normalize these primitives, allowing the network to automatically adjust sensitivity to spatial and shape variations:

$$\mathbf{r}^d_{i,j} = \psi(d_{i,j}/e^{\sigma_d}), \quad \mathbf{r}^c_{i,j} = \psi(\delta c_{i,j}/e^{\sigma_c}), \tag{5}$$

where $\psi(\cdot)$ denotes the sinusoidal positional encoding function [21].

To preserve angular correlations within the local reference frame, we concatenate the embeddings of three angles prior to projection. Furthermore, a deep MLP fuses the distance, angle, and curvature embeddings. This design enables the distinction between close-range points with sharp edges and those on flat surfaces. The final embedding $\mathbf{r}^{aa}_{i,j}$ is:

$$\mathbf{r}^f_{i,j} = \mathcal{L}\left([\psi(\alpha_{i,j}/e^{\sigma_a}) \oplus \psi(\beta_{i,j}/e^{\sigma_a}) \oplus \psi(\gamma_{i,j}/e^{\sigma_a})]\right), \tag{6}$$

$$\mathbf{r}^{aa}_{i,j} = \mathrm{MLP}(\mathbf{r}^d_{i,j}\mathbf{W}^D \oplus \mathbf{r}^c_{i,j}\mathbf{W}^C \oplus \mathbf{r}^f_{i,j}\mathbf{W}^F). \tag{7}$$

where $\oplus$ denotes concatenation, $\mathcal{L}$ represents a linear projection and $\mathbf{W}^D$, $\mathbf{W}^C$, $\mathbf{W}^F$ are trainable weights. $\mathbf{r}^{aa}_{i,j}$ is subsequently fed into the **AATransformer** for attention learning.

2.3 Anatomy-Aware Transformer (AATransformer)

We introduce the **Anatomy-Aware Transformer (AATransformer)** to explicitly encode both rotation-invariant and equivariant surface geometry. Given input $H \in R^{N \times d_t}$, the output Z is computed via weighted aggregation:

$$z_i = \sum_{j=1}^{N} \alpha_{i,j}(h_j \mathbf{W}_V), \tag{8}$$

where attention weights $\alpha_{i,j}$ are derived via a row-wise Softmax on the alignment scores $e_{i,j}$. We inject geometric priors to resolve ambiguity in feature-poor regions, defining the alignment score $e_{i,j}$ as:

$$e_{i,j} = \frac{(h_i \mathbf{W}_Q)(h_j \mathbf{W}_K + \mathbf{r}^p_{i,j}\mathbf{W}_P + \mathbf{r}^{aa}_{i,j}\mathbf{W}_{AA} w_i)^T}{\sqrt{d_t}}. \tag{9}$$

Here, $\mathbf{W}_Q, \mathbf{W}_K, \mathbf{W}_V, \mathbf{W}_P, \mathbf{W}_{AA} \in R^{d_t \times d_t}$ represent the learnable projection matrices for Query, Key, Value [21], point-wise structure embedding [24], and anatomy-aware prior embedding, respectively. The terms w_i, $\mathbf{r}^{aa}_{i,j}$, and $\mathbf{r}^p_{i,j}$ integrate features from **ESWM**, **AAEmbedding**, and GeometricEmbedding [18], respectively, ensuring robustness against complex dental topology. Finally, to bridge local descriptor isolation, we employ Cross-Attention [18] to facilitate interaction between source P and target Q.

2.4 Loss Functions

To supervise the ESWM, we employ a weighted Binary Cross-Entropy (BCE) loss. Ground-truth labels $y_i \in \{0, 1\}$ are generated by transforming the source P via T_{gt}, where points within distance τ_{dist} are labeled positive. The loss is defined as:

$$L_{weight} = \sum_{X \in P} \frac{1}{|X|} \sum_{i=1}^{|X|} -[\alpha_i y_i \log(s_i) + (1 - y_i) \log(1 - s_i)] \tag{10}$$

where s_i denotes the predicted saliency and $\alpha_i = \lambda_{pos} \cdot y_i + (1 - y_i)$ applies a penalty factor λ_{pos} to upweight positive samples, compelling the network to focus on overlapping structures. The overall objective is then formulated as $L_{total} = L_c + L_f + L_r + L_{weight}$, where the first three terms follow PARE-Net [24].

