MT3D: 3D Missing Tooth Detection and Dataset

Junsu Kim, 20235090

UNIST Artificial Intelligence Graduate School
joonsu0109@unist.ac.kr

1 Introduction

In recent years, robotic systems have been increasingly integrated into surgical procedures, including dental implant surgery [1], promising enhanced precision, efficiency, and reproducibility. However, the advancement of fully autonomous robotic dental surgery is significantly hampered by data scarcity. The reliance on manual data annotation and the general lack of extensive missing tooth datasets restrict the capabilities of current systems. This motivates the development of an automated pipeline for generating labeled data for missing teeth from complete intraoral scans.

Previous deep learning approaches for detecting missing teeth have primarily focused on 2D X-ray images [3, 7]. While useful for initial diagnosis, these 2D methods are inadequate for robotic dental implant surgery, which demands precise 3D localization and orientation of the implant site within a real-world coordinate system. The information from a 2D image lacks the necessary depth and rotational data, limiting its practical application in surgical robotics. Therefore, there is a clear need for a 3D missing tooth detection model that can directly predict the precise position and orientation of point cloud missing tooth data.

To address this gap, we introduce `MT3D`, a novel dataset for 3D missing tooth detection We generated this dataset by removing individual teeth from 175 fully-shaped 3D intraoral scans selected from the public Teeth3DS dataset [2]. For each of the 175 scans, we created 12 distinct tooth loss cases, resulting in a comprehensive dataset. Furthermore, we propose a novel 3D detection framework that predicts 3D bounding box of the missing tooth area from point cloud. This bounding box accurately defines the position and orientation for subsequent implant placement. Our experimental results demonstrate high detection accuracy, presenting a clinically viable solution that paves the way for autonomous dental implant surgery.

2 Related Works

While deep learning has broadly advanced dental diagnostics in areas like caries detection and segmentation [4, 5], methods for guiding implant surgery face distinct challenges. Current approaches for detecting missing tooth often rely on either 2D radiographs or high-quality 3D pre-operative scans. For instance, Park et al. [7] used 2D panoramic images for preliminary planning, but such methods inherently lack the depth information crucial for surgical navigation. More

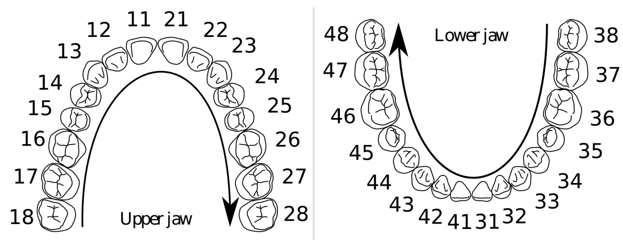

Fig. 1. FDI World Dental Federation notation.

advanced frameworks, like that of Ma et al. [6], use 3D scan data for precise pre-operative planning. However, their reliance on complete, clean scans makes them ill-suited for the data captured during a live surgery. A critical gap therefore exists for real-time, intra-operative guidance. Our work addresses this by proposing a model that operates directly on 3D point cloud data, which can be readily acquired during surgery, making it suitable for live surgical applications.

3 Method

3.1 MT3D: 3D missing tooth dataset

Teeth3DS dataset. To generate 3D missing tooth dataset, we utilized the open-source Teeth3DS [2] dataset. Teeth3DS is composed of 1,800 3D intraoral scans from 900 real patients, complete with instance labels for the upper and lower jaws. However, since the Teeth3DS dataset is based on real clinical cases, it contains many scans with imperfect or damaged teeth, which could be confounded with actual missing tooth. Therefore, to establish a reliable base data, we filtered the dataset. Following the annotation criteria shown in Fig. 1, we selected the scans that contained 14 or more teeth. This filtering process yielded a curated subset of 175 scans. By using this refined data, we were able to construct a more complete missing tooth dataset for training deep learning model.
3D missing tooth dataset. Acquiring real-world data for missing teeth is exceptionally challenging, leading to a significant data deficit in the clinical domain. To address this, we propose a pipeline that generates a synthetic missing tooth dataset as point cloud with 3D bounding box labels, as illustrated in Fig.2. This process leverages the intraoral scans and instance labels from the existing Teeth3DS dataset. Our method creates a diverse data by generating multiple samples from each scan. For each scan, we remove a single tooth from one of 12 target positions ($x1 - x6$), as shown in Fig.3
Processing pipeline. The pipeline consists of two main stages: generating the healed-ridge surface and creating the 3D bounding box annotation. First, to generate the point cloud for the edentulous surface, we identify the adjacent points—the points on the gingiva and proximal surfaces of neighboring teeth, using the instance label of the tooth designated for removal. Concurrently, we

