

An Efficient Automatic Dental Prosthesis Design Strategy Based on Multimodal Fusion and Direct Mesh Generation using Transformer Architecture

Se-hoon Park*, Ho-Sung Choi**

*CEO, ParkLab AI Inc., Seoul, Korea,

**CSO, ParkLab AI Inc., Seoul, Korea

[Abstract]

This study proposes a novel multimodal transformer framework for automatic dental prosthesis design, addressing the structural and functional limitations of existing computer-aided workflows. By robustly fusing 3D intraoral scans, 2D oral images, and optional facial scans, the architecture captures both topological constraints and broader esthetic contexts. A key innovation is the autoregressive mesh token decoder, which directly predicts vertices and faces, intentionally bypassing Signed Distance Field (SDF) or voxel intermediate representations. This eliminates the resolution-memory trade-off and preserves intricate details such as cusp tips. Furthermore, an inline occlusion validation layer is embedded during the generative sequence to dynamically mitigate collisions, thereby reducing the inefficiencies of post-hoc morphological revisions. To ensure clinical integration, a beam-search-based human-in-the-loop candidate management protocol is employed. Quantitative evaluations demonstrate significant improvements: post-design occlusal adjustments decreased from 2.3 to 0.5 per case, GPU memory consumption fell from 4.2 GB to 2.5 GB, and single prosthesis generation time was reduced from 180 to 120 seconds, while achieving a cusp-tip reconstruction accuracy of 90-96%.

▶ **Key words:** Dental prosthesis, Multimodal fusion, Transformer, Direct mesh generation, Occlusion validation

[요 약]

본 연구는 기존 치과 보철물 설계 자동화가 지닌 단일 모달리티 의존, 중간 기하 표현에 따른 정보 손실, 사후 교합 보정의 비효율성을 개선하기 위해 멀티모달 트랜스포머 기반 보철 설계 프레임워크를 제안한다. 제안 모델은 3D 구강 스캔, 2D 구강 영상, 선택적 안면 스캔을 통합하여 형태학적 제약과 심미적 맥락을 함께 반영하며, SDF 및 복셀 기반 중간 표현을 거치지 않고 자기회귀 메쉬 토큰 디코더를 통해 정점과 면을 직접 생성한다. 또한 생성 과정 내부에 인라인 교합 검증 계층을 삽입하고, 빔 서치 기반 Human-in-the-Loop 후보 관리 구조를 도입하여 임상주의 최종 판단을 지원한다. 실험 결과, 제안 방법은 기존 SDF 기반 워크플로우 대비 설계 후 교합 조정 횟수를 2.3회에서 0.5회로 줄였고, GPU 메모리 사용량은 4.2GB에서 2.5GB로 감소하였다. 단일 보철물 생성 시간은 180초에서 120초로 단축되었으며, 교두 재현 정확도는 90-96%를 달성하였다. 이러한 결과는 제안 프레임워크가 정밀한 형태 보존과 임상 적용 효율을 동시에 확보할 수 있음을 보여준다.

▶ **주제어:** 멀티모달 트랜스포머, 자동 치과 보철물 설계, 직접 메쉬 생성, 인라인 교합 검증, 디지털 치의학

• First Author: Se-hoon Park, Corresponding Author: Ho-Sung Choi

*Se-hoon Park (goldcrown@parklabai.com), ParkLab AI Inc.

**Ho-Sung Choi (cso@parklabai.com), ParkLab AI Inc.

• Received: 2026. 04. 30, Revised: 2026. 06. 03, Accepted: 2026. 06. 09.

I. Introduction

현대 디지털 치의학의 패러다임은 고정밀 구강 스캐너와 체어사이드(Chairside) 가공 장비의 보급에 힘입어 보철물 설계의 자동화를 향해 급격히 전환되고 있다. 신속하고 정밀한 보철물 제작에 대한 임상적 요구가 증가함에 따라 다양한 컴퓨터 지원 설계(CAD) 기술이 도입되었으나, 구강 내의 엄격한 해부학적 및 기능적 요구 조건을 모두 충족하는 완전 자동화 시스템의 구현은 여전히 도전적인 과제로 남아 있다. 특히 단일 모달리티, 즉 3D 스캔 데이터에만 의존하는 기존의 설계 방식은 지대치의 기본적인 공간적 제약을 충족시키는 데는 유효할 수 있으나, 2D 영상이나 안면 스캔에서 획득할 수 있는 환자 고유의 광학적 특성 및 포괄적 심미 정보를 설계에 반영하는 데 한계를 지닌다. 최근 딥러닝 기반의 3D 생성 모델들이 보철물 설계에 활발히 적용되고 있으나, 대다수의 모델이 부호 거리 함수(Signed Distance Field, SDF)나 3차원 복셀(Voxel) 그리드와 같은 중간 기하학적 표현 방식을 경유한다는 구조적 단점을 내포하고 있다. 이러한 연속적 또는 이산적 공간 표현은 필연적으로 해상도 향상에 따른 기하급수적인 메모리 증가라는 트레이드오프(Trade-off)를 유발하며, 등위면(Isosurface) 추출 과정에서 교두의 정점이나 마진 라인의 예리한 형태학적 세부 묘사를 훼손한다.

