Titre: Title:	3D Shape Generation: Geometrical and Functional Methods for Dental Crown Design
Auteur: Author:	Golriz Hosseinimanesh
Date:	2025
Type:	Mémoire ou thèse / Dissertation or Thesis
Référence: Citation:	Hosseinimanesh, G. (2025). 3D Shape Generation: Geometrical and Functional Methods for Dental Crown Design [Thèse de doctorat, Polytechnique Montréal]. PolyPublie. https://publications.polymtl.ca/70235/

URL de PolyPublie: PolyPublie URL:	https://publications.polymtl.ca/70235/
Directeurs de recherche: Advisors:	François Guibault, & Farida Cheriet
Programme: Program:	Génie informatique

POLYTECHNIQUE MONTRÉAL
affiliée à l'Université de Montréal

3D Shape Generation: Geometrical and Functional Methods for Dental Crown Design

GOLRIZ HOSSEINIMANESH
Département de génie informatique et génie logiciel

Thèse présentée en vue de l'obtention du diplôme de *Philosophiæ Doctor*
Génie informatique

Novembre 2025

POLYTECHNIQUE MONTRÉAL
affiliée à l'Université de Montréal

Cette thèse intitulée :

3D Shape Generation: Geometrical and Functional Methods for Dental Crown Design

présentée par **Golriz HOSSEINIMANESH**
en vue de l'obtention du diplôme de *Philosophiæ Doctor*
a été dûment acceptée par le jury d'examen constitué de :

Guillaume-Alexandre BILODEAU, président
Francois GUIBAULT, membre et directeur de recherche
Farida CHERIET, membre et codirectrice de recherche
Lama SÉOUD, membre
Matthew TOEWS, membre externe

DEDICATION

To all my family...

ACKNOWLEDGEMENTS

I would like to sincerely thank my supervisor, Prof. François Guibault, for the opportunity to pursue my PhD under his guidance, and for the freedom to explore research aligned with my interests. I am especially grateful for his patience, support, and kindness throughout this journey.

My heartfelt thanks also go to my co-supervisor, Prof. Farida Cheriet, for sharing her expertise in computer vision and her commitment to clinical innovation. I'm also grateful to Prof. Aaron Courville for his excellent courses in deep learning, which laid a solid foundation for this research.

I extend my appreciation to the research team at Polytechnique de Montréal for their support, thoughtful feedback, and collaborative spirit. Special thanks to Mrs. Julia Keren from Intellident Dentaire Lab for her help in gathering clinical data and related information.

I am also thankful to the NSERC and MEDTEQ Institutes for their financial support during this project.

Finally, I owe the deepest gratitude to my family. To my parents, Mohammad and Zahra, for their unconditional love and encouragement; to my brothers, Hosein and Ahmad, for always believing in me; and to my husband, Ehsan, whose unwavering support and love gave me strength through every challenge. This accomplishment would not have been possible without them.

RÉSUMÉ

La restauration de couronnes dentaires représente l'une des procédures les plus exigeantes techniquement en dentisterie restauratrice, nécessitant traditionnellement une conception manuelle approfondie par des techniciens qualifiés pour chaque couronne. Ce processus intensif en temps et dépendant de l'opérateur crée des goulots d'étranglement significatifs dans les flux de travail cliniques tout en introduisant de la variabilité dans la qualité et la précision. Alors que la demande clinique croît pour des restaurations plus rapides, plus cohérentes et personnalisées, le besoin de solutions automatisées intelligentes devient de plus en plus urgent. De plus, les outils de conception assistée par ordinateur existants s'appuient fortement sur des bibliothèques de modèles et des ajustements manuels, limitant leur capacité à gérer les contraintes géométriques complexes requises pour un ajustement approprié de la couronne, les relations occlusales et les exigences esthétiques.

Pour répondre à ces limitations, les chercheurs ont exploré des approches d'apprentissage automatique pour la génération automatisée de couronnes. Les premières méthodes ont formulé la conception de couronnes comme un problème de synthèse d'images 2D, mais ces approches étaient fondamentalement limitées par la perte d'information de la projection 3D vers 2D et ne pouvaient pas capturer la morphologie complète de la couronne. Les approches subséquentes ont employé des représentations 3D telles que les nuages de points et les grilles de voxels, mais manquaient de modélisation explicite des relations spatiales critiques incluant les limites de la ligne marginale, les motifs de contact occlusal et les interactions avec les dents adjacentes. Bien que ces méthodes aient démontré la faisabilité de la génération automatisée de couronnes, elles ont produit des surfaces géométriquement précises mais lisses manquant des détails morphologiques fins essentiels pour l'acceptation clinique, et ont échoué à prévenir la pénétration de la couronne dans les structures adjacentes ou à optimiser les relations de contact avec la dentition opposée.

L'objectif principal de cette thèse est de développer un cadre complet pour la génération entièrement automatisée de couronnes dentaires qui adresse le pipeline clinique complet de la reconstruction géométrique à la modélisation des contraintes spatiales. Nous formulons la conception de couronnes comme un problème de complétion de forme 3D spatialement contraint, où une architecture basée sur un transformateur génère des couronnes anatomiquement précises à partir de dents préparées et de leur contexte dentaire environnant. Pour relever ces défis, nous présentons quatre avancées méthodologiques complémentaires. Premièrement, nous avons établi une approche de complétion de nuages de points utilisant une architecture

encodeur-décodeur de transformateur avec intégration d'entrée de ligne marginale, démontrant des améliorations significatives dans la précision des limites mais nécessitant un post-traitement pour la conversion en maille. Deuxièmement, nous avons développé le premier pipeline de génération de mailles de bout en bout (DMC) utilisant la reconstruction de surface de Poisson différentiable, éliminant les exigences de post-traitement et atteignant une sortie de maille directe avec une distance de Chamfer moyenne de 0,062. Troisièmement, nous avons introduit des fonctions de perte avancées incluant une fonction de perte dédiée à la ligne marginale et la perte InfoCD pour un alignement de points amélioré, atteignant une réduction de 12,32% dans la distance de Chamfer et une réduction de 46,43% dans l'EQM par rapport à DMC. Quatrièmement, nous avons implémenté une modélisation complète des contraintes spatiales par l'intégration d'entrée de ligne marginale combinée avec la perte d'interaction antagoniste et les fonctions de perte d'intersection, assurant des relations occlusales appropriées et prévenant la pénétration de la couronne dans les structures adjacentes.

L'évaluation complète sur des ensembles de données cliniques réelles comprenant 557 cas à travers toutes les positions dentaires démontre des améliorations substantielles par rapport aux méthodes de pointe. Notre cadre atteint 35,9-40,6% d'améliorations dans la précision géométrique comparé aux approches existantes. L'intégration des contraintes spatiales réduit les erreurs maximales d'ajustement des limites de 1,37 à 0,74 mm et une réduction de 58,4% dans la variabilité, tandis que la perte d'interaction antagoniste contribue à une amélioration de 9,51% dans l'alignement occlusal et la perte d'intersection réduit substantiellement la pénétration de la couronne dans les dents adjacentes. Le cadre complet réduit significativement le temps de conception tout en maintenant les exigences de précision clinique. L'analyse qualitative utilisant des systèmes de classification de contact validés cliniquement montre des améliorations mesurables dans les motifs de contact occlusal, avec l'approche de contrainte spatiale réduisant les zones de contact problématiques par rapport aux méthodes sans modélisation spatiale explicite.

La validation technique à travers les ensembles de données expérimentales confirme l'efficacité des composants du cadre proposé. Les améliorations systématiques démontrent l'importance de la modélisation explicite des contraintes spatiales et de l'intégration de la ligne marginale pour la génération automatisée de couronnes. Les gains de performance substantiels à travers plusieurs métriques d'évaluation valident la progression de la complétion de nuages de points à la génération directe de mailles avec application complète des contraintes de relations spatiales.

Les méthodes développées dans ce projet contribuent significativement au développement de

systèmes de conception dentaire automatisés. Les travaux futurs devraient se concentrer sur les améliorations d'efficacité computationnelle par des structures de données spatiales basées sur l'octree pour réduire le temps de traitement tout en maintenant la précision. De plus, les approches alternatives de raffinement de surface utilisant les harmoniques sphériques pourraient fournir un contrôle morphologique plus précis par des méthodes analytiques plutôt que des approches basées sur l'apprentissage. Suite à la validation complète de la précision de génération de couronnes, des efforts supplémentaires devraient être dirigés vers l'intégration de modèles de diffusion pour améliorer la diversité et la qualité des géométries de couronnes générées, et l'extension du cadre aux flux de travail prothétiques complets incluant la conception de ponts et la planification de restauration d'arcade complète.