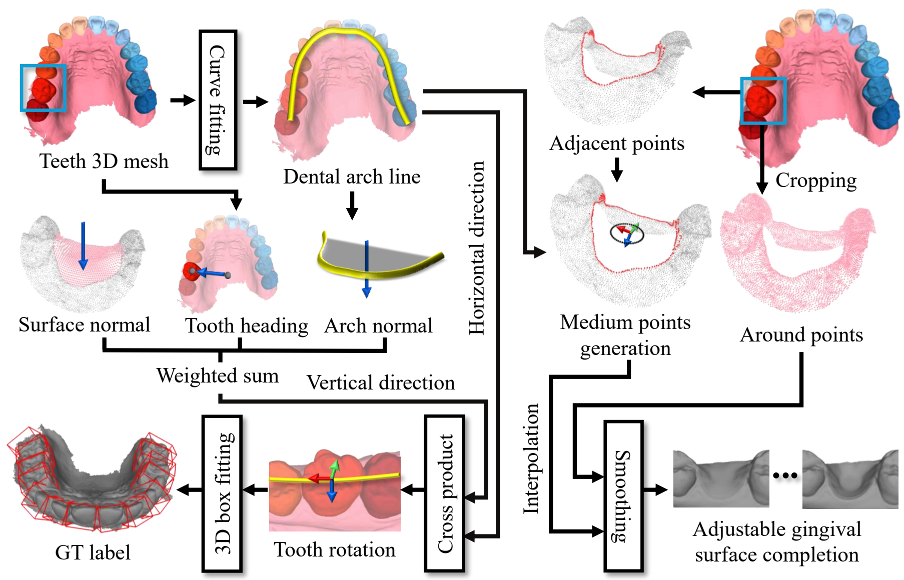

Fig. 2. Missing tooth data generation pipeline.

determine the dental arch line via spline curve fitting using the center point of each tooth. Principal Component Analysis (PCA) is then applied to this curve to compute the arch normal vector. Along this vector's direction, we generate a set of medium points that serve as a scaffold for the new surface. An initial surface is then created by applying grid interpolation between the adjacent points and medium points. Finally, this surface is smoothed using a moving average filter with the surrounding point cloud, resulting in a natural-looking edentulous ridge. Second, we construct the 3D bounding box annotation for the removed tooth. The box's vertical direction vector is determined by a weighted sum of three vectors: the local surface normal, the arch normal, and a tooth heading vector (defined from the center of the teeth scan to the center of the missing tooth). The horizontal direction is derived from the two points on the dental arch line closest to the missing tooth's center. With these two vectors defining the box's orientation, we perform a final box fitting on the original point cloud of the removed tooth to generate its precise 3D bounding box label.

3.2 3D missing tooth detection

Architecture. As illustrated in Fig.4, we built the 3D missing tooth detection model that takes a point cloud as input and predicts the 3D bounding box of a missing tooth, trained on our generated `MT3D` dataset. The model consists of an encoder to extract point cloud features and a prediction head to regress the bounding box parameters. The encoder is based on Point Transformer [9]

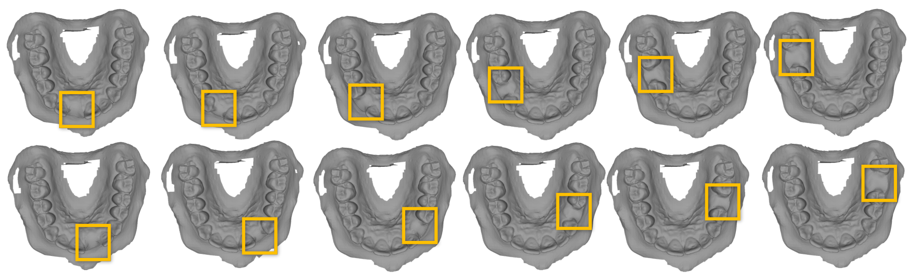

Fig. 3. Examples of the `MT3D` missing tooth dataset, with 12 teeth removed per scan. The figure is shown as a mesh for visualization purposes only.