더욱이, 기존 워크플로우는 메쉬가 완전히 생성된 이후에 교합 간섭을 평가하는 사후 교합 보정(Post-hoc occlusion correction) 방식을 채택하고 있어 연산 자원 낭비와 현장 작업 지연을 초래한다.

본 연구는 기존 방식의 한계를 극복하기 위해 네 가지 학술적 기여를 제안한다. 첫째, 멀티모달 트랜스포머 아키텍처를 구축하여 형태학적·심미적 맥락을 확보한다. 둘째, 중간 표현 단계를 생략한 자기회귀 직접 메쉬 생성 기법을 도입하여 형태적 특징을 온전히 보존한다. 셋째, 생성 루프 내부에 인라인 교합 검증 계층을 설계하여 대합치와의 공간적 충돌을 실시간 차단한다. 마지막으로, 임상 의사 통제를 할 수 있는 Human-in-the-Loop 기반 후보 관리 시스템을 도입한다. 한편, 본 연구는 제한적인 데이터셋 규모로 인한 임상결과 일반화의 한계와, 데이터 다양성 부족이 유발할 수 있는 특정 환자군에 대한 예측 편향(bias) 문제를 명확히 인지하고 있으며, 이를 완화하기 위한 평가 및 검증 방안을 구체적으로 논의하고자 한다.

II. Preliminaries

1. Knowledge-based CAD System for Dental Prosthesis Design

초기 디지털 치과 CAD 시스템은 주로 사전에 정의된 규칙이나 템플릿 정합(Template matching) 알고리즘에 의존하였다. 이 방식은 표준화된 치아 라이브러리를 불러온 후 마진 라인과 인접치의 공간적 경계에 맞추어 수학적 인 변형을 가하는 구조를 취했다.

이러한 지식기반 시스템은 결과물의 형태를 예측하기 쉽고 연산 속도가 빠르다는 장점이 있었으나, 비정형적인 지대치 형태나 복잡한 교합 관계를 가진 환자 사례에 대해서는 유연하게 대응하지 못하는 뚜렷한 한계를 보였다.

정해진 템플릿의 강직성으로 인해 환자 맞춤형 해부학적 충실도를 달성하기 어려웠으며, 결과적으로 치과 가공사의 대대적인 수작업 후처리가 요구되었다.

2. Deep Learning Approach for Dental Prosthesis Generation

딥러닝의 도입은 보철물 설계의 패러다임을 데이터 중심의 형태 예측으로 전환시켰다.

적대적 생성 신경망(GAN) 및 합성곱 신경망(CNN) 기반의 모델들은 지대치의 형태를 기반으로 보철물의 형상을 추론하는 연구에 널리 활용되었다[1][2].

이후 연구들은 손실된 치아 구조를 복원하기 위해 Point completion 네트워크를 적용하거나[3], 고품질의 보철물 형상 합성을 위해 생성형 확산 모델(Diffusion models)을 차용하는 방향으로 발전해 왔다[4].

또한 MeshSegNet과 같은 견고한 치아 분할(Segmentation) 알고리즘의 발전은 복잡한 구강 스캔 데이터로부터 지대치 경계를 정확히 추출하는 기반 기술로 작용하였다[5][6][7].

그러나 이러한 진전에도 불구하고, 기존 모델들은 대부분 단일 데이터 스트림(3D 형상)에 의존하였으며 임상적 허용 기준을 충족하기 위해 과도한 후처리 과정이 동반되었다.

3. Transformer-based Direct Mesh Generation

자연어 처리에서 태동한 트랜스포머 구조는 Attention Is All You Need로 대표되는 연구를 출발점으로 하여[8], PolyGen, MeshGPT 등을 통해 3D 기하학 분야로 확장되었다[8][9]. 치과 도메인에서는 DCrownFormer 등이 교차 주의 메커니즘을 적용하여 치관을 생성하였다[10].