ABSTRACT

Dental crown restoration represents one of the most technically demanding procedures in restorative dentistry, traditionally requiring extensive manual design by skilled technicians for each crown. This time-intensive, operator-dependent process creates significant bottlenecks in clinical workflows while introducing variability in quality and precision. As clinical demands grow for faster, more consistent, and personalized restorations, the need for intelligent, automated solutions becomes increasingly urgent. Furthermore, existing computer-aided design tools rely heavily on template libraries and manual adjustments, limiting their ability to handle the complex geometric constraints required for proper crown fit, occlusal relationships, and aesthetic requirements. To address these limitations, researchers have explored machine learning approaches for automated crown generation. Early methods formulated crown design as a 2D image synthesis problem, but these approaches were fundamentally limited by information loss from 3D-to-2D projection and could not capture complete crown morphology. Subsequent approaches employed 3D representations such as point clouds and voxel grids, but lacked explicit modeling of critical spatial relationships including margin line boundaries, occlusal contact patterns, and adjacent tooth interactions. While these methods demonstrated feasibility of automated crown generation, they produced geometrically accurate but smooth surfaces lacking the fine-scale morphological details essential for clinical acceptance, and failed to prevent crown penetration into adjacent structures or optimize contact relationships with opposing dentition. The main objective of this thesis is to develop a comprehensive framework for fully automated dental crown generation that addresses the complete clinical pipeline from geometric reconstruction to spatial constraint modeling. We formulate crown design as a spatially-constrained 3D shape completion problem, where a transformer-based architecture generates anatomically accurate crowns from prepared teeth and their surrounding dental context. To address these challenges, we present four complementary methodological advances. First, we established a point cloud completion approach using transformer encoder-decoder architecture with margin line input integration, demonstrating significant improvements in boundary accuracy but requiring post-processing for mesh conversion. Second, we developed the first end-to-end mesh generation pipeline (DMC) using differentiable Poisson surface reconstruction, eliminating post-processing requirements and achieving direct mesh output with 0.062 average Chamfer Distance. Third, we introduced advanced loss functions including a dedicated margin line loss function and InfoCD loss for enhanced point alignment, achieving 12.32% reduction in Chamfer Distance and 46.43% reduction in MSE compared to DMC. Fourth, we implemented comprehensive

spatial constraint modeling through margin line input integration combined with antagonist interaction loss and intersection loss functions, ensuring proper occlusal relationships and preventing crown penetration into adjacent structures. Comprehensive evaluation on real clinical datasets comprising 557 cases across all tooth positions demonstrates substantial improvements over state-of-the-art methods. Our framework achieves 35.9-40.6% improvements in geometric accuracy compared to existing approaches. The spatial constraint integration reduces maximum boundary fitting errors from 1.37 to 0.74 mm and 58.4% reduction in variability, while antagonist interaction loss contributes 9.51% improvement in occlusal alignment and intersection loss substantially reduces crown penetration into adjacent teeth. The complete framework dramatically reduces design time while maintaining clinical accuracy requirements. Qualitative analysis using clinically validated contact classification systems shows measurable improvements in occlusal contact patterns, with the spatial constraint approach reducing problematic contact areas relative to methods without explicit spatial modeling. Technical validation across the experimental datasets confirms the effectiveness of the proposed framework components. The systematic improvements demonstrate the importance of explicit spatial constraint modeling and margin line integration for automated crown generation. The substantial performance gains across multiple evaluation metrics validate the progression from point cloud completion to direct mesh generation with comprehensive spatial relationship enforcement. The methods developed in this project significantly contribute to the development of automated dental design systems. Future work should focus on computational efficiency improvements through octree-based spatial data structures to reduce processing time while maintaining accuracy. Additionally, alternative surface refinement approaches using spherical harmonics could provide more precise morphological control through analytical methods rather than learning-based approaches. Following comprehensive validation of crown generation accuracy, further efforts should be directed towards integrating diffusion models to enhance the diversity and quality of generated crown geometries, and extending the framework to complete prosthetic workflows including bridge design and full-arch restoration planning.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ACRONYMS

AdaIN	Adaptive Instance Normalization
AI	Artificial Intelligence
CAD	Computer-Aided Design
CAM	Computer-Aided Manufacturing
CBCT	Cone-Beam Computed Tomography
CD	Chamfer Distance
CNN	Convolutional Neural Network
COV	Coverage
DAIS	Dental AI System
DBSCAN	Density-Based Spatial Clustering of Applications with Noise
DCD	Density-aware Chamfer Distance
DCPR-GAN	Dental Crown Prosthetic Restoration GAN
DGCNN	Dynamic Graph Convolutional Neural Network
DMC	Dental Mesh Completion
DPSR	Differentiable Poisson Surface Reconstruction
EMD	Earth Mover's Distance
FDI	World Dental Federation
FFT	Fast Fourier Transform
FPD	Fréchet Point Cloud Distance
FPS	Farthest Point Sampling
FSC	Few-point Shape Completion
GAN	Generative Adversarial Network
GCN	Graph Convolutional Network
GPU	Graphics Processing Unit
HyperCD	Hyperbolic Chamfer Distance
InfoCD	Information-theoretic Contrastive Chamfer Distance
IVADO	Institut de valorisation des données
KNN	K-Nearest Neighbors
LSGAN	Least Squares GAN
MEDTEQ	Medical Technologies
MICCAI	Medical Image Computing and Computer Assisted Intervention
MLP	Multi-Layer Perceptron
MMD	Minimum Matching Distance

MSE	Mean Square Error
NLP	Natural Language Processing
NKSR	Neural Kernel Surface Reconstruction
NSERC	Natural Science and Engineering Research Council of Canada
OBB	Oriented Bounding Box
PCA	Principal Component Analysis
PDE	Partial Differential Equation
PSR	Poisson Surface Reconstruction
PT	Point Transformer
SA-IFN	Self-Attention Implicit Function Networks
SAP	Shape as Points
SDF	Signed Distance Field
SP-GAN	Sphere-Guided 3D Shape Point Cloud GAN
SPIE	International Society for Optics and Photonics
TSP	Traveling Salesman Problem
VAE	Variational Autoencoder

CHAPTER 1 INTRODUCTION

Dental restorations play a vital role in maintaining oral health, functionality, and aesthetics. When a tooth is damaged, decayed, or missing, it often requires reconstruction through various forms of dental restoration, such as fillings, bridges, implants, dentures, or crowns. Among these, the design and fabrication of dental crowns, also referred to as dental caps, remain one of the most common and intricate procedures in restorative dentistry.

Teeth are subject to wear and decay, which may alter their shape and size. Restoring their original form not only enhances chewing efficiency and speech clarity but also contributes to overall facial appearance and confidence. Dental crowns, in particular, are customized to restore the external surface and occlusal function of individual teeth. As shown in Figure 1.1, the crown replaces the visible portion of a prepared tooth. Due to anatomical variability among patients, each crown must be uniquely designed to match the shape, alignment, and functional dynamics of surrounding teeth. The current workflow for crown fabrication is

Figure 1.1 Dental crown restoration showing the prosthetic cap covering a prepared tooth [2].

heavily reliant on the skill and expertise of dental technicians, even with the aid of modern computer-aided design (CAD) tools. While these tools assist in digital modeling and simulation, the process remains time-consuming and labor-intensive. Each case demands careful attention to occlusal relationships, anatomical fit, and material constraints, making the manual design of crowns susceptible to inconsistencies in both quality and functional performance.

One of the most critical constraints in crown generation is occlusion, the way upper and lower

teeth align and contact during various mandibular movements such as biting or chewing. Improper occlusal relationships can result in discomfort, impaired function, and long-term dental complications. Figure 1.2 illustrates the concept of dental occlusion, highlighting the importance of contact points between opposing teeth.

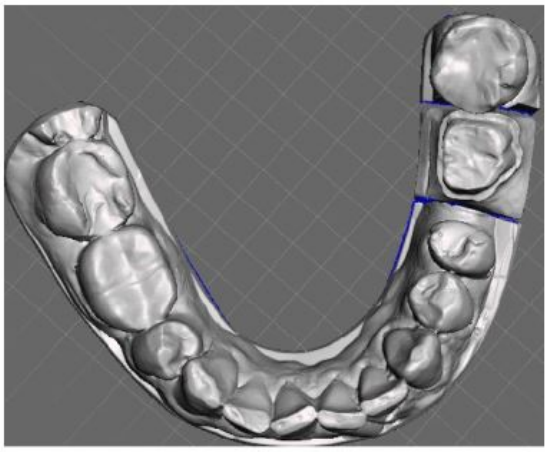
Figure 1.2 Dental occlusion showing the critical contact relationships between upper and lower [3].

Additional constraints include the necessity for proper spacing between the crown and adjacent or opposing teeth, ensuring that the crown does not interfere with neighboring structures or gingival tissue. The crown must also maintain a minimum material thickness to withstand functional loads while providing sufficient retention and comfort. These clinical requirements underscore the complexity of dental crown design.

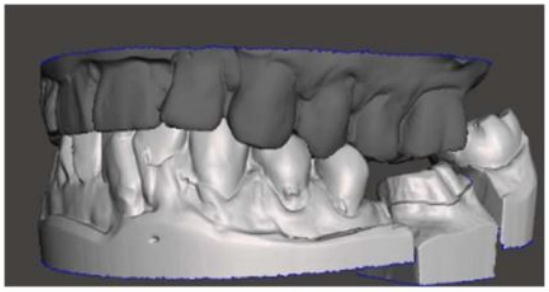
Each dental crown is composed of distinct geometric regions: the crown shell, which defines the outer surface; the crown bottom, which fits onto the prepared tooth structure; and the margin line, which marks the boundary of contact between the crown and the tooth preparation. Figure 1.3 illustrates these components, alongside the input data used in digital restoration workflows. These inputs typically include a digital scan of the dental arch, highlighting the prepared tooth, and a contextual model representing neighboring and opposing teeth. This data provides the anatomical and functional context required to design a personalized, clinically viable crown.

Computer vision has progressed significantly in recent years, leading to the successful deployment of discriminative tasks such as object detection, recognition, completion, and semantic segmentation across real-world applications. Among these advances, the understanding and processing of three-dimensional (3D) shapes has gained particular attention due to their relevance in medical diagnostics and industrial design applications requiring spatial understanding and precision.

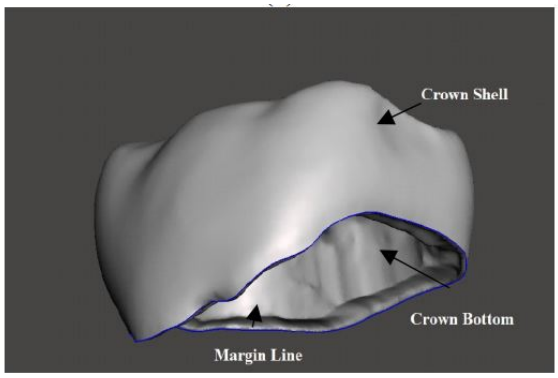
3D models provide natural and detailed representations of anatomical structures, aligning



(a) Dental arch with the prepared tooth

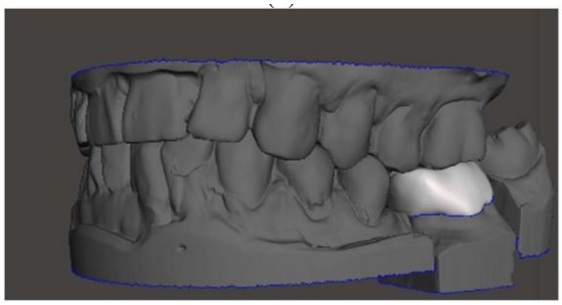


(b) Dental context, master and antagonist arch



(c) Designed crown

(d) Margin line



(e) Designed crown in context

Figure 1.3 Illustration of dental crown design workflow: (a) dental arch with prepared tooth, (b) anatomical context including neighboring and opposing teeth, (c) designed crown, (d) margin line extracted from the shell, and (e) crown integrated within the dental context. [4].

closely with clinical requirements for geometric accuracy and functional relationships. The ability to generate high-quality 3D shapes that exhibit fine geometric detail and structural

integrity is paramount in medical applications. Moreover, practical applications increasingly demand controllable generation processes that allow clinical constraints and anatomical context to guide the final output.

With the growing accessibility of 3D scanning technologies, capturing anatomical structures has become routine in medical practice. However, due to occlusions, scanning artifacts, or procedural limitations, many acquired models are incomplete, missing critical geometric information. This has led to the emergence of 3D shape completion as a fundamental challenge: reconstructing missing portions of 3D geometry in ways that are faithful to observed structures and anatomically plausible in extrapolated regions.

A central difficulty lies in the representation of 3D data. Unlike 2D images with regular grid structures, 3D shapes can be represented through various forms, including point clouds, voxel grids, triangle meshes, and implicit surfaces, each with distinct advantages and limitations. Point clouds offer computational efficiency but lack topological information, while voxel grids provide regular structure but are memory-intensive and resolution-limited. Meshes encode surface connectivity but present irregular structures challenging to process, whereas implicit representations offer modeling flexibility at increased computational cost. The choice of representation directly impacts both performance and clinical applicability of shape completion models.

The convergence of recent advances in deep learning and 3D processing with the limitations identified in current dental CAD workflows presents compelling opportunities for intelligent automation [16]. Crown generation can be effectively formulated as a constrained 3D shape completion problem, where the objective is to reconstruct complete crown geometry from partial dental scan data while satisfying clinical requirements for spatial relationships, occlusal function, and morphological accuracy. This formulation transforms the traditional manual design challenge into a learning problem where anatomical context guides automated generation of patient-specific restorations.

Unlike general 3D shape completion tasks, dental crown generation requires explicit modeling of clinical constraints, including margin line precision, antagonist relationships, and adjacent tooth interactions. The integration of spatial constraint modeling with advanced 3D generation techniques offers the potential to address these requirements systematically, moving beyond current CAD limitations toward fully automated systems capable of generating clinically viable crowns that meet both geometric and functional specifications.

1.1 Dental Crown Application

1.1.1 Dental anatomy

To understand the task of dental crown generation from a computational perspective, it is essential to first have a foundational understanding of human dental anatomy. Teeth are vital for a range of biological and functional purposes, including food digestion, speech articulation, and facial structure support. In a fully developed adult dentition, each dental arch typically contains 16 teeth: 2 incisors, 1 canine, 2 premolars, and up to 3 molars on each side. However, third molars, commonly known as wisdom teeth, are often absent in many adults due to impaction or extraction.

The precise location and categorization of each tooth are critical in both clinical and computational design processes. Tooth positions are labeled using the World Dental Federation (FDI) notation system. According to this system, the dental arches are divided into four quadrants: upper right (quadrant 1), upper left (quadrant 2), lower left (quadrant 3), and lower right (quadrant 4), as illustrated in Figure 1.4. The system utilizes two digits to identify each tooth, where the first digit indicates the quadrant and the second digit indicates the tooth position within that quadrant (numbered 1-8 from the center outward). For example, tooth 11 refers to the central incisor in the upper right quadrant, tooth 21 is the central incisor in the upper left quadrant, and tooth 48 represents the third molar (wisdom tooth) in the lower right quadrant [5].

1.1.2 Current approach to design a crown by a technician

The traditional dental crown fabrication process includes multiple manual and semi-digital steps. Initially, the dentist prepares the affected tooth by removing any decay and shaping it to fit a crown. Following the preparation, a dental impression is taken, either through physical molding or intraoral scanning, and sent to a dental laboratory.

In the laboratory, the design and fabrication of the crown may proceed via manual methods or through the use of CAD software such as ExoCAD. Manual design typically involves creating a negative mold of the prepared and adjacent teeth, into which a casting material is poured and cured at high temperatures to produce the crown. This process, although effective, is time-intensive and requires expert handling.

In digital workflows, CAD software assists technicians by providing modeling tools and visualization aids. However, a significant amount of human input remains necessary. The technician first orients the scanned model since software cannot autonomously distinguish

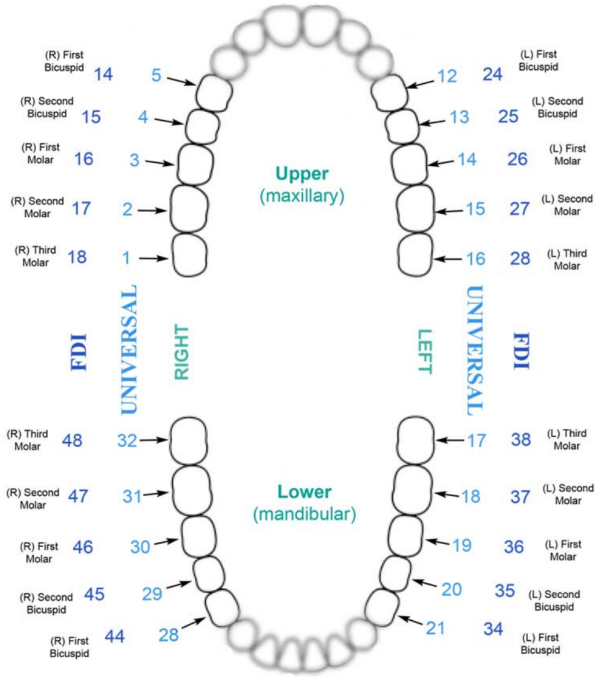


Figure 1.4 one complete dental arch [5]

between the occlusal and gingival directions. Next, the margin line, representing the boundary between the prepared tooth and gingiva, is manually defined and adjusted. A cement gap, typically around 0.08 mm, is introduced between the crown and the preparation to accommodate manufacturing tolerances and adhesive space.

Further, the technician controls the margin line thickness at the base of the crown, usually adjusting it to about 0.15 mm. CAD systems often allow mirroring a contralateral crown or copying an adjacent tooth to guide the new design. These templates serve as initial shapes, which are then modified to match the patient's anatomy.

The design of the crown's outer shell is semi-automated. CAD tools suggest a generic morphology for the crown shell, but the technician must manually refine its size, orientation, and alignment to achieve proper occlusion with the antagonist (opposing tooth) and establish contact with adjacent teeth. This step involves adjusting the cusps, ridge height, and contact points and often requires iterative refinement.

Ensuring precise contact with the neighboring teeth and maintaining an occlusal gap of approximately 0.1 mm from the opposing tooth are key clinical constraints. These adjustments are crucial to allow natural jaw closure and proper chewing function. However, some intricate details, such as crafting natural grooves or determining optimal contact patterns, remain largely dependent on the technician's experience and manual effort. As in Figure 1.5.

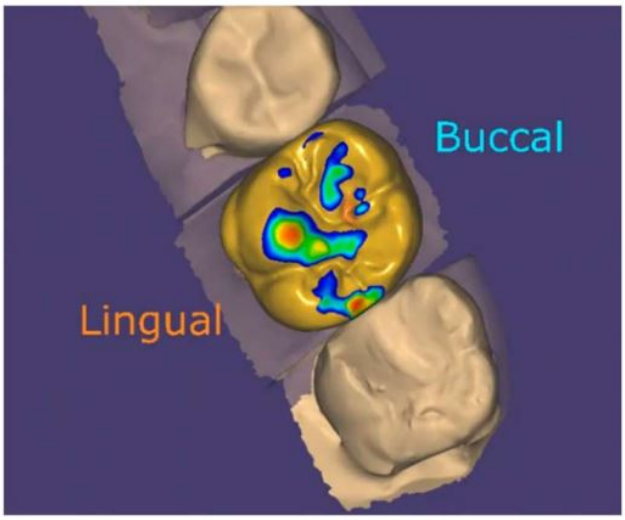


Figure 1.5 Design of dental crown with CAD software [4].

Overall, designing dental crowns that restore both aesthetic and functional properties demands a high level of expertise. Even with the aid of CAD tools, the current workflow is time-consuming and dependent on human input. This highlights the need for intelligent, automated systems capable of generating personalized crown designs based on the specific anatomy and occlusion context of each patient.

1.2 Goal of the research

The primary objective of this research is to develop a comprehensive deep learning framework for fully automated dental crown generation that addresses the complete restoration pipeline from geometric reconstruction to functional integration. Crown design represents a fundamental 3D shape completion challenge—reconstructing the missing portion of a prepared tooth while maintaining precise geometric and functional relationships with surrounding structures.

The complexity of this task extends far beyond conventional 3D shape completion applications. As detailed in the current workflow, dental technicians must simultaneously address multiple competing objectives: achieving precise margin line fit, establishing proper contact points with adjacent teeth, ensuring appropriate occlusal relationships with opposing dentition, and creating morphologically accurate surface features that support masticatory function. Each of these requirements demands geometric precision and functional understanding that exceeds typical shape completion tasks, where visual similarity alone is often the primary concern.

This research systematically addresses three fundamental challenges that bridge the gap between automated generation capabilities and clinical requirements: precise boundary fitting through margin line integration, anatomically valid spatial relationships through constraint modeling, and clinically realistic surface morphology that preserves essential functional features. The work extends beyond geometric accuracy to ensure that generated crowns maintain proper contact with adjacent teeth and establish appropriate occlusal relationships with opposing dentition.

The work progresses through four interconnected innovations that demonstrate systematic advancement from foundational concepts to practically viable solutions. Beginning with transformer-based architectures that leverage attention mechanisms for complex spatial relationship learning, advancing through end-to-end mesh generation that eliminates error accumulation from multi-stage pipelines, refining through advanced loss functions that enhance geometric accuracy and point distribution quality, and culminating in comprehensive spatial constraint modeling that addresses the complete requirements for patient-specific crown generation. These contributions establish new benchmarks in medical shape completion while addressing the practical requirements for transforming dental restoration workflows through reduced manual design time, improved consistency across different skill levels, and enhanced accessibility to high-quality dental care.

1.3 Thesis overview

The remainder of this thesis is organized as follows:

- **Chapter 2** presents a comprehensive review of related work. It introduces foundational concepts in deep learning for 3D shape completion, surveys standard approaches to 3D data representation, and discusses recent trends and persistent challenges. The chapter also highlights gaps in the current literature that motivate this research and outlines the specific objectives formulated to address them.

- **Chapter 3** defines the core problem addressed in this thesis and presents the detailed research objectives. It frames the development of a fully automated dental crown generation system as a structured 3D shape completion task.

- **Chapter 4** presents the first peer-reviewed publication arising from this research, which introduces a transformer-based framework for dental crown generation using conditional shape completion with margin line integration for improved boundary accuracy.

- **Chapter 5** details the second publication, which advances the initial approach by

developing an end-to-end mesh generation pipeline using differentiable Poisson surface reconstruction, eliminating post-processing requirements while maintaining geometric precision.

- **Chapter 6** showcases the third publication, which introduces advanced loss functions, including InfoCD loss for enhanced point alignment and adaptive query generation strategies, demonstrating systematic improvements in crown generation quality through architectural refinements.

- **Chapter 7** presents the fourth publication, which develops a comprehensive framework integrating spatial constraint modeling to generate dental crowns with proper morphology and functionality that satisfy patient comfort. This work incorporates antagonist interaction loss for occlusal contact patterns and intersection loss to prevent crown penetration into adjacent teeth, while integrating margin line data as direct network input to ensure both geometric precision and functional requirements essential for patient satisfaction.

- **Chapter 8** provides a general discussion of the contributions, evaluating how the proposed methods advance the state of the art while also addressing their limitations in comparison to existing techniques.

- **Chapter 9** concludes the thesis by summarizing the main contributions and outlining future research directions focused on computational efficiency improvements, alternative surface refinement approaches, and integration with advanced generative models for enhanced crown generation capabilities.

CHAPTER 2 LITERATURE REVIEW

The literature review is structured into several sections. It begins with an overview of three-dimensional (3D) data representation, followed by a discussion of non-learning-based methods in the shape completion domain. Subsequently, it examines state-of-the-art deep learning approaches, including mesh completion and generation techniques. The review also explores functional algorithms designed to control 3D predictions. Finally, it concludes by emphasizing automatic dental restoration as a pivotal application of 3D shape completion.

2.1 3D data representation

The growing availability of 3D data has motivated researchers to leverage it for addressing various computer vision challenges such as 3D object classification, recognition, and shape completion. The additional depth and complete geometric information provided by 3D data offer significant advantages, greatly enhancing the performance of numerous applications. However, raw 3D data collected from different scanning devices comes in diverse formats with varying structures and characteristics. The choice of data format used to represent 3D geometry is critical, as it directly impacts the types of algorithms that can be applied for learning from this data.

Common 3D data representations include point clouds, meshes, voxels, and spectral representations. Each of these representations has unique features with distinct advantages and limitations that we discuss below.

2.1.1 Point Cloud

A point cloud is a collection of 3D points defined by their coordinates (x, y, z), which collectively represent the external surface of an object. Point clouds are commonly generated by advanced 3D scanning technologies, including LiDAR, photogrammetry, and depth cameras, making them essential for applications such as autonomous vehicles and architecture [17].

The key strength of point clouds lies in their simplicity and direct representation of the 3D world. Unlike meshes or voxel grids, point clouds do not require connectivity information or volumetric data, making them easier to generate and manipulate. Their memory-efficient and sparse nature allows processing with relatively low computational requirements compared to denser representations [17].

However, point clouds are inherently unstructured and unordered, making tasks such as

segmentation and surface reconstruction challenging. The finite number of points can result in loss of detail, making it difficult to accurately model complex objects or produce smooth, continuous surfaces. Additionally, noise and irregularities are common in raw point clouds, necessitating preprocessing techniques before further analysis [17].

2.1.2 3D meshes

Mesh representation serves as the primary method for representing 3D objects' surfaces in computer graphics and computational geometry. A mesh consists of vertices, edges, and faces that work together to create a polyhedral shape in 3D space. Vertices serve as points in 3D space, edges form connections between pairs of vertices, and faces (commonly triangles) are defined by edges connecting vertices to create the surface structure [17].

The popularity of mesh representations stems from their balance between simplicity and expressive power. They can approximate complex geometries with arbitrary precision through careful adjustment of vertex and face density. This flexibility, combined with extensive hardware support in modern graphics systems, makes meshes efficient for both rendering and simulation tasks [17].

Despite their advantages, mesh representations face significant challenges. The complexity derives from their multi-element nature, where various connections exist between components, creating intricate relationships. The inherent irregularity of mesh structures—where the number of elements varies dramatically between geometries—poses challenges for processing algorithms. Additionally, their fixed vertex connection patterns make topology modifications difficult [18].

2.1.3 Voxel grid representation

A voxel (volumetric pixel) is the 3D counterpart of a 2D pixel. In 3D shape representation, a voxel grid provides a structured way to describe volumetric data by discretizing space into small cubic elements, each carrying an occupancy value indicating whether it is inside or outside the target shape [17].

Voxel grids are particularly advantageous for 3D convolutional neural networks, as they provide a structured format that enables direct application of convolutional operations. This makes them powerful for deep learning-based 3D shape analysis and medical imaging applications [17].

However, voxel grids suffer from high memory consumption, requiring $\mathcal{O}(n^3)$ memory complexity for an $n \times n \times n$ grid. Most voxel grids contain a large proportion of empty voxels,

leading to inefficient storage and increased computational overhead. Techniques such as sparse voxel representations and octrees are often employed to optimize storage efficiency while maintaining geometric fidelity [17].

2.1.4 Spectral and Spherical Harmonic Representations

Spectral methods provide a powerful alternative approach to representing and analyzing 3D geometry by projecting spatial information into the frequency domain. Spherical harmonics (SH) stand out as particularly effective for representing functions defined on spherical surfaces, offering compact and rotation-invariant shape analysis [19, 20]. This approach proves especially valuable for anatomical structures that can be approximated as genus-zero and closed, such as dental structures [21].

Spherical harmonics decompose shapes into frequency components, yielding coefficients that encode geometric information at varying levels of detail. Lower-order coefficients capture global structure, while higher-order coefficients preserve fine-grained surface details [19]. This hierarchical representation enables effective trade-offs between compactness and fidelity, making SH particularly suitable for shape completion tasks where global structure must be inferred from incomplete data [22].

The rotational invariance of SH coefficients makes them attractive for learning-based pipelines where geometric alignment cannot be guaranteed [23]. However, SH methods require shape parameterization onto a spherical domain, which may introduce distortion for complex topologies or highly non-spherical surfaces [21, 24].

2.1.5 Surface reconstruction

Surface reconstruction represents a fundamental process in 3D modeling, transforming point cloud data into complete surface representations. This process addresses the challenge of generating continuous, watertight surfaces from discrete point samples, which is essential for applications requiring manufacturable 3D models [25].

Surface reconstruction methods can be classified into explicit and implicit approaches. Explicit methods directly generate surface representations through triangulation or parametric modeling, offering computational efficiency but struggling with complex topologies [26]. Implicit methods define surfaces as iso-contours of mathematical functions that approximate input data, providing superior handling of topological changes and noise robustness at increased computational cost [27].

Signed distance functions (SDF) represent a particularly powerful implicit approach, encod-

ing surface position by computing the signed distance from any point to the closest surface location. The function equals zero on the surface, is negative inside the object, and positive outside, providing complete spatial understanding [28]. While implicit approaches offer flexibility in achieving arbitrary resolutions and handling complex geometries, they require post-processing through methods like marching cubes to generate visualizable surfaces.

Modern differentiable reconstruction techniques enable end-to-end optimization, allowing surface quality to be directly optimized within learning frameworks. The choice between reconstruction approaches depends on application requirements: explicit methods suit real-time applications requiring computational efficiency, while implicit methods excel in scenarios demanding topological robustness and noise handling [17].

In conclusion, the various forms of 3D data representation serve as the foundation for structuring, processing, and interpreting geometric information in computer vision. Each format offers unique strengths depending on resolution, application, and context. Understanding these representations is crucial before exploring the algorithms built upon them, particularly as modern approaches increasingly leverage hybrid representations that combine multiple formats to exploit their complementary advantages. The next section focuses on 3D shape completion, examining both traditional rule-based techniques and modern deep learning approaches that have significantly advanced the field.

2.2 3D shape completion

3D shape completion represents a fundamental challenge in computer vision: reconstructing missing portions of 3D geometry from incomplete observations while maintaining geometric fidelity and structural plausibility. This capability is essential for dental crown generation, where the objective is to reconstruct complete crown geometry from prepared tooth data and surrounding anatomical context. Understanding the evolution from conventional to deep learning-based approaches provides crucial insights into the methodological foundations underlying automated dental restoration systems.

2.2.1 Conventional methods in 3D shape completion

Prior to deep learning adoption, 3D shape completion relied on conventional approaches that can be categorized into geometry-based, alignment-based, and template-based methods. These techniques typically employ handcrafted priors, geometric assumptions, or predefined model databases to infer missing regions.

Geometry-based methods leverage intrinsic geometric properties such as surface continuity

and spatial symmetry to infer missing regions. Surface-focused strategies employ smooth interpolation techniques like Laplacian smoothing [29] and Poisson surface reconstruction [30] to locally fill holes while maintaining geometric consistency. Symmetry-driven methods detect potential axes or planes of symmetry and replicate structures across these symmetries to reconstruct missing geometry [31, 32]. While effective for well-structured objects with mild incompleteness, these approaches rely on strong geometric assumptions that limit their applicability to complex shapes or substantial missing data.

Alignment-based methods retrieve similar models from databases and align them with incomplete inputs, assuming missing parts can be inferred from structurally similar models [33,34]. Some approaches directly transfer entire models or sections, while others use deformed synthetic shapes or geometric primitives to fill incomplete regions [35, 36]. These methods offer versatility across shape categories but depend heavily on database quality and involve computationally intensive optimization procedures.

Template-based approaches leverage prior templates through direct retrieval, partial alignment, deformation, or geometric substitution strategies [33, 35, 36]. While interpretable and grounded in explicit geometric priors, their dependence on symmetry assumptions, continuity constraints, and database similarity restricts applicability in scenarios involving noise, severe incompleteness, or novel object categories. These limitations have motivated the shift toward data-driven, learning-based approaches capable of learning shape priors directly from datasets.

2.2.2 Deep learning methods for 3D shape completion

Deep learning approaches operate by stacking multiple transformation layers to extract abstract features from raw input, requiring minimal domain-specific prior knowledge. These data-centric methods have gained significant traction in 3D scene understanding tasks, including shape recognition, segmentation, and completion [6]. Deep generative models facilitate dimensionality reduction and feature extraction, enabling generation of new 3D shapes that adhere to learned semantic and spatial constraints.

Machine learning-based techniques learn mappings between partial and complete 3D shapes using encoder-decoder architectures. Many methods represent shapes using voxel grids suited for 3D convolution, while alternatives work directly with point sets to preserve local geometric details. A key challenge arises when predicting complete shapes from significantly incomplete inputs, as both point clouds and voxels provide limited input modality [37]. This limitation reduces generalizability, with some techniques performing well only on specific object categories or minor incompleteness cases.

The fundamental difficulties in 3D shape completion include accurately recovering both structural layout and detailed geometry. The unordered nature of point clouds complicates reconstruction of structurally coherent shapes, while achieving fine-grained accuracy requires restoring subtle features like symmetry and surface regularity. Recent approaches address these challenges through iterative refinement [38], global-local descriptor combination [39], and architectural enhancements like skip connections [40].

Modern deep learning approaches can be categorized into several methodological families:

Point-based methods

Point-based methods directly process unordered 3D point sets, pioneered by PointNet [41] for permutation-invariant point cloud processing. PointNet++ [42] introduced hierarchical architectures with multi-resolution feature learning, while FoldingNet [43] developed decoders that fold 2D grids into 3D surfaces. PCN (Point Completion Network) [44] established influential encoder-decoder frameworks as illustrated in Figure 2.1, combining Chamfer Distance and Earth Mover's Distance losses for geometric alignment.

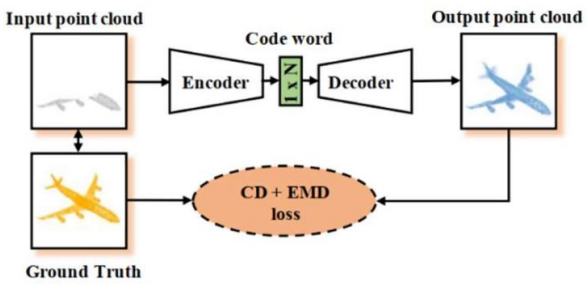


Figure 2.1 The illustration of an end-to-end network for point clouds completion [6]

Recent advances include ASFM-Net [38] with asymmetric Siamese autoencoders for iterative refinement, and capsule-based models leveraging dynamic routing for spatial relationship preservation. However, challenges remain in handling point density variations and capturing larger structural patterns due to limited receptive fields.

Convolution-based methods

Convolution-based methods adapt CNNs to volumetric representations, transforming point clouds into voxel grids for 3D convolution application as shown in Figure 2.2. Early approaches [45, 46] converted point clouds into binary occupancy grids, while GRNet [47] proposed gridding mechanisms to preserve contextual relationships during voxelization.

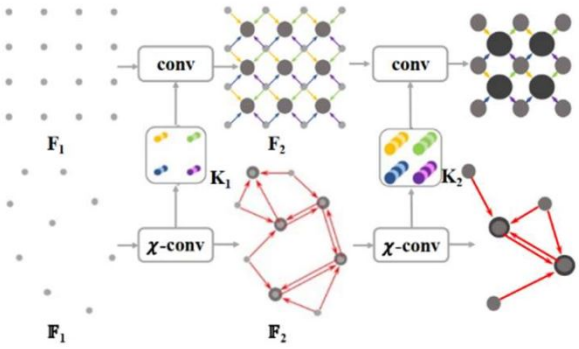


Figure 2.2 The illustration of hierarchical convolution on regular grids (top) and point clouds (bottom) [6]

Advanced methods like KPConv [48] define convolution operations directly in continuous 3D space using deformable kernels, while PointConv [49] computes convolutional weights dynamically based on input features. However, volumetric representations suffer from cubic memory scaling and inefficient storage due to empty voxel prevalence.

Graph-based Methods

Graph-based methods model point clouds as graphs where points serve as nodes with dynamically established connections, as illustrated in Figure 2.3. DGCNN [50] introduced EdgeConv for adaptive neighborhood capture at multiple abstraction levels, while LDGCNN [51] improved efficiency through direct multi-level feature aggregation.

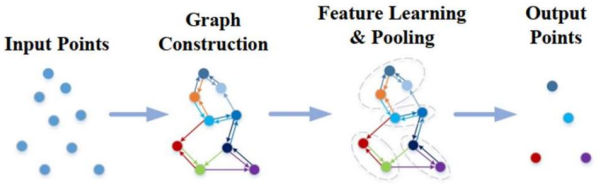


Figure 2.3 Graph-based network [6]

Recent approaches incorporate attention mechanisms for improved local context aggregation, with methods like CRA-Net [52] modeling spatial region dependencies through global contextual features. However, challenges persist in designing convolutional operators that adapt to dynamically scaled neighborhoods while maintaining computational tractability.

Folding-based methods

Folding-based decoders reconstruct detailed 3D point clouds by learning to deform 2D grids into 3D surfaces, as shown in Figure 2.4. FoldingNet [43] pioneered this approach by mapping regular 2D grids to 3D object surfaces through learnable folding functions that mimic physical transformations.

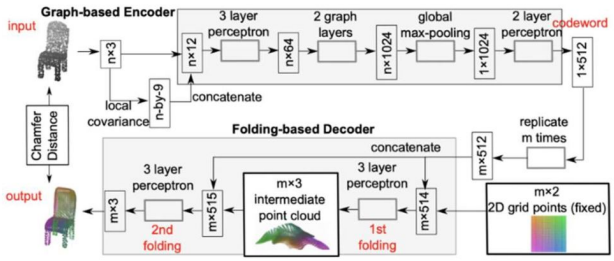


Figure 2.4 Folding-based network [6]

AtlasNet [53] extended this concept by treating reconstruction as unions of multiple parametric surface patches, while MSN [54] proposed coarse-to-fine frameworks using morphing decoders. Recent advances include hierarchical approaches like TopNet [55] with tree-structured decoders and adaptive mechanisms addressing static grid limitations.

GAN-based Methods

Generative Adversarial Networks extend 2D synthesis principles to 3D shape completion, with generators hallucinating missing geometry while discriminators enforce plausibility through learned shape priors, as illustrated in Figure 2.5. Early approaches like 3D-ED-GAN [56] combined voxel-based encoder-decoder GANs with recurrent networks for temporal dependency learning.

Advanced methods include hierarchical architectures like PF-Net [57] implementing coarse-to-fine strategies, and recent approaches integrating reinforcement learning [58] or style-based generation [59]. However, GAN-based methods face inherent training instability and mode collapse challenges while struggling to preserve local surface details.

Transformer-based methods

Transformer architectures leverage self-attention mechanisms for long-range dependency modeling without fixed spatial structures, as shown in Figure 2.6. PoinTr [9] reformulated comple-

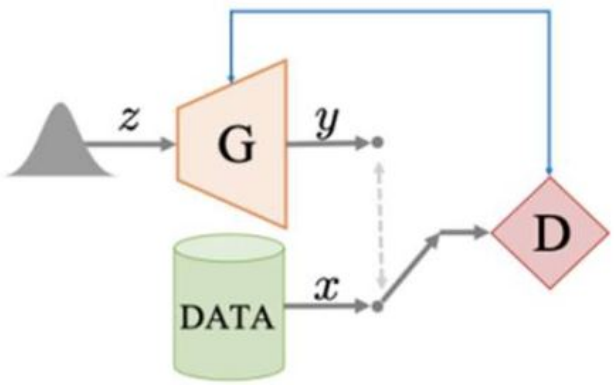


Figure 2.5 GAN-based network [6]

tion as set-to-set translation using position embeddings and geometry-aware modules, while AdaPoinTr [12] enhanced robustness through adaptive query generation and denoising tasks.

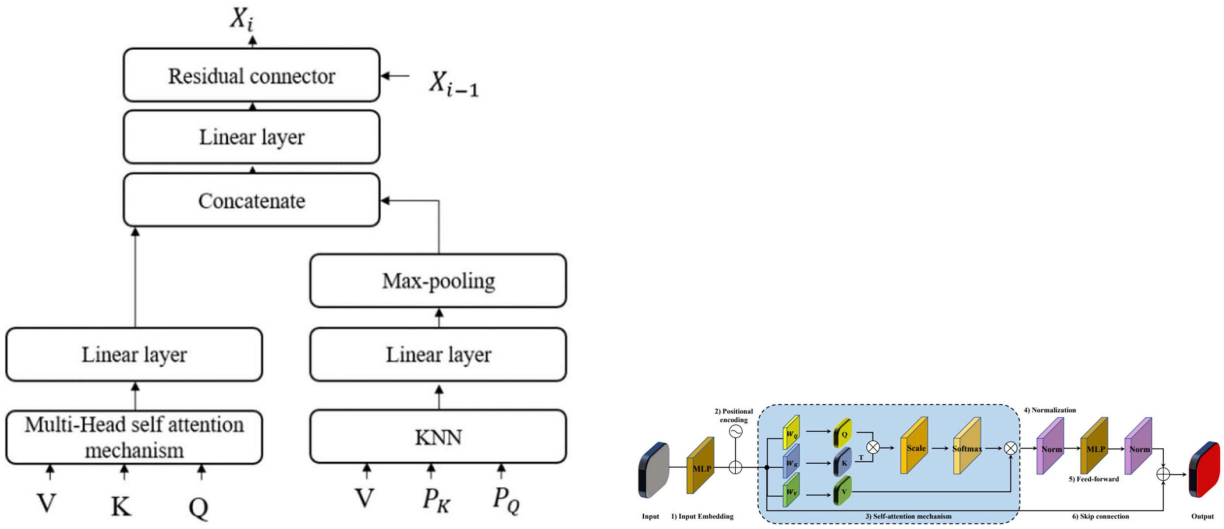


Figure 2.6 (a, b) shows the transformer architecture in general and in details [6]

Recent advances include SnowflakeNet [60] with iterative point deconvolution mechanisms and ShapeFormer [61] using vector quantized implicit functions for autoregressive completion. These methods excel at capturing global context and structural relationships essential for complex completion tasks.

Diffusion-based models

Diffusion-based models iteratively refine random noise into structured outputs through learned denoising steps, offering advantages in sample diversity and training stability over GANs, as illustrated in Figure 2.7. DiffComplete [62] frames completion as conditional generation with transformer-based denoising networks, while score-based approaches [63] use learned gradient fields for point distribution recovery.

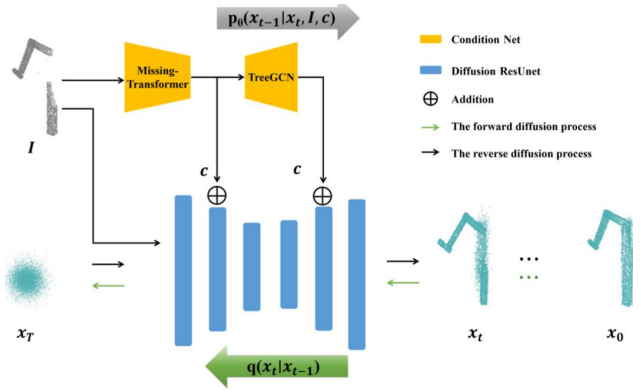


Figure 2.7 Diffusion-based network [7]

Recent developments integrate diffusion with other paradigms, including vector quantization [64] and multi-view guidance [65]. While offering high-fidelity reconstruction and stable training, diffusion models face computational challenges due to iterative sampling requirements and long inference times.

3D Mesh deformation

Mesh deformation methods address point cloud connectivity limitations by progressively deforming predefined mesh templates to match target shapes while maintaining vertex connectivity and topological structure. Pixel2Mesh [66] uses ellipsoid templates with fixed topology, while AtlasNet [67] employs surface parameterization for greater flexibility.

Recent approaches enable direct deformation parameter prediction from sparse inputs [68], improving computational efficiency through single forward passes. These methods excel at preserving surface continuity and enabling controllable deformation but may limit shape diversity due to template constraints.

Validation datasets

For validation of these methods, researchers predominantly rely on several public datasets that provide standardized benchmarks for 3D shape completion evaluation. ShapeNet [69] serves as the most comprehensive CAD-based dataset with 30,974 3D models across eight categories, typically using 16,384-point ground truth representations. Real-world validation employs datasets like KITTI [70] for LiDAR-captured automotive scenarios and ScanNet for indoor RGB-D scenes. Specialized benchmarks include Completion3D [55] for low-resolution testing, ShapeNet-34/55 [9] and MVP [71] for enhanced category coverage, and viewpoint variation for geometric deep learning evaluation. These datasets enable systematic comparison of completion methods across synthetic and real-world scenarios with varying complexity levels.

The evolution from conventional to deep learning approaches demonstrates significant advances in handling complex geometries, large missing regions, and diverse shape categories. However, challenges remain in achieving fine-grained geometric accuracy, maintaining structural consistency, and ensuring computational efficiency for practical applications. These considerations are particularly critical for dental applications requiring clinical-grade precision and reliability.

2.3 Automatic Dental restoration

Automatic dental restoration encompasses the computational methods required to process, analyze, and generate dental structures for clinical applications. This field relies on two fundamental components that directly support automated crown generation: segmentation for preprocessing dental scan data into usable inputs, and generation for creating complete crown geometries. Understanding these foundational technologies provides essential context for automated dental crown design systems.

2.3.1 Segmentation

Semantic segmentation in dental applications refers to the process of identifying and separating individual teeth from one another, as well as from adjacent anatomical structures such as the gingiva, as illustrated in Figure 2.8. In our research, segmentation serves a crucial preprocessing role by enabling the creation of structured input data for deep learning models. Specifically, we utilize segmentation to parse complete dental arches according to tooth labels and die identification, generating the anatomical context required for crown generation networks.

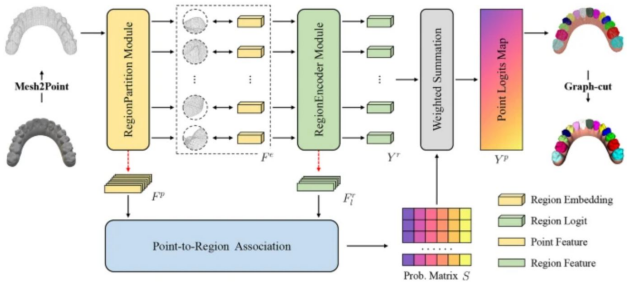


Figure 2.8 Dental arch segmentation [8]

The segmentation process involves partitioning scanned dental arches into distinct anatomical regions, including the prepared tooth (die), adjacent teeth, and opposing arch structures. This structured representation enables crown generation networks to understand spatial relationships and anatomical context essential for producing clinically viable restorations. While manual segmentation using commercial dental CAD software remains possible, this approach is labor-intensive and prone to operator variability [72].

Modern approaches leverage geometric deep learning to process non-Euclidean data structures inherent in 3D dental meshes and point clouds [73]. MeshSegNet [74] represents a foundational architecture that operates directly on 3D dental meshes, achieving high accuracy in tooth-level segmentation through spatial and geometric feature learning. However, performance degrades with atypical dental structures, motivating enhanced approaches that incorporate self-supervised pretraining [75] and hierarchical feature modeling [76].

Recent advances include transformer-based architectures like ToothFormer [77] that capture long-range spatial dependencies, multi-modal approaches like DentSeg++ [78] that fuse 2D and 3D features, and domain adaptation methods like AdaToothSeg [79] that enable robust performance across diverse scanning conditions. These developments have significantly improved segmentation accuracy and robustness, though challenges remain in handling incomplete scans, anatomical variability, and clinical data diversity.

2.3.2 Generation

Crown generation represents the core challenge in automated dental restoration, requiring the synthesis of anatomically accurate and functionally viable crown geometries from prepared tooth data and surrounding anatomical context. This process must satisfy both geometric precision requirements for proper fitting and functional constraints for occlusal relationships with opposing dentition [80].

Early generative approaches employed 2D methodologies, projecting 3D dental geometry into depth map representations for processing with established image generation frameworks. Yuan et al. [81] developed conditional GANs that generate crown surface structures from intraoral scanner data, incorporating perceptual loss, occlusal groove filters, and contact constraints to enhance surface realism. Tian et al. [82] extended this paradigm with two-stage conditional GANs designed to model occlusal relationships more accurately, while Hwang et al. [1] leveraged spatial profiles between opposing teeth to learn design rules from human-created crowns.

Despite promising results, 2D approaches suffer from information loss during 3D-to-2D projection and underutilization of spatial context, motivating the transition to fully 3D learning frameworks. SA-IFN [83] employs voxel-based approaches integrating self-attention mechanisms with implicit neural representations to model tooth morphology at multiple scales. While achieving low geometric error (Chamfer Distance of 0.069), voxel-based representations face resolution constraints and over-prediction in non-surface regions due to Marching Cubes reconstruction dependencies.

Recent approaches have emphasized volumetric completeness and transformer-based architectures. Zhu et al. [84] combine transformer-based point cloud completion with voxel reconstruction to generate watertight 3D meshes, prioritizing manufacturability over boundary sharpness. These developments demonstrate the field's evolution toward hybrid models that combine geometric priors with deep learning capabilities, incorporating anatomical knowledge and multi-modal signals to create restorations that satisfy both functional and manufacturing requirements.

The progression from 2D GAN-based pipelines to sophisticated 3D learning architectures reflects the growing emphasis on preserving fine-grained geometry, integrating geometry-aware mechanisms, and leveraging attention-based representations. Current state-of-the-art approaches increasingly incorporate opposing dentition modeling, implicit surface representations, and multi-resolution strategies to address the complex requirements of automated dental crown generation [85].

2.4 Summary of the literature

At the time of this research, automated dental crown generation represented an emerging field with limited dedicated solutions, creating substantial gaps between existing geometric/morphological capabilities and functional requirements. The few available dental crown generation approaches demonstrated significant limitations that prevented direct application

in dental restoration workflows.

Early dental crown generation methods relied primarily on 2D approaches, projecting 3D dental scans into depth map representations for processing with conditional GANs. While these methods achieved reasonable geometric similarity, they suffered from fundamental information loss during 3D-to-2D conversion and failed to capture complete crown morphology necessary for functional viability. The limited 3D approaches available employed voxel-based representations with implicit neural networks, achieving low geometric error but facing resolution constraints and over-prediction issues that compromised surface quality.

Furthermore, available methods lacked end-to-end mesh generation capabilities, requiring separate post-processing steps that introduced error accumulation and prevented comprehensive optimization. Loss function design remained limited to geometric similarity metrics without incorporating functional constraints such as margin line accuracy essential for dental applications.

Other limitations across existing work were the absence of spatial constraint modeling, margin line precision, adjacent and antagonist tooth interactions—fundamental requirements for functional crown generation. Existing approaches operated in isolation from surrounding anatomical context, producing geometrically plausible results that required extensive manual correction due to spatial relationship failures. This gap prevented direct application and impacted restoration functionality through compromised spatial accuracy and extended design times.

Transformer-based architectures showed promise for capturing long-range dependencies in point cloud completion tasks, yet their application to dental crown generation remained unexplored. The potential for integrating spatial constraint modeling with transformer-based shape completion represented an unaddressed opportunity for achieving functionally viable automated dental restoration systems.

These fundamental gaps in morphological precision, functional constraint modeling, and end-to-end optimization motivated the development of a comprehensive framework that addresses the complete restoration pipeline from geometric reconstruction to spatial relationship enforcement. The following chapters present systematic solutions to these limitations, establishing foundations for morphologically and functionally automated dental crown generation.

CHAPTER 3 OBJECTIVE AND GENERAL METHODOLOGY

3.1 Problem Statement

Building on the identified limitations in current dental crown generation approaches, the fundamental challenge lies in the absence of integrated frameworks that simultaneously address morphological precision and functional constraint satisfaction. Existing automated methods produce results requiring extensive manual correction, failing to achieve the dual requirements necessary for direct application in restoration workflows.

The core problem manifests in two dimensions: morphological limitations where current methods lack the precision required for accurate boundary fitting and surface detail preservation, and functional gaps where spatial relationships, occlusal interactions, and adjacent tooth constraints are not incorporated into the generation process. Current approaches focus primarily on shape similarity without considering the structural information of pair-wise interactions and global correlations essential for high-precision crown generation. Additionally, the appropriateness of different 3D representations for dental crown generation remains unclear, with most methods requiring separate post-processing steps that introduce errors and prevent end-to-end optimization.

Furthermore, existing automated crown generation methods operate in isolation from the complete dental arch context, failing to incorporate functionality restrictions that ensure proper integration with adjacent and opposing teeth. This dual limitation prevents the realization of truly automated, personalized crown generation systems that can produce clinically viable restorations without extensive manual intervention.

This thesis addresses the fundamental research question of whether automated dental crown generation can simultaneously achieve the morphological precision and functional constraint satisfaction required for practical application.

3.2 Objectives

To answer this research question, this thesis establishes two primary objectives that systematically address the identified limitations through integrated frameworks combining geometric accuracy with spatial relationship modeling. These objectives are detailed in the following sections.

3.2.1 Objective 1: Morphological Crown Generation

Our first objective consists of developing a deep learning method that learns structural information of pair-wise interactions and global correlations for dental crown generation with high precision. The proposed method determines whether mesh-based 3D representation is appropriate for dental crown generation and investigates novel approaches for 3D mesh completion that satisfy geometric constraints and preserve patient-specific details. After the training process, these networks should be able to generate end-to-end 3D mesh outputs without requiring separate post-processing steps. Therefore, this can achieve geometric accuracy and morphological precision suitable for direct manufacturing applications.

3.2.2 Objective 2: Functional Integration

Our second objective consists of developing automated crown generation that incorporates functionality restrictions and ensures proper integration within the complete dental arch context. The proposed approach investigates and develops unique architectures for patient-specific dental crown generation that satisfy functional requirements including proper spatial relationships with adjacent teeth, appropriate occlusal contact patterns with opposing dentition, and prevention of structural penetration. After the training process, the generated crowns should maintain adequate chewing function while ensuring aesthetic integration. Therefore, this can produce crowns suitable for direct clinical application without extensive manual correction.

3.3 General Methodology

The methodology addresses both morphological precision and functional constraint satisfaction through a systematic four-stage framework that progressively builds from foundational architectural development to comprehensive spatial relationship modeling. Each stage represents a distinct contribution that systematically advances toward the overarching goal of automated dental crown generation suitable for direct clinical application.

The methodology begins by establishing transformer-based architectures for learning structural information through pair-wise interactions and global correlations in dental crown generation. This foundational stage addresses the core challenge of achieving high-precision morphological reconstruction by developing deep learning methods that can process 3D dental context with explicit margin line integration. The approach formulates crown generation as spatially-constrained shape completion, incorporating boundary constraints directly into the learning process to ensure accurate crown-tooth interface fitting. This stage establishes

the fundamental capability for automated crown generation while providing explicit geometric constraints essential for subsequent developments. The detailed implementation and evaluation of this foundational framework is presented in Chapter 4.

Building upon the foundational framework, the methodology advances to investigate whether mesh-based 3D representation is appropriate for dental crown generation through end-to-end mesh completion. This stage addresses the critical limitation of separate post-processing steps by integrating differentiable surface reconstruction directly into the neural network architecture. The approach enables direct mesh generation without separate reconstruction procedures, eliminating error accumulation while maintaining geometric precision essential for dental applications. The framework incorporates dual supervision through both point cloud and mesh representations, ensuring high-quality mesh outputs suitable for direct manufacturing while determining the optimal 3D representation strategy for dental crown generation. The comprehensive analysis and results of this end-to-end approach are detailed in Chapter 5.

The methodology progresses to achieve state-of-the-art morphological precision through sophisticated architectural refinements and specialized loss functions. This stage incorporates adaptive query generation capabilities and advanced loss formulations including contrastive Chamfer distance that combines traditional metrics with contrastive learning principles. The approach addresses the challenge that existing generative models struggle with fine details crucial for dental function by encouraging uniform point distribution while maintaining geometric accuracy. Specialized margin line loss functions operate on critical boundary regions, ensuring that crown generation preserves significant local information required for patient-specific design rather than synthesizing average shapes. The advanced architectural innovations and their impact on morphological precision are thoroughly examined in Chapter 6.

The final stage achieves comprehensive functionality by incorporating spatial constraint modeling that ensures generated crowns satisfy functional requirements within the complete dental arch context. This stage develops unique architectures that integrate functionality restrictions including spatial occlusion relationships with opposing teeth and non-overlapping constraints with existing teeth and gingiva. The framework incorporates antagonist interaction loss for proper occlusal contact patterns and intersection loss preventing crown penetration, transforming crown generation from isolated shape completion to comprehensive spatial relationship modeling. The approach addresses the complete clinical pipeline by ensuring generated crowns maintain adequate chewing function while achieving aesthetic integration suitable for direct application. The complete functional integration framework and

its clinical validation are presented in Chapter 7.

The unified methodology integrates all four stages through progressive development that systematically addresses morphological precision and functional constraint satisfaction. The approach establishes a coherent progression from basic automation capabilities to clinical-ready deployment, with each stage building upon previous achievements while introducing new capabilities. The framework balances geometric accuracy with spatial relationship enforcement through multi-objective optimization strategies, enabling the production of patient-specific crowns that satisfy both morphological precision and functional viability requirements. This comprehensive approach addresses the complete restoration pipeline from geometric reconstruction to functional validation, establishing foundations for automated crown generation systems that meet clinical standards for direct application in dental restoration workflows.

CHAPTER 4 ARTICLE 1: IMPROVING THE QUALITY OF DENTAL CROWN USING A TRANSFORMER-BASED METHOD

Golriz Hosseinimanesh[a], Farnoosh Ghadiri[b], Ammar Alsheghri[a], Ying Zhang[a], Julia Keren[c], Farida Cheriet[a], François Guibault[a]

[a]Polytechnique Montréal, Department of Computer and Software Engineering, P.O. Box 6079, Montréal, H3C 3A7, Québec, Canada

[b]Centre d'intelligence artificielle appliquée (JACOBB), Montréal, Québec, Canada

[c]Intellident Dentaire Inc., Montréal, Québec, Canada

Presentation: This work was published in the conference proceedings of SPIE Medical Imaging 2023: Image Processing on November 1, 2022 [86].

4.1 CRediT authorship contribution statement

Golriz Hosseinimanesh: Methodology, Software, Validation, Investigation, Data Curation, Writing - Original Draft. **Farnoosh Ghadiri**: Writing - Review& Editing. **Ammar Alsheghri**: Writing - Review& Editing. **Ying Zhang**: Writing - Review& Editing. **Julia Keren**: Conceptualization, Resources. **Farida Cheriet**: Supervision. Writing - Review & Editing. **François Guibault**: Supervision, Project administration, Funding acquisition, Resources, Writing - Review & Editing.

4.2 Abstract

Designing a synthetic crown is a time-consuming, inconsistent, and labