3D Dynamic Prediction of Missing Teeth in Diverse Patterns via Centroid-prompted Diffusion Model

Zongrui Ji[1,2], Na Li[3], Peng Xue[4], Yi Dong[1,2,✉], and Lei Ma[1,2,✉]

[1] Department of Control Science and Engineering, School of Electronics and Information Engineering, Tongji University, Shanghai,201804,China
[2] State Key Laboratory of Autonomous Intelligent Unmanned Systems, Tongji University, Shanghai, 200092, China
[3] Stomatological Hospital of Henan Province, The First Affiliated Hospital of Zhengzhou University, School and Hospital of Stomatology of Zhengzhou University, Zhenzhou, 450052, China
[4] School of Mechanical, Electrical and Information Engineering, Shandong University, Weihai, 264209, China
yidong@tongji.edu.cn, ma_lei@tongji.edu.cn

Abstract. Dental implantation restores missing teeth through surgical insertion of artificial roots, relying on preoperative digital planning to ensure precision and efficiency. However, critical challenges persist in virtual tooth positioning: this process demands extensive clinical expertise and time-consuming manual adjustments due to ambiguous anatomical references from missing teeth. To address these limitations, we propose a unified framework that accurately predicts the original three dimensional (3D) shapes and positions of missing teeth in diverse patterns, enabling anatomy-aware preoperative planning. Our proposal introduces two technical innovations: (1) A dynamic iterative generation strategy is proposed to progressively predict multiple missing teeth one by one using a target tooth identification module, accommodating arbitrary tooth loss patterns without case-specific retraining; (2) A tooth-centroid-prompted conditional diffusion model is developed to leverage geometric constraints from predicted tooth centroid and adjacent teeth to generate high-fidelity point cloud reconstructions. Extensive experiments show that our model outperforms conventional U-net based framework in predicting multiple missing teeth, achieving a prediction accuracy of 1.30mm (Chamfer Distance) and an angular error of 5.42 degrees. This improvement has the potential to enhance the accuracy and efficiency of dental implant planning by providing precise anatomical references for missing teeth, potentially revolutionizing digital dentistry workflows.

Keywords: Digital dentistry · Dental implant · Deep learning · Diffusion model · Point cloud generation

1 Introduction

Dental implantation has become a foundational component of contemporary restorative dentistry, providing reliable functional and aesthetic results through the surgical insertion of artificial roots into the jawbone [9]. The procedure relies critically on preoperative digital planning, which integrates advanced imaging modalities like cone-beam computed tomography (CBCT) and computer-aided design (CAD) software to optimize implant positioning and prosthetic outcomes [15]. Central to this workflow is virtual tooth positioning, a process that determines the optimal three-dimensional (3D) implant location, angulation, and depth by simulating anatomical relationships between missing teeth and adjacent structures [3]. Despite advancements, virtual tooth positioning remains a time-intensive and expertise-dependent task, as its efficacy and accuracy are still constrained by the need for unambiguous anatomical references, which are often absent in cases of tooth loss [16,4]. In current clinical practise, clinicians must manually perform tooth positioning through iterative trial-and-error adjustments, a process that demands profound familiarity with occlusal dynamics and regional bone morphology [17]. This challenge is exacerbated in cases of multiple missing teeth, where inter-tooth spatial relationships and arch continuity are disrupted, further complicating anatomical inference [12].

Accurate 3D reconstruction of missing teeth, encompassing spatial coordinates, axial orientation, and morphological details, provides foundational references for precision implant planning, yet remains clinically elusive due to two inherent complexities [2]. First, dental anatomy exhibits intricate inter-tooth biomechanical relationships, where subtle variations in crown-root angulation or proximal contact points critically influence occlusal dynamics [12,13]. Second, tooth loss patterns span a high-dimensional spectrum, ranging from single-tooth gaps to complex multi-unit edentulism with non-linear spatial dependencies [13]. Despite preliminary efforts to address these challenges, existing methodologies suffer from fundamental limitations: Panoramic radiograph analysis [14] and neighbor-guided localization [2] reduce the problem to 2D detection or heuristic positioning, neglecting 3D shape reconstruction. Conventional deep learning methods like U-net and its variants often fail to capture the intricate morphological variability of dental structures or enforce biomechanically valid spatial constraints [6,7,8]. Moreover, most frameworks require case-specific retraining when confronted with arbitrary tooth loss configurations, undermining their practicality in real-world clinical workflows. These limitations highlight an unmet need for robust, generalizable systems capable of reconstructing both 3D shapes and positions of multiple missing teeth while preserving anatomical coherence with adjacent dentition.