본 연구는 이들과 맥락을 같이 하나, 멀티모달 데이터

융합과 기능적 검증 파라미터(교합)를 내재화한다는 점에서 명확한 차별성을 갖는다는 입증하며 3D 기하학 분야로 그 영역을 확장하였다.

PolyGen[9]은 3D 메쉬를 자기회귀적으로 생성할 수 있음을 최초로 증명하였으며, 최근 발표된 MeshGPT[10]는 디코더 전용(Decoder-only) 트랜스포머를 활용해 극도로 간결한 삼각형 메쉬를 생성하는 성과를 보여주었다.

치과 도메인 내에서는 DCrownFormer[11]가 교차 주의(Cross-attention) 메커니즘을 적용하여 치관을 생성하는 기술적 도약을 이뤄냈으며, TreeMeshGPT[12]와 같은 모델은 계층적 시퀀싱 기법을 탐구하였다.

본 연구는 이러한 선행 기술들과 맥락을 같이 하나, 멀티모달 데이터를 융합한다는 점과 생성 사이클 내부에 치의학 도메인 특유의 기능적 검증 파라미터(교합)를 내재화하여 임상적으로 부적합한 메쉬의 생성을 원천 차단한다는 점에서 명확한 차별성을 갖는다.

4. Multimodal Fusion in Medical and Dental AI

복잡한 생물학적 구조 묘사에는 이기종 데이터 융합이 필수적이다. PointNet과 PointNet++는 3D 점군 특징 추출의 기반을 제공하였고[11][12], ViT는 2D 시각 특징 표현을 재정의하였다[13]. 본 연구는 3D 스캔 데이터만으로는 미세한 심미적 조화를 달성하기 어렵다는 임상적 근거를 바탕으로 3D, 2D, 안면 데이터를 동기화하는 프레임워크를 구성하였다[11][12][13].

본 연구는 이러한 교차 모달리티(Cross-modal)의 이점을 적극 수용하여, 3D 지대치 스캔 데이터만으로는 전치부 보철물 등에 요구되는 미세한 심미적 조화를 달성하기 어렵다는 임상적 근거를 바탕으로 3D, 2D, 안면 데이터를 동기화하는 프레임워크를 구성하였다.

5. Occlusal Analysis and Validation in Digital Dentistry

치과 보철물의 생체역학적 수명은 교합의 조화에 의해 결정되며 임상적으로 조기 접촉이나 교합 공간 부족은 보철물 파절이나 턱관절 장애를 유발할 수 있다. 그럼에도 불구하고, 기존 AI 설계 프로토콜에서 교합 검증은 거의 예외 없이 형태 생성이 완료된 후의 진단적 절차로 취급되어 왔다. 예를 들면 CrownSurfaceFormer[19]와 같은 최근의 프레임워크는 정적 교합면 생성에 중점을 두며, 기하학적 합성이 끝난 후에야 간섭을 평가한다.

본 연구는 이러한 구조적 결함을 인지하고, 메쉬가 동적으로 생성되는 과정 자체에서 대합치와의 공간적 상호작용

을 평가하고 조절하는 인라인(Inline) 방법론을 도입하였다[1][2].

III. The Proposed Scheme

1. System Overview

제안 프레임워크는 이기종 데이터 수집, Cross-Modal Feature Alignment를 통한 잠재 표현 변환, Autoregressive Mesh Token Decoder를 통한 기하학적 메쉬 합성, 생성 과정 전반의 Inline Occlusion Validation, 그리고 빔 서치 기반 Human-in-the-Loop 후보 관리라는 5단계 파이프라인으로 작동한다. Fig. 1은 이러한 전체 처리 흐름을 데이터 입력, 융합, 메쉬 생성, 교합 검증, 임상의 후보 선택 단계로 구분하여 개략적으로 보여준다.