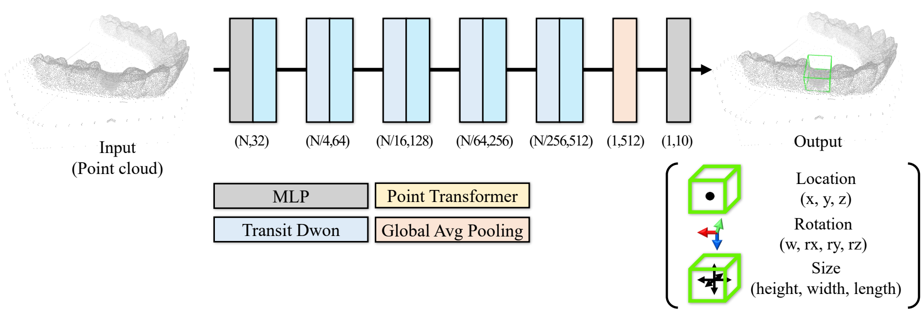

Fig. 4. Missing tooth detection framework.

architecture. Unlike the original Transformer [8], which inefficiently computes attention between all point cloud, Point Transformer introduces a more effective approach for point cloud data. It first uses k-Nearest Neighbors (kNN) to define a local neighborhood around each query point. The self-attention mechanism is then applied only within these local points. This strategy allows the model to efficiently capture detailed local geometric structures, leading to a more powerful feature representation. The encoded features are then processed by a prediction head composed of a multi-layer perceptron (MLP). This head is responsible for predicting the 10 parameters that define the final 3D bounding box for the missing tooth.

Training objective. To train our model, we utilize a multi-task training objective composed of four distinct loss functions. For the regression of the 3D bounding box's center point and its size (dimensions), we employ an L1 Loss. The orientation of the bounding box is represented as a quaternion, and we use a Mean Squared Error (MSE) loss to regress its components. Finally, to ensure more stable convergence of the orientation, we introduce an auxiliary loss that maximizes the cosine similarity between the predicted and ground-truth rotational axis vectors. The total loss is a sum of these four components: the L1 loss for the center, the L1 loss for the size , the MSE loss for the quaternion , and the cosine similarity loss for the axes.

4 Experiments

4.1 Implementation details

We utilize a total of 4,200 samples for training, generated from our pipeline (175 scans $\times$ 12 missing teeth $\times$ 2 jaws). For evaluation, we use a validation set of 1,152 samples and a test set of 888 samples. All models were trained for 100 epochs on a single NVIDIA RTX 3090 GPU with a batch size of 8. We employed an SGD optimizer and a cosine learning rate scheduler to manage the training process. We evaluate our model's performance using three distinct metrics. The overall accuracy of the bounding box shape and overlap is assessed using 3D IoU. The positional accuracy is measured by the center distance between the predicted and ground-truth box centers. Finally, the rotational accuracy of the bounding box is evaluated using cosine similarity on its orientation vectors.

4.2 Quantitative results

We conducted an ablation study to validate the contribution of each component in our training objective, with the results summarized in Tab.1. First, when we removed both the size loss and the axis loss, the model failed to regress meaningful bounding boxes, making an IoU measurement impossible. Next, when training without only the axis loss, the model achieved decent performance but showed poor convergence for the center distance, indicating positional instability. Our full model, incorporating all four loss functions, demonstrated the best performance across all evaluation metrics. This highlights the importance of the auxiliary axis loss. By stabilizing the rotational regression, it indirectly aids the entire bounding box estimation process, proving crucial for achieving robust and accurate results.

Loss	**IoU(%) ↑**	**Center Distance (cm) ↓**	**Cosine Similarity ↑**		
			X	**Y**	**Z**
w/o size, axis loss	–	1.2900	0.9820	0.9797	0.9884
w/o axis loss	62.91	3.0140	0.9739	0.9703	0.9607
Ours	**68.85**	**0.9196**	**0.9923**	**0.9897**	**0.9934**

Table 1. Ablation study of different loss configurations.