3 Experiment

3.1 Experiment setup

Tasks and Datasets To address the challenges of registration in complex clinical scenarios, we construct and release a large-scale, challenging dataset for dental multi-modality registration. The de-identified dataset was collected at Peking University School and Hospital of Stomatology with IRB approval (PKUSSIRB-2025115199). This dataset comprises 425 subjects, consisting of paired maxillary and mandibular models, totaling 850 pairs of intraoral scans. Facilitating the fusion of Digital Dental Models, our dataset consists of paired CBCT and IOS scans acquired from the same patients. This configuration establishes a rigorous benchmark for cross-modality registration, challenging algorithms to resolve the metal artifacts and inherent resolution between volumetric reconstruction and optical surface acquisition. Following the standard protocols established in Geo [18], we partition the 425 subjects into training, validation, and testing sets with a ratio of 315 (630 pairs), 55 (110 pairs), and 55 (110 pairs), respectively.

Implementation Details Our framework was implemented in PyTorch and trained on a NVIDIA RTX 4090 GPU. The network parameters were optimized using the Adam optimizer [12] for 250 epochs. The weight decay was set to 1×10^{-4}. To comprehensively assess point cloud registration quality, we report standard geometric metrics: Inlier Ratio (IR), Relative Rotation Error (RRE), Relative Translation Error (RTE), Root Mean Square Error (RMSE), and Registration Recall (RR) [10]. Beyond geometric metrics, we also conducted a rigorous external validation to verify the potential of our method in clinical workflows.

3.2 Results

Quantitative Comparison. Table 1 presents the quantitative comparison against state-of-the-art methods. Our proposed AAGT demonstrates superior

Table 1. Quantitative results on our dataset. We compare the IR, RRE, RTE, RMSE, and RR with different methods. **Bold** text indicates the best results, and our method outperforms state-of-the-art baseline methods.

Methods	*IR* (%) ↑	*RRE* (°) ↓	*RTE* (mm) ↓	*RMSE* (mm) ↓	*RR* (%) ↑
Predator [9]	0.470	29.055	2.633	7.403	0.319
Geo [18]	0.545	1.679	0.187	0.503	0.982
PARE [24]	0.604	0.256	0.095	0.118	0.991
AAGT(ours)	**0.635**	**0.226**	**0.078**	**0.104**	**0.999**

performance across most key metrics. Specifically, AAGT achieves the highest IR of **0.635**, indicating its exceptional capability in identifying reliable correspondences even in dental scans with limited overlapping regions. Crucially, for clinical applications where precision is paramount, our method significantly reduces registration errors. Compared to the second-best method (PARE), AAGT reduces the RRE by **11.7%** and RTE by **17.9%**. The RMSE of **0.104** and RR of **0.999** suggest that our registration accuracy reaches the sub-millimeter level, which satisfies the strict requirements for orthodontic treatment planning.

Qualitative Visualization and Error Distribution Due to the scarcity of deep learning methods specifically designed for IOS-CBCT registration, we benchmark our proposed method against the current SOTA models in generic 3D point cloud registration, namely Geo [18] and PARE [24]. To comprehensively assess registration fidelity, we present a multi-view visualization analysis in Fig. 3 comparing AAGT against these strong baselines. We first visualize the spatial distribution of RMSE and KD-Tree [2] distance weighted error on both maxillary and mandibular models (a-b). The deviation error is color-coded, with blue representing negligible error (< 0.06 mm) and red indicating significant misalignment (> 0.24 mm). As illustrated, baseline methods exhibit considerable errors, particularly around the gingival margins and molar cusps, where curvature changes sharply. In contrast, AAGT achieves a uniformly low-error alignment across the entire dental arch. Clinically, this high precision on the occlusal surfaces is critical, as it guarantees the accurate fabrication of orthodontic appliances and ensures proper bite function post-treatment.

To verify internal anatomical correspondence, we projected the registered IOS contours back onto the CBCT slices (c-d) as rigorous external validation. We performed external validation using data containing orthodontic brackets. As evident in the CBCT slices, the presence of high-density metal brackets induces severe metal artifacts, which obscure distinct tooth boundaries and introduce significant noise. Meanwhile, we selected two challenging Regions of Interest: the Anterior Region (high aesthetic requirement) and the Molar Region (high occlusion stability requirement). Visual inspection reveals that the contours generated by AAGT tightly delineate the tooth boundaries in the CT images, achieving near-perfect overlap with the anatomical structures. Conversely, Geo [18] and

Table 2. Ablation study on our dataset. The table is split by study groups (ESWM & AATransformer and Embedding module). **Bold** text indicates the best results.

Type		*IR*(%) ↑	*RRE*(°) ↓	*RTE*(mm) ↓	*RMSE*(mm) ↓
baseline		0.604	0.256	0.095	0.118
w/AATr, w/o ESWM		0.614	0.343	0.129	0.161
w/AATr, w/ ESWM		**0.627**	**0.237**	**0.094**	**0.115**
Extra_Points	Ori_Points	*IR*(%) ↑	*RRE*(°) ↓	*RTE*(mm) ↓	*RMSE*(mm) ↓
w/ GeoEmbed	w/ GeoEmbed	0.627	0.237	0.094	0.115
w/ AAEmbed	w/ AAEmbed	0.601	0.257	0.090	0.120
w/ AAEmbed	w/ GeoEmbed	**0.635**	**0.226**	**0.078**	**0.104**

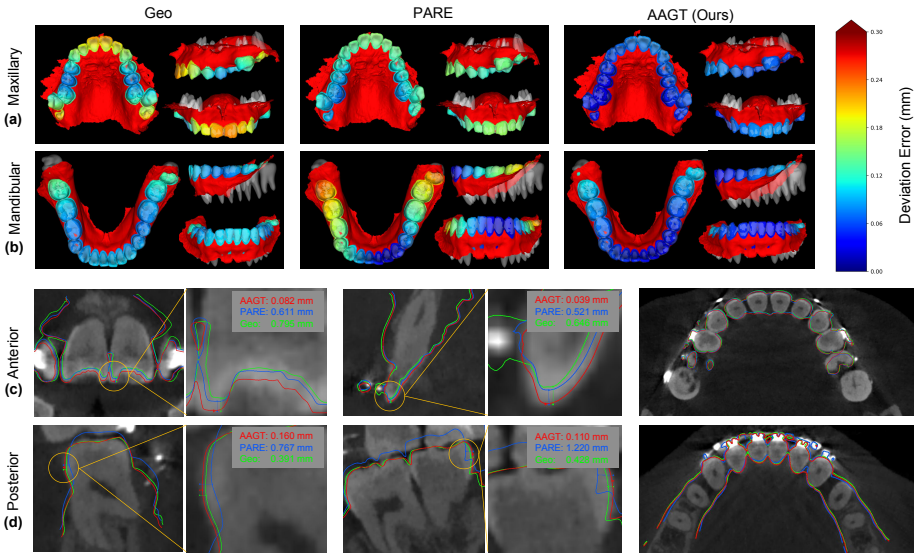

Fig. 3. Visual comparison between our method, Geo [18], and PARE [24]. (a)-(b) are the heatmaps of RMSE and KD-Tree [2] distance weighted error for the maxillary and mandibular dentition, respectively; (c)-(d) are the results of external validation. We projected the registered IOS onto the CT image and compared the registration errors.

PARE [24] show visible gaps and drift, with deviations exceeding 0.6mm in some areas. This further confirms that our method achieves sub-pixel level accuracy, providing reliable data fusion for root-crown integration in stomatology.

Table 2 details our ablation studies. Removing individual components consistently degrades performance, confirming their indispensability for robust registration. Notably, applying AAEmbedding directly to raw superpoints with reconstructed curvature and normals yielded suboptimal results. We attribute this to sparse superpoints lacking local manifold continuity; estimating differential

geometry on such centroids introduces high-frequency noise and normal perturbations, corrupting the embedding space and degrading registration fidelity.

4 Conclusion

This paper introduces **AAGT**, a geometry-guided framework for automated dental image registration that facilitates the precise construction of digital dental models. By explicitly embedding Anatomy Aware priors into a transformer architecture, our method effectively preserves high-frequency anatomical details. Experiments on a large-scale clinical dataset demonstrate that AAGT achieves state-of-the-art performance and exhibits superior robustness against metal artifacts and resolution discrepancies compared to existing methods. Meanwhile, a challenging multi-modal clinical dataset will be publicly available.

Acknowledgments. This research was supported by the Fundamental Research Funds for the Beijing University of Posts and Telecommunications under Grant 2025AI4S19, the National Natural Science Foundation of China (82571131) and Beijing Municipal Natural Science Foundation (L252199, L242134, L242059).

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Bai, S., Liao, S., Han, F., Peng, X., Li, Z., Yang, W., Wu, R.: Automatic real-time space registration application for simulating dental and maxillofacial surgery. Journal of Craniofacial Surgery **33**(6), 1698–1704 (2022)
2. Bentley, J.L.: Multidimensional binary search trees used for associative searching. Communications of the ACM **18**(9), 509–517 (1975)
3. Besl, P.J., McKay, N.D.: Method for registration of 3-d shapes. In: Sensor fusion IV: control paradigms and data structures. vol. 1611, pp. 586–606. Spie (1992)
4. Codari, M., de Faria Vasconcelos, K., Ferreira Pinheiro Nicolielo, L., Haiter Neto, F., Jacobs, R.: Quantitative evaluation of metal artifacts using different cbct devices, high-density materials and field of views. Clinical oral implants research **28**(12), 1509–1514 (2017)
5. Dai, N., Li, D., Yang, X., Cheng, C., Sun, Y.: Research and primary evaluation of an automatic fusion method for multisource tooth crown data. International Journal for Numerical Methods in Biomedical Engineering **33**(11), e2878 (2017)
6. Deng, C., Litany, O., Duan, Y., Poulenard, A., Tagliasacchi, A., Guibas, L.J.: Vector neurons: A general framework for so (3)-equivariant networks. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 12200–12209 (2021)
7. Fischler, M.A., Bolles, R.C.: Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. Communications of the ACM **24**(6), 381–395 (1981)

8. Flügge, T., Derksen, W., Te Poel, J., Hassan, B., Nelson, K., Wismeijer, D.: Registration of cone beam computed tomography data and intraoral surface scans–a prerequisite for guided implant surgery with cad/cam drilling guides. Clinical Oral Implants Research **28**(9), 1113–1118 (2017)
9. Huang, S., Gojcic, Z., Usvyatsov, M., Wieser, A., Schindler, K.: Predator: Registration of 3d point clouds with low overlap. In: Proceedings of the IEEE/CVF Conference on computer vision and pattern recognition. pp. 4267–4276 (2021)
10. Huang, X., Mei, G., Zhang, J., Abbas, R.: A comprehensive survey on point cloud registration. arXiv preprint arXiv:2103.02690 (2021)
11. Isensee, F., Jaeger, P.F., Kohl, S.A., Petersen, J., Maier-Hein, K.H.: nnu-net: a self-configuring method for deep learning-based biomedical image segmentation. Nature methods **18**(2), 203–211 (2021)
12. Kingma, D.P.: Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980 (2014)
13. Lorensen, W.E., Cline, H.E.: Marching cubes: A high resolution 3d surface construction algorithm. In: Seminal graphics: pioneering efforts that shaped the field, pp. 347–353 (1998)
14. Mohammad-Rahimi, H., Nadimi, M., Rohban, M.H., Shamsoddin, E., Lee, V.Y., Motamedian, S.R.: Machine learning and orthodontics, current trends and the future opportunities: A scoping review. American Journal of Orthodontics and Dentofacial Orthopedics **160**(2), 170–192 (2021)
15. Park, J.H., Hwang, C.J., Choi, Y.J., Houschyar, K.S., Yu, J.H., Bae, S.Y., Cha, J.Y.: Registration of digital dental models and cone-beam computed tomography images using 3-dimensional planning software: Comparison of the accuracy according to scanning methods and software. American Journal of Orthodontics and Dentofacial Orthopedics **157**(6), 843–851 (2020)
16. Preda, F., Nogueira-Reis, F., Stanciu, E.M., Smolders, A., Jacobs, R., Shaheen, E.: Validation of automated registration of intraoral scan onto cone beam computed tomography for an efficient digital dental workflow. Journal of Dentistry **149**, 105282 (2024)
17. Qi, C.R., Su, H., Mo, K., Guibas, L.J.: Pointnet: Deep learning on point sets for 3d classification and segmentation. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 652–660 (2017)
18. Qin, Z., Yu, H., Wang, C., Guo, Y., Peng, Y., Ilic, S., Hu, D., Xu, K.: Geotransformer: Fast and robust point cloud registration with geometric transformer. IEEE Transactions on Pattern Analysis and Machine Intelligence **45**(8), 9806–9821 (2023)
19. Reduwan, N.H., Aziz, A.A., Mohd Razi, R., Abdullah, E.R.M.F., Mazloom Nezhad, S.M., Gohain, M., Ibrahim, N.: Application of deep learning and feature selection technique on external root resorption identification on cbct images. BMC Oral Health **24**(1), 252 (2024)
20. Revilla-León, M., Gómez-Polo, M., Sallorenzo, A., Rutkunas, V., Pérez-Barquero, J.A., Fernández-Estevan, L., Agustín-Panadero, R., Kois, J.C.: A novel protocol for evaluating and classifying tooth preparations being scanned by using intraoral scanners: A critical review of how tooth preparation factors affect intraoral scanning accuracy. Journal of Dentistry **157**, 105740 (2025)
21. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł., Polosukhin, I.: Attention is all you need. Advances in neural information processing systems **30** (2017)

22. Woo, S., Lee, S., Chae, J., Rim, J., Lee, J., Seo, J., Lee, C.: Automatic matching of computed tomography and stereolithography data. Computer Methods and Programs in Biomedicine **175**, 215–222 (2019)
23. Wu, D., Jiang, J., Wang, J., Zhou, S., Qian, K.: Accuracy evaluation of dental cbct and scanned model registration method based on pulp horn mapping surface: an in vitro proof-of-concept. BMC Oral Health **24**(1), 827 (2024)
24. Yao, R., Du, S., Cui, W., Tang, C., Yang, C.: Pare-net: Position-aware rotation-equivariant networks for robust point cloud registration. In: European Conference on Computer Vision. pp. 287–303. Springer (2024)
25. Yi, C., Li, S., Wen, A., Wang, Y., Zhao, Y., Zhang, Y.: Digital versus radiographic accuracy evaluation of guided implant surgery: an in vitro study. BMC Oral health **22**(1), 540 (2022)
26. Yu, H., Li, F., Saleh, M., Busam, B., Ilic, S.: Cofinet: Reliable coarse-to-fine correspondences for robust pointcloud registration. Advances in Neural Information Processing Systems **34**, 23872–23884 (2021)
27. Yu, H., Qin, Z., Hou, J., Saleh, M., Li, D., Busam, B., Ilic, S.: Rotation-invariant transformer for point cloud matching. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 5384–5393 (2023)
28. Yu, J., Ren, L., Zhang, Y., Zhou, W., Lin, L., Dai, G.: Peal: Prior-embedded explicit attention learning for low-overlap point cloud registration. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 17702–17711 (2023)