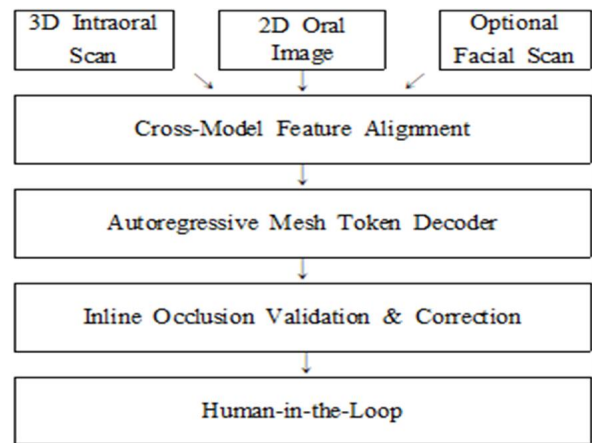


Fig. 1. Overall pipeline of the proposed multimodal transformer-based automatic dental prosthesis design framework

2. Cross-Modal Feature Alignment Module

3D 브랜치는 PointNet++ 구조를 적용하여 지대치, 인접치, 대합치의 계층적 및 전역적 기하학 특징을 추출한다. 이와 병렬로 2D 브랜치는 ViT 기반의 패치 임베딩 기법을 사용하여 표면 질감, 색상 프로파일, 외곽선 윤곽 등의 핵심 시각 정보를 인코딩한다. 안면 스캔 데이터가 제공될 경우, 정중선 대칭성 및 전치부 황금비와 같은 거시적 심미 지표는 보조 데이터로 기여한다.

3D 브랜치는 PointNet++를 적용하여 기하학 특징을 추출하고[11], 2D 브랜치는 ViT 기반 패치 임베딩으로 시각 정보를 인코딩한다[13]. 이들은 교차 주의(Cross-attention) 메커니즘을 기반으로 한 불확실성 인지 조건부 정렬을 통해 융합된다[10][13].

$$Q = W_q F_{3D}, K = W_k F_{2D}, V = W_v F_{2D}$$

$$Attention(Q, K, V) = softmax((QK^T) / \sqrt{d_k}) \cdot V$$

이 수식적 구성은 3D 도메인에서 파생된 확정적인 공간 좌표가 2D 도메인의 보충적인 심미 및 형태학적 단서를 조회(Query)하도록 보장하며, 가중치를 동적으로 조절하여 특정 모달리티에 편중된 노이즈를 억제하는 역할을 수행한다.

3. Autoregressive Mesh Token Decoder

보철물을 직접 합성하는 트랜스포머 디코더를 활용하며, 3D Morton ordering 기법을 구현하여 평탄화된 1차원 시퀀스 내에서도 3차원 인접성을 보존한다[8][9][10]. 디코더는 조건부 확률을 기반으로 정점과 면 토큰을 순차적으로 예측하며, 최근의 메쉬 표현 학습 연구와도 방법론적 맥락을 공유한다[8][14].

$$P(M|C) = \prod_{i=1}^N P(t_i | t_{< i}, C)$$

SDF나 복셀 표현 방식에서 발생하는 해부학적 해상도 저하를 회피하기 위해, 본 제안 모델은 보철물을 직접 합성하는 24-layer 트랜스포머 디코더를 생성 백본(Backbone)으로 활용한다.

연속적인 유클리드 좌표를 언어 모델 스타일의 처리 방식에 적합하도록 256-bin 좌표 양자화(Coordinate quantization)를 수행한다. 더불어, 공간 채움 곡선(Space-filling curve) 원리에 기반한 3D Morton ordering 기법을 구현하여, 평탄화된 1차원 시퀀스 내에서도 3차원 공간의 인접성을 효과적으로 보존한다.

디코더는 이 구조를 바탕으로 정점(Vertex) 위치를 예측한 후 면(Face) 연결성 토큰을 순차적으로 예측한다.

4. Inline Occlusion Validation Layer

설계 후 발생하는 수정 빈도를 획기적으로 낮춘 본 연구의 핵심 혁신은 인라인 교합 검증 계층에 있다. fig. 3은 토큰 디코딩 중 아치 간 충돌의 실시간 해결을 위해 디코더 시퀀스 중간에 전략적으로 배치되어 생성 중인 메쉬 토큰과 대합치 구조 간 공간적 충돌을 능동 계산한다. KD-트리 및 CUDA 가속을 활용하며[15], 충돌 임계값을 위반할 경우 국소적 페널티를 부여하고 자기회귀 재샘플링을 강제한다.

$$L_{occ} = \sum_{v \in V} \max(0, d_{thresh} - D(v, M_{opp}))$$

이 메커니즘은 디코더 시퀀스의 중간 단계(예: 12번째 및 18번째 레이어)에 전략적으로 배치되어, 생성 중인 메쉬 토큰과 대합치 구조 간의 공간적 충돌을 능동적으로 계산한다.