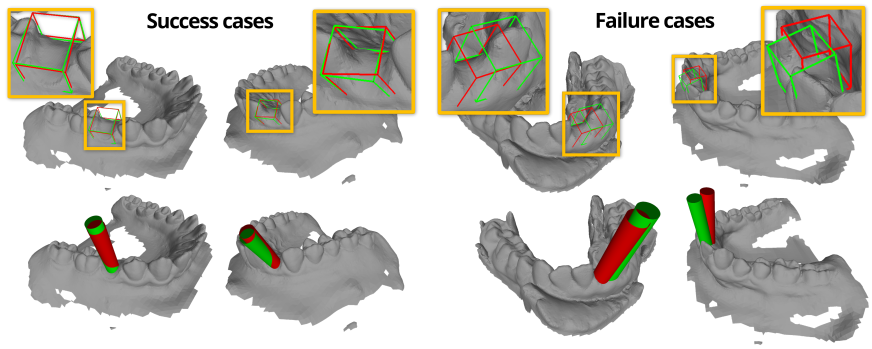

Fig. 5. Qualitative results of 3D missing tooth detection.

4.3 Qualitative results

We compare successful and failed predictions qualitatively in Fig.5. Success cases show that our model achieves high prediction accuracy, particularly in the molar regions. This is likely because the edentulous ridge surface in these areas is relatively flat and wide, providing clearer geometric features for the model to analyze. Conversely, failure cases primarily occur in the anterior (incisor) region. Unlike the broad molar areas, the ridge surface for incisors is much narrower and often more sharply curved. This complex topology appears to create challenges for our model, leading to difficulties in accurately predicting the 3D bounding box.

5 Conclusion

In this paper, we address the critical challenge of data scarcity for automating dental implant surgery. We introduced a novel pipeline to generate `MT3D`, a synthetic 3D point cloud dataset for missing tooth detection. On this dataset, we trained a Point Transformer-based 3D missing tooth detection model and experimentally validated the effectiveness of our entire framework. While challenges in real-time application and further model optimization remain, this work provides a foundational dataset and methodology. We believe our contribution represents a significant step toward the automation of dental implant surgery.

References

1. Bahrami, R., Pourhajibagher, M., Nikparto, N., Bahador, A.: Robot-assisted dental implant surgery procedure: A literature review. Journal of Dental Sciences (2024)
2. Ben-Hamadou, A., Smaoui, O., Rekik, A., Pujades, S., Boyer, E., Lim, H., Kim, M., Lee, M., Chung, M., Shin, Y.G., Leclercq, M., Cevidanes, L., Prieto, J.C., Zhuang,

S., Wei, G., Cui, Z., Zhou, Y., Dascalu, T., Ibragimov, B., Yong, T.H., Ahn, H.G., Kim, W., Han, J.H., Choi, B., van Nistelrooij, N., Kempers, S., Vinayahalingam, S., Strippoli, J., Thollot, A., Setbon, H., Trosset, C., Ladroit, E.: 3dteethseg'22: 3d teeth scan segmentation and labeling challenge (2023), https://arxiv.org/abs/2305.18277
3. Chen, S.L., Chen, T.Y., Huang, Y.C., Chen, C.A., Chou, H.S., Huang, Y.Y., Lin, W.C., Li, T.C., Yuan, J.J., Abu, P.A.R., et al.: Missing teeth and restoration detection using dental panoramic radiography based on transfer learning with cnns. IEEE Access **10**, 118654–118664 (2022)
4. Kim, M.S., Amm, E., Parsi, G., ElShebiny, T., Motro, M.: Automated dentition segmentation: 3d unet-based approach with miscnn framework. Journal of the World Federation of Orthodontists **14**(2), 84–90 (2025)
5. Lee, J.H., Kim, D.H., Jeong, S.N., Choi, S.H.: Detection and diagnosis of dental caries using a deep learning-based convolutional neural network algorithm. Journal of dentistry **77**, 106–111 (2018)
6. Ma, L., Xue, P., Gu, Y., Zhao, Y., Zhu, M., Ding, Z., Shen, D.: Accurate 3d prediction of missing teeth in diverse patterns for precise dental implant planning. arXiv preprint arXiv:2307.07953 (2023)
7. 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)
8. 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)
9. Zhao, H., Jiang, L., Jia, J., Torr, P.H., Koltun, V.: Point transformer. In: Proceedings of the IEEE/CVF international conference on computer vision. pp. 16259–16268 (2021)