To address these challenges, we propose a unified framework for anatomy-aware preoperative planning that accurately predicts the original 3D shapes and spatial positions of missing teeth in diverse patterns. Our methodology introduces two interconnected innovations designed to overcome limitations in existing computational approaches. First, we develop a dynamic iterative generation strategy that progressively reconstructs multiple missing teeth one by one through a neighborhood-aware updates. This mechanism autonomously identifies prediction sequences of the multiple missing teeth, effectively decomposing complex multi-tooth reconstruction tasks into sequential single-tooth predictions. Crucially, this architecture adapts to arbitrary tooth loss patterns without requiring case-specific retraining, significantly enhancing clinical adaptability. Second, we propose a tooth-centroid-prompted conditional diffusion model that generates 3D point clouds by leveraging geometric constraints from predicted tooth centroid and its adjacent teeth. This dual-conditioning strategy ensures high-fidelity reconstructions while preserving occlusal functionality. These two innovations create a robust framework for accurately predicting missing teeth in any loss pattern while maintaining clinical flexibility. This may open new possibilities for enhancing the quality and precision of dental implantation in clinical practice.

2 Method

2.1 Dynamic Iterative Generation Strategy

In this study, we propose a unified framework to iteratively predict multiple missing teeth using a dynamic iterative generation strategy (DIGS). The core of DIGS is a target tooth identification (TTI) module, which dynamically selects a target missing tooth for prediction in current iteration. As shown in Fig. 1, given a dental model with three missing teeth, the TTI module first selects a missing tooth for prediction. The centroid of the target tooth is predicted and merged with its adjacent teeth as input to a conditional point cloud diffusion model. After predicting the first tooth, we update the dental model and iterate the process for the remaining missing teeth. Mesh model of the predicted tooth point cloud is reconstructed using Poisson surface reconstruction algorithm.

The TTI module ensures that the multiple missing teeth are predicted one by one, accommodating various tooth loss patterns without case-specific retraining. Specifically, given a dental model reconstructed from labeled CBCT segmentation images, we first identify the tooth numbers of the missing teeth based on the FDI table [1]. We then analyze their adjacent teeth, categorizing them into direct-adjacent and indirect-adjacent teeth. Direct-adjacent teeth are the immediate neighbors of the target tooth, while indirect-adjacent teeth are the immediate neighbors of the direct-adjacent ones. For a standard dental model with 28 teeth, each tooth (except the second molars) has five direct-adjacent and four indirect-adjacent teeth in two sides, while second molars (i.e., 17, 27, 37, 47) have three direct-adjacent and two indirect-adjacent teeth. For a missing tooth with four direct-adjacent teeth, we prioritize predicting it by feeding the point cloud of its direct-adjacent teeth into a multiple-missing-teeth (MMT) diffusion model, which is trained to predict more than two missing teeth in the input adjacent teeth point cloud. If no missing tooth with four direct-adjacent teeth exists, we include indirect-adjacent teeth in one side and prioritize the missing tooth with four adjacent teeth. This design ensures the consistency of input data for the MMT diffusion model. For the last missing tooth, we treat it