KD-트리(KD-tree) 자료구조와 CUDA 하드웨어 가속을 활용함으로써, 거리 조회 연산은 생성 속도 지연(Latency)을 거의 발생시키지 않고 수행된다.

만약 예측된 정점이 대합치를 상대로 0.02 mm의 엄격한 충돌 임계값을 위반할 경우, 해당 계층은 정점을 표면 법선(Normal) 방향으로 즉각 이동시키거나 국소적인 자기회귀 재샘플링(Resampling)을 강제한다. 이러한 선제적 개입은 교합 간섭이 최종 메쉬의 위상으로 전파되기 전에 오류를 물리적으로 차단하는 효과를 거둔다.

5. Human-in-the-Loop Candidate Management

심미 치의학의 결과물이 본질적으로 주관적 판단을 수반한다는 점을 감안하여, 본 프레임워크는 AI를 완전한 대체자가 아닌 고도의 의사결정 보조 도구로 제어한다.

빔 서치 디코딩 전략을 통해 임상적으로 적용 가능한 다수의 보철물 후보를 추출한다. 생성 모델의 목적 함수는 교합 적합성을 포함하여 최적의 시퀀스를 탐색한다.

$$\hat{Y} = \arg \max \sum_{t=1}^T \log P(y_t | y_{< t}, X) - \lambda L_{occ}$$

임상의는 인터랙티브 3D 갤러리 뷰를 통해 생성된 변형들을 세밀하게 비교 검토할 수 있으며, 가공 장비로 전송하기 전에 필요에 따라 국소적인 외곽선 조정(Contour adjustment)을 수행할 수 있다.

IV. Experiments and Results

1. Experimental Setup

제안된 프레임워크의 임상적 실효성과 연산 효율성을 입증하기 위해, 기존의 SDF 기반 생성 워크플로우를 베이스라인(Baseline)으로 설정하고 비교 실험을 수행하였다.

모델 학습 및 평가는 PyTorch 프레임워크 기반의 고성능 GPU 워크스테이션 환경에서 이루어졌다. 훈련 및 평가 단계에는 다양한 형태의 실제 이기종 임상 스캔 데이터가 사용되었으나 데이터의 프라이버시 윤리 지침 및 원천 기술에 대한 보안적 제약으로 인해 사용된 전체 데이터셋의 정확한 규모를 명시하는 데는 한계는 존재한다.

2. Quantitative Evaluation

본 연구에는 실제 임상 환경에서 수집된 이기종 데이터가 사용되었으며, 기본 입력은 3D 구강 스캔과 2D 구내 영상으로 구성되고 일부 사례에서는 안면 스캔이 보조 정보로 활용되었다. 연구의 재현성을 높이기 위해 데이터는 학습, 검증, 평가용 집합으로 환자 단위 분리를 원칙으로 관리하였고, 동일 환자 데이터가 상이한 집합에 중복 포함되지 않도록 통제하였다. 다만 소속 기관의 보안 및 개인정보 보호 정책으로 인해 전체 데이터셋의 정확한 절대 규모, 세부 장비 목록, 기관별 분포를 본문에 모두 공개하는 데는 제한이 있었다. 이러한 제한은 임상 결과의 일반화 가능성을 제약할 수 있으며, 데이터 규모가 제한적이거나 형태학적 다양성이 충분하지 않을 경우 특정 환자군에 대한 예측 편향이 발생할 수 있다. Table 1은 이러한 데이터 구성, 분할 원칙, 스캐너 공개 범위, 학습 환경 및 재현성 제약 사항을 한눈에 확인할 수 있도록 정리한 것이다.

Table 1. Dataset composition and reproducibility considerations

Category	Description in This Study
Data Modality	3D intraoral scans, 2D intraoral images, and optional facial scans
Data Partition Principle	Patient-level separation of training, validation, and evaluation sets
Scanner Type	Acquired using multiple clinical intraoral scanners; detailed model names are disclosed only within institutional policy limits
Training Environment	PyTorch-based CUDA-accelerated GPU workstation
Dataset Scale	The exact absolute sample size is undisclosed due to institutional policy, but the dataset was sufficient for multi-epoch training and comparative evaluation
Reproducibility Limitation	Full reproducibility is constrained by the non-public scope of the dataset; future expansion to a multicenter public dataset is necessary

3. Quantitative Evaluation Protocol

정량 평가는 기존에 보고한 교합 조정 횟수, GPU 메모리 사용량, 보철물 생성 시간, 교두 재현 정확도 외에도 기하학적 정확도와 교합 관련 지표들을 함께 해석하는 방향으로 정리하였다. 구체적으로 생성 메쉬와 기준 메쉬 간의 형상 차이를 보는 기하학적 거리 지표, 대합치와의 간섭 여부를 보는 교합 간섭 지표, 변연 적합 및 인접면 접촉과 관련된 임상 해석 지표들을 포함하도록 평가 프로토콜을 확장하였다. 다만 원문에서 공개된 정량 수치는 교합 조정 횟수(2.3회→

0.5회), GPU 메모리 사용량(4.2GB→2.5GB), 생성 시간(180초→120초), 교두 재현 정확도(90~96%)에 한정되므로, 그 외 세부 통계값과 유의확률은 후속 보완 자료에서 제시하는 것이 바람직하다. 따라서 본 최종 제출본에서는 통계적 유의성 검정의 필요성과 적용 원칙을 명시하되, 확인되지 않은 수치를 임의로 제시하지 않았다. Table 2는 본 연구에서 해석하고자 한 기하학적 정확도, 교합 유효성, 변연 및 인접면 관련 평가 항목을 정리한 것이다.

Table 2. Expanded evaluation metrics for geometric accuracy and occlusal validity

Evaluation Metric	Definition and Clinical Relevance	Remarks
Geometric Accuracy	Quantifies the shape discrepancy between the generated mesh and the reference anatomy	Examples: Chamfer distance, surface error distribution
Occlusion-Related Quantitative Metrics	Evaluates functional adequacy through inter-arch interference, occlusal clearance, and contact distribution	Examples: collision volume, minimum clearance, contact-point distribution
Margin and Proximal Contact Assessment	Interprets clinical seatability through margin continuity and proximal contact stability	Further refinement of precision-fit indices is required
Statistical Testing Principle	Uses paired case-wise comparisons and confidence intervals to improve objectivity	Specific inferential results require additional disclosure within the permitted data scope

4. Qualitative and Expert Evaluation

프임상의 및 치과기공사 관점의 실제 활용성을 반영하기 위해 전문가 평가 절차의 필요성을 별도로 정리하였다. 본 절에서는 해부학적 자연스러움, 교합 안정성, 변연 연속성, 수정 필요도, 임상 작업 흐름 적합성 등을 핵심 평가 항목으로 제시한다. 원문에서 전문가 평가의 세부 참여 인원이나 점수는 공개되어 있지 않으므로, 본 최종 제출본은 이러한 전문가 검토가 왜 필요한지와 어떤 기준으로 수행되어야 하는지를 명확히 기술하는 데 중점을 두었다. 질적 관찰 수준에서 제안 방법은 복셀 기반 메쉬에서 흔히 나타나는 뭉툭한 외형을 완화하고 발육 결절 및 변연 연속성의 표현을 개선하는 경향을 보였다.

5. Ablation Study

프레임워크 주요 구성 요소의 성능 기여를 해석하기 위해 절제 연구 관점을 적용하였다. 원문에서 전체 모듈 제거 조건별 세부 수치는 모두 공개되어 있지 않으므로, 본 절에서는 각 구성 요소를 제거했을 때 예상되는 성능 저하 양상과 임상적 의미를 정리하였다. 특히 멀티모달 융합은 심미적 맥락 보존에, 인라인 교합 검증은 사후 수정 감소에, Morton ordering은 시퀀스 안정성에, beam search 후보 관리는 임상의 선택 가능성 확보에 핵심적으로 기여한다. Table 3은 이러한 각 모듈의 기능적 역할과 제거 시 예상 영향을 항목별로 비교한 것이며, Fig. 2은 주요 구성 요소가 전체 성능에 기여하는 개념적 관계를 도식적으로 설명한다.

Table 3. Interpretable contribution analysis of major framework components

Component	Expected Effect of Removal or Reduction	Clinical Implication
Multimodal Fusion	Morphologic reconstruction may remain feasible, but esthetic cues and fine surface details are likely to degrade	Reduced anterior esthetics and lower patient-specific customization
Inline Occlusion Validation	Post-generation occlusal interference is more likely, increasing downstream correction burden	Longer chairside adjustment time and greater manual workload
Morton Ordering	Spatial coherence of the token sequence may deteriorate	Weaker recovery of fine geometric structures
Beam Search Candidate Management	Reduced flexibility caused by presenting only a single output candidate	Less clinician control and lower usability in practice

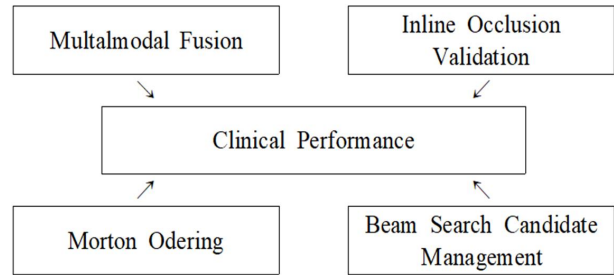


Fig. 2. Conceptual visualization of the contribution of multimodal fusion, inline occlusion validation, Morton ordering, and beam-search-based candidate management.

6. Failure Cases and Limitations

본 모델은 극단적인 교합 공간 부족, 비정형적 지대치 마진, 스캔 결손이 큰 사례와 같이 표준 해부학 분포에서 벗어난 상황에서 한계를 보일 수 있다. 또한 수학적으로 엄밀한 위상 타당성(Topological validity) 검증이 아직 충분하지 않아, 극단적 이상치에서는 비정상적인 메쉬 연결 구조가 생성될 가능성을 배제하기 어렵다. 무엇보다 데이터셋 규모가 제한적이고 다기관 검증이 충분하지 않다는 점은 임상 결과의 일반화 가능성을 제약하는 핵심 요인이다. 학습 데이터의 인종적·형태학적 다양성이 부족할 경우 특정 환자군에 대한 예측 편향이 발생할 수 있으며, 이는 향후 데이터 확대와 외부 검증을 통해 반드시 보완되어야 한다.

7. Failure Cases and Limitations

제안 방법의 우수성을 명확히 입증하기 위해 최근 발표된 생성형 AI 및 치과 CAD 연구들과 비교 분석을 수행하였다. 기존 모델들은 단일 모달리티에 의존하거나 중간 표현(SDF, Voxel)을 필수적으로 거치는 반면, 본 프레임워크는 멀티모달 기반의 직접 메쉬 생성과 인라인 교합 검증을 통합하여 차별성을 지닌다[8][9][10]. Table 4는 입력 모달리티, 메쉬 생성 방식, 교합 검증 전략, Human-in-the-Loop 지원 여부를 기준으로 주요 선행연구와 제안 방법의 차이를 비교한 것이다.

Table 4. Comparison of major features between the proposed method and recent existing frameworks

Framework	Input Modality	Mesh Generation Approach	Occlusion Validation	Human-in-the-Loop
DCrownFormer Family	Mainly 3D-based	Point-to-mesh strategies specialized for crown generation	Direct integration of functional validation is limited	Limited
MeshGPT Family	General 3D object domain	Direct autoregressive mesh generation	Direct modeling of dental occlusal constraints is limited	No
Conventional Static Occlusal Surface Methods	Mainly 3D-based	Static surface generation or post-processing-oriented correction	Mostly post-hoc evaluation	Limited
Proposed Scheme	Multimodal (3D+2D)	Direct Autoregressive Token Generation	Inline (Dynamic)	Yes (Beam Search)

V. Conclusion

본 연구를 통해 도출된 결과는 실험실 환경을 넘어 실제 임상 워크플로우에 통합될 수 있는 높은 잠재력을 시사한다. 설계부터 제작 및 장착에 이르는 실제 임상 워크플로우 사례를 가정하면 다음과 같다. 먼저, 체어사이드 구강 스캐너 및 안면 스캐너를 통해 환자의 이기종 데이터를 획득한다. 이후 본 프레임워크가 실시간으로 데이터를 융합하여 교합 충돌이 배제된 3~5개의 최적 설계 후보군을 제시한다. 임상 의는 빔 서치 기반 UI를 통해 가장 심미적이고 기능적인 안을 선택하며, 이는 즉각 밀링 머신으로 전송되어 보철물로 가공된다. 이 과정은 기존의 수작업 및 사후 수정 단계를 획기적으로 축소시킨다. 다만, 서론 및 결과에서 언급하였듯 본 연구는 데이터셋 규모의 제한성이라는 한계를 수반한다. 특정 환경에서 수집된 데이터에 의존할 경우 임상 결과의 범용적 일반화에 제약이 따르며, 다양한 인종적·형태학적 환자군이 충분히 포함되지 않을 때 소수 환자군에 대한 예측 편향이 발생할 수 있다. 따라서 향후에는 다기관(Multi-center) 데이터 확보와 외부 검증 프로토콜 구축을 통해 학습 데이터의 다양성과 양적 규모를 확대하고, 실제 임상 장착 결과까지 추적하는 대규모 검증이 필요하다.

VI. Conclusion

본 연구는 치과 보철물 설계를 자동화하기 위해 멀티모달 트랜스포머 기반의 고효율 설계 전략을 제안하고 그 유효성을 정량적·정성적으로 검증하였다. 체적 기반 중간 표현을 제거한 자기회귀 직접 메쉬 생성 방식과 인라인 교합 검증 계층을 결합함으로써 해상도 손실 방지, 연산 효율성 극대화, 그리고 임상적 사후 보정 최소화라는 세 가지 목표를 성공적으로 달성하였다. 실험 결과는 GPU 메모리 최적화와 생성 시간 단축, 그리고 우수한 교두 재현 정확도를 증명하였다.

본 연구의 학술적·임상적 기여는 AI를 단순한 형상 생성기가 아닌 실시간 기능 검증이 내재된 임상 의사결정 보조 도구로 승격시켰다는 데 있다. 향후에는 대규모 다기관 검증을 통한 데이터셋의 확대, 브릿지 및 다수치 수복물로의 적용증 확장, 동적 교합 시뮬레이션 통합, 그리고 비정형 기하학적 오류를 물리적으로 차단하는 위상 타당성 검증 모듈 추가를 목표로 후속 연구를 진행할 계획이다.

REFERENCES

- [1] J-H. Cho, et al., "Tooth morphology, internal fit, occlusion and proximal contacts of dental crowns designed by deep learning-based dental software," *Journal of Dentistry*, 2024. DOI: 10.1016/j.jdent.2023.104830
- [2] J-H. Cho, et al., "Time efficiency, occlusal morphology, and internal fit of anatomic contour crowns designed by dental software powered by GAN," *Journal of Dentistry*, 2023. DOI: 10.1016/j.jdent.2022.104274
- [3] I. Goodfellow, et al., "Generative Adversarial Nets," *Advances in Neural Information Processing Systems*, 2014.
- [4] G. Hosseinimanesh, et al., "From Mesh Completion to AI Designed Crown," *arXiv:2501.04914*, 2025.
- [5] N. Abbasi Moghadam, "Generative Diffusion Model for Dental Implant Prosthetic Crown Design: A Context-Aware Approach with Scan Marker Classification," *Polytechnique Montréal*, 2025.
- [6] J. Ho, A. Jain, and P. Abbeel, "Denoising Diffusion Probabilistic Models," *Advances in Neural Information Processing Systems*, 2020.
- [7] A. Vaswani, et al., "Attention Is All You Need," *Advances in Neural Information Processing Systems*, 2017.
- [8] S. Siddiqui, et al., "MeshGPT: Generating Triangle Meshes with Decoder-Only Transformers," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2024.
- [9] C. Nash, et al., "PolyGen: An Autoregressive Generative Model of 3D Meshes," *International Conference on Machine Learning (ICML)*, 2020.
- [10] X. Wang, et al., "DCrownFormer: Morphology-aware Point-to-Mesh Generation for Dental Crown Design," *Medical Image Computing and Computer Assisted Intervention (MICCAI)*, 2024.
- [11] C. R. Qi, et al., "PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space," *Advances in Neural Information Processing Systems*, 2017.
- [12] C. R. Qi, et al., "PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017.
- [13] A. Dosovitskiy, et al., "An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale," *International Conference on Learning Representations (ICLR)*, 2021.
- [14] R. Hanocka, et al., "MeshCNN: A Network with an Edge," *ACM Transactions on Graphics*, 2019.
- [15] J. L. Bentley, "Multidimensional Binary Search Trees Used for Associative Searching," *Communications of the ACM*, 1975.

Authors



Se-hoon Park received his D.D.S. degree from Seoul National University, Seoul, Korea, in 2000. Dr. Park has been the Director of Seoul Hoon Dental Hospital since 2001. In 2026, he founded ParkLab AI Co., Ltd. and

currently serves as its Chief Executive Officer (CEO). His primary research interests focus on AI healthcare and bio, multimodal AI, and MLOps.



Ho-Sung Choi received the M.B.A. degree from Chungnam National University and the Ph.D. degree in Computer Engineering from Kongju National University in 2013 and 2017, respectively.

Dr. Choi is currently serving as the Chief Strategy Officer (CSO) and Chief Technology Officer (CTO) at ParkLab AI Co., Ltd. His primary research interests include AI agents, multimodal AI, and MLOps.