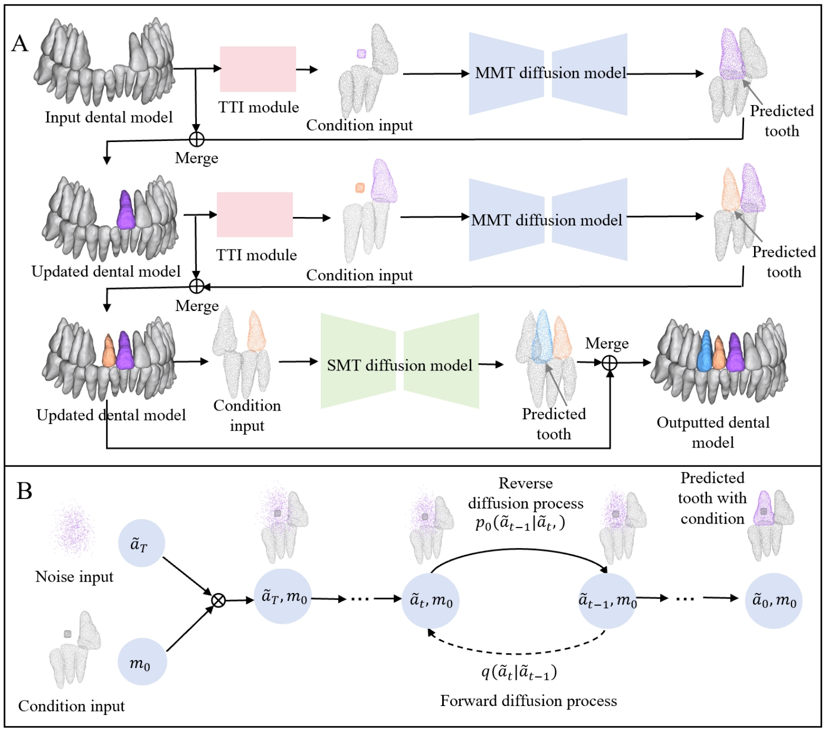

Fig. 1. (A) The proposed framework for precise prediction of multiple missing teeth. (B) Tooth-centroid-prompted conditional diffusion model.

as a single tooth, using its five adjacent teeth as input to a single-missing-tooth (SMT) diffusion model trained for single-tooth prediction, as shown in Fig. 1. After predicting one missing tooth, we update the input dental model by including the predicted tooth, and continue to perform the TTI module to select the next target missing tooth to predict until all the missing teeth are predicted.

2.2 Tooth-centroid-prompted Diffusion Model

To improve prediction accuracy and ensure the network identifies the correct tooth to predict in cases of multiple missing teeth, we propose a tooth-centroid-prompted diffusion model. In this approach, we first predict the centroid of the missing tooth based on its adjacent teeth, then feed both the predicted centroid and the adjacent teeth into the diffusion model as conditional input. This approach enhances the model's ability to accurately predict the missing tooth by leveraging both its spatial position and its anatomical context.

Tooth Centroid Prediction To accurately generate the centroid of the target missing tooth, we employ a tooth centroid regression method based on its adjacent teeth determined by the TTI module. This method adopts the Random Forest Regression (RFR) algorithm to perform this task. The specific steps are as follows: First, a random forest regression model is constructed and trained on the labeled dataset with the 3D centroids of the neighboring teeth as input and the missing tooth centroid as output. Then, for a target missing tooth, the trained regression model is used to predict its centroids based on its adjacent teeth centroids extracted from the 3D point cloud of the adjacent teeth. Next, using the predicted centroid as the origin, a small cluster of point cloud (500 points) is generated to represent the centroid. The centroid point cloud is then merged with the point cloud of the adjacent teeth as the conditional input to the MMT diffusion model.

Conditional Diffusion Model In this study, we utilize conditional point cloud diffusion models for accurate 3D reconstruction of missing teeth [18]. The point cloud of adjacent teeth serves as the conditional input for these models. Specifically, we train two diffusion models: MMT diffusion model for predicting multiple missing teeth, and SMT diffusion model for predicting a single missing tooth. For multiple missing teeth prediction, the adjacent teeth inputted into the MMT diffusion model may involve two missing teeth. To ensure the model accurately predicts the location of the target missing tooth, we incorporate the centroid of the missing tooth as a guide. This centroid is merged with the point cloud of the selected adjacent teeth to form the conditional input for MMT diffusion model, helping the model focus on the correct target tooth. In contrast, for single missing tooth prediction, we use the point cloud of five direct-adjacent teeth as the conditional input to SMT diffusion model. In this case, the position of the target missing tooth is clearly defined within the conditional point cloud, simplifying the model's task.

Both the MMT diffusion model and SMT diffusion model diffusion models are built upon the denoising diffusion probabilistic model (DDPM) framework [18]. The loss function of the diffusion model is defined as follows:

$$\mathcal{L}_t = \left\| \epsilon - \epsilon_\theta \left(\sqrt{\tilde{\alpha}_t} a_0 + \sqrt{1-\tilde{\alpha}_t}\epsilon, m_0, t \right) \right\|^2 . \tag{1}$$

During the training process, the forward diffusion process gradually adds noise to the ground truth of missing tooth a_0. Here, $\tilde{\alpha}_t$ is a coefficient related to the time step t that controls the degree of noise accumulation. And ϵ is the actual added noise. The model uses the neural network ϵ_θ, combined with the conditional information (i.e., the point cloud of adjacent teeth) m_0 and the time step t, to estimate the noise in the noisy data. The noisy data is given by the formula $(\sqrt{\tilde{\alpha}_t}a_0+\sqrt{1-\tilde{\alpha}_t}\epsilon)$. The loss function measures the difference between the actual noise ϵ and the estimated noise $\epsilon_\theta \left(\sqrt{\tilde{\alpha}_t} a_0 + \sqrt{1-\tilde{\alpha}_t}\epsilon, m_0, t \right)$. By minimizing this loss, the model continuously adjusts its parameters to learn the true distribution of the tooth point cloud. This enables it to accurately reconstruct 3D shape of missing tooth from the noise during the reverse diffusion process and accomplish the missing tooth prediction task.

3 Experimental Results

Dataset and Preprocessing This study employed CBCT scans collected from Stomatological Hospital of Henan Province Zhengzhou University, consisting of 130 healthy subjects, with the third molars excluded. Data preprocessing involved segmenting and labeling the teeth from the CBCT images, followed by reconstructing dental mesh models based on these segmentations. These models were then used to generate the training and testing datasets by simulating different tooth loss patterns, with 100 subjects allocated for training and the remaining 30 for testing.

Experimental Setup We carried out two experiments to evaluate the performance of the proposed framework on single missing tooth prediction and multiple missing teeth prediction, respectively. For comparison, we also conducted similar experiments using the proposed framework but with U-net as the backbone network on volumetric data (U-net framework) motivated by its widespread adoption in medical imaging and 3D reconstruction tasks, as well as its role as a standard benchmark in prior literature [6,7,8]. Note that the U-Net baseline in our experiments utilized a high-resolution volumetric grid of $0.4{\times}0.4{\times}0.4$mm (image size:$144{\times}144{\times}224$), minimizing discretization errors to a clinically acceptable level. In the evaluation process, we used three numerical metrics, Chamfer distance (CD) [10], Hausdorff distance (HD) [11], and Angular error (AE) [5], in order to systematically compare the performance of each method. And all metrics were computed directly on the raw point clouds with the same point number. In addition, we also performed qualitative visual comparisons to observe the effect of different methods under different tooth loss patterns.

In the single missing experiment, there were 2800 pairs of training data and 840 pairs of test data (28 pairs of data were obtained from each subject). In the experiments of multiple missing teeth (in this conference version, only the right side of teeth, 11-17 and 41-47, were used for method development and evaluation). We considered all the possible patterns of missing teeth with their corresponding adjacent teeth, and finally generated 3800 pairs of training data and 1140 pairs of test data. Each tooth has the point number of 1000, and point number of the centroid point cloud is 500. Therefore, the point clouds inputted to MMT and SMT diffusion models have 4500 point and 5000 points, respectively. The diffusion model MMT and SMT were trained for 2000 epoch, while the U-net framework was trained for 6000 epochs with the image size of $144 \times 144 \times 224$.

Results Our method demonstrated superior performance over the U-Net framework across all quantitative metrics in single missing tooth prediction, as quantified in Table 1. The U-Net framework achieved Chamfer distance (CD), Hausdorff distance (HD), and angular error (AE) values of 1.67 mm, 3.16 mm, and 7.16° respectively, while our method significantly reduced these errors to 1.26

Table 1. Comparison results between our method and U-net framework on the Single missing tooth prediction. CD, HD and AE represent Chamfer distance, Hausdorff distance and Angular error, respectively.

Tooth ID number		11	12	13	14	15	16	17
CD(mm)	U-net	1.72	1.88	1.46	1.60	1.61	1.99	1.99
	Ours	**1.20**	**1.14**	**1.22**	**1.16**	**1.15**	**1.40**	**1.54**
HD(mm)	U-net	2.99	3.24	3.40	2.81	2.65	4.31	3.70
	Ours	**1.96**	**2.16**	**2.17**	**2.35**	**2.20**	**2.67**	**2.56**
AE(degree)	U-net	5.83	4.80	5.03	4.60	5.19	6.67	6.08
	Ours	**3.69**	**3.08**	**3.53**	**3.08**	**4.24**	**4.40**	**4.53**
Tooth ID number		41	42	43	44	45	46	47
CD(mm)	U-net	1.23	1.24	1.43	1.22	1.34	1.94	1.71
	Ours	**0.98**	**1.10**	**1.28**	**1.14**	**1.19**	**1.32**	**1.45**
HD(mm)	U-net	2.09	2.22	3.03	2.07	2.64	4.54	3.59
	Ours	**1.53**	**2.03**	**2.16**	**1.95**	**2.15**	**2.71**	**2.58**
AE(degree)	U-net	4.30	4.12	5.20	4.37	6.34	6.47	7.24
	Ours	**3.38**	**2.46**	**3.66**	**3.89**	**4.14**	**4.24**	**4.03**

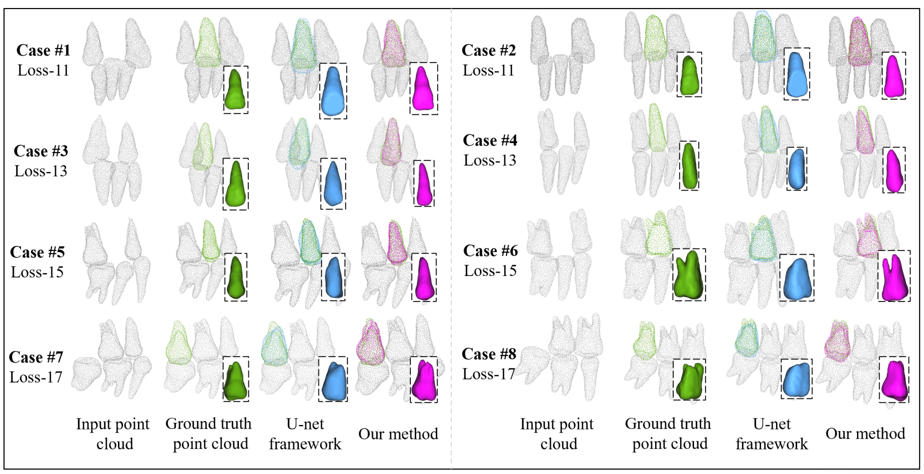

Fig. 2. Comparison results between our method and U-net framework in the case of single missing teeth. Tooth mesh models in the dotted boxes are reconstructed from the corresponding tooth point cloud using the Poisson surface reconstruction algorithm.

mm, 2.43 mm, and 4.48°. These enhancements confirm our method's improved precision in both spatial alignment and anatomical orientation prediction. As shown in Fig. 2, visual comparisons across four representative tooth types (incisor 11, cuspid 13, premolar 15, and second-molar 17) further substantiate these quantitative findings. Our method preserves complete morphological features while achieving exceptional detail reconstruction. Notably, our framework elimi-

Table 2. Comparison results between our method and U-net framework on the multiple missing teeth in two patterns.

		Pattern 1			**Pattern 2**		
Tooth ID number		11	12	13	15	16	17
CD(mm)	U-net	2.93	3.61	2.77	2.84	2.17	2.57
	Ours	**1.24**	**1.16**	**1.46**	**1.22**	**1.46**	**1.81**
HD(mm)	U-net	3.51	4.12	4.01	5.04	5.58	4.55
	Ours	**2.08**	**2.29**	**3.52**	**2.85**	**3.39**	**3.59**
AE(degree)	U-net	7.22	5.38	7.71	10.23	9.70	10.29
	Ours	**3.69**	**4.39**	**6.84**	**5.46**	**6.97**	**7.40**

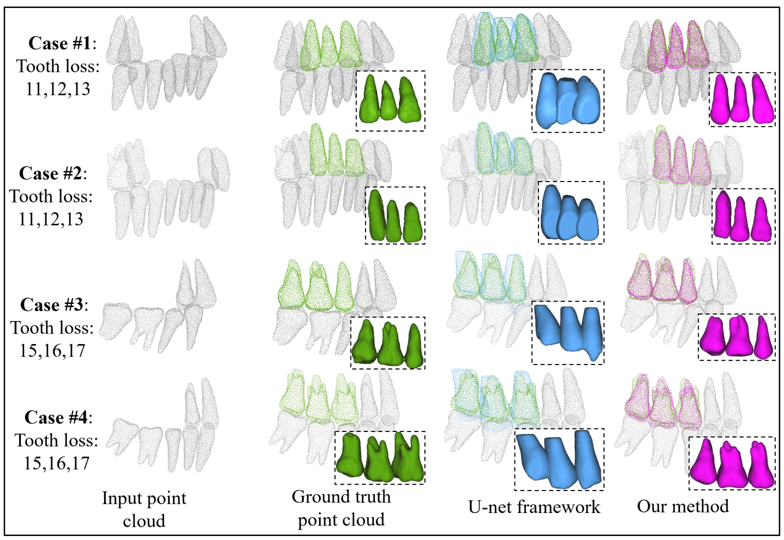

Fig. 3. Comparison results between our method and U-net framework in multiple missing teeth prediction.

nates the dependency on auxiliary positional cues required by the U-Net framework, which necessitates tooth centroid prompts for marginal tooth generation in single missing tooth prediction.

Table 2 summarizes the quantitative results of the experiment of multiple missing teeth prediction across two clinical scenarios: consecutive anterior tooth loss (teeth 11-13) and posterior marginal tooth loss (teeth 15-17). The proposed method demonstrates superior measurement accuracy compared to the U-net framework, achieving mean reductions of 50.0% in Chamfer distance (CD: 1.30 vs. 2.60 mm), 32.5% in Hausdorff distance (HD: 2.57 vs. 3.81 mm), and 27.6% in angular error (AE: 5.42° vs. 7.49°). These improvements validate the framework's enhanced capability in handling complex edentulous patterns. Fig. 3 provides visual evidence supporting these quantitative findings. Our method generates anatomically plausible tooth arrangements that better approximate the ground truth morphology, particularly in maintaining proper occlusal relationships and interproximal contacts. The results of the ablation experiment demonstrate that without the centroid-prompt, the generation of teeth in the multiple missing teeth prediction becomes disordered, leading to a significant decrease in accuracy.

4 Conclusion

This study presents a unified deep learning framework to address the critical challenges of predicting the original 3D shapes and positions of missing teeth for dental implantation planning. By introducing a dynamic iterative generation strategy and a tooth-centroid-prompted conditional diffusion model, our approach achieves precise 3D reconstruction of missing teeth across diverse loss patterns without case-specific retraining. The proposed framework demonstrates clinically relevant accuracy, outperforming conventional U-net based framework across key metrics. Our results suggest transformative potential for improving surgical planning efficiency, minimizing intraoperative risks, and enabling data-driven personalized implantation.

Acknowledgments This work was supported by National Natural Science Foundation of China (grant numbers 62473287, 62303320).

Disclosure of Interests The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. Fdi tooth numbering chart. American Journal of Orthodontics and Dentofacial Orthopedics **121**(1), 56 (2002)
2. Bayrakdar, S.K., Orhan, K., Bayrakdar, I.S., Bilgir, E., Ezhov, M., Gusarev, M., Shumilov, E.: A deep learning approach for dental implant planning in cone-beam computed tomography images. BMC Medical Imaging **21**(1), 86 (2021)
3. D'haese, J., Ackhurst, J., Wismeijer, D., De Bruyn, H., Tahmaseb, A.: Current state of the art of computer-guided implant surgery. Periodontology 2000 **73**(1), 121–133 (2017)
4. Flügge, T., Kramer, J., Nelson, K., Nahles, S., Kernen, F.: Digital implantology—a review of virtual planning software for guided implant surgery. part ii: Prosthetic set-up and virtual implant planning. BMC oral health **22**(1), 1–11 (2022)
5. Guentsch, A., Sukhtankar, L., An, H., Luepke, P.G.: Precision and trueness of implant placement with and without static surgical guides: an in vitro study. The Journal of prosthetic dentistry **126**(3), 398–404 (2021)
6. Li, J., Egger, J.: Towards the automatization of cranial implant design in cranioplasty. In: Computer Assisted Radiology and Surgery, vol. 12439, pp. 1–8. Springer International Publishing (2020)

7. Li, J., Egger, J.: Towards the automatization of cranial implant design in cranioplasty ii. In: Computer Assisted Radiology and Surgery, vol. 13123, pp. 1–8. Springer International Publishing (2021)
8. Li, J., et al.: Autoimplant 2020-first miccai challenge on automatic cranial implant design. IEEE Transactions on Medical Imaging **40**(9), 2329–2342 (2021)
9. Londoño, J.J., Ramos, A.M., Correa, S.A., Mesnard, M.: Review of expandable dental implants. British Journal of Oral and Maxillofacial Surgery **59**(5), 546–554 (2021)
10. Ma, L., Lian, C., Kim, D., Xiao, D., Wei, D., Liu, Q., Kuang, T., Ghanbari, M., Li, G., Gateno, J., et al.: Bidirectional prediction of facial and bony shapes for orthognathic surgical planning. Medical image analysis **83**, 102644 (2023)
11. Ma, L., Xiao, D., Kim, D., Lian, C., Kuang, T., Liu, Q., Deng, H., Yang, E., Liebschner, M.A., Gateno, J., et al.: Simulation of postoperative facial appearances via geometric deep learning for efficient orthognathic surgical planning. IEEE transactions on medical imaging **42**(2), 336–345 (2022)
12. Ma, L., Zhang, J., Deng, K., Xue, P., Cui, Z., Fang, Y., Tang, M., Zhao, Y., Zhu, M., Ding, Z., et al.: Construction of unbiased dental template and parametric dental model for precision digital dentistry. arXiv preprint arXiv:2304.03556 (2023)
13. Needleman, I., Garcia, R., Gkranias, N., Kirkwood, K.L., Kocher, T., Iorio, A.D., Moreno, F., Petrie, A.: Mean annual attachment, bone level, and tooth loss: A systematic review. Journal of Clinical Periodontology **45**, S112–S129 (2018)
14. Park, J., Lee, J., Moon, S., Lee, K.: Deep learning based detection of missing tooth regions for dental implant planning in panoramic radiographic images. Applied Sciences **12**(3), 1595 (2022)
15. Schubert, O., Schweiger, J., Stimmelmayr, M., Nold, E., Güth, J.F.: Digital implant planning and guided implant surgery—workflow and reliability. British dental journal **226**(2), 101–108 (2019)
16. Talmazov, G., Bencharit, S., Waldrop, T.C., Ammoun, R.: Accuracy of implant placement position using nondental open-source software: An in vitro study. Journal of Prosthodontics **29**(7), 604–610 (2020)
17. Whitley III, D., Eidson, R.S., Rudek, I., Bencharit, S.: In-office fabrication of dental implant surgical guides using desktop stereolithographic printing and implant treatment planning software: A clinical report. The Journal of prosthetic dentistry **118**(3), 256–263 (2017)
18. Zhou, L., Du, Y., Wu, J.: 3d shape generation and completion through point-voxel diffusion pp. 5826–5835 (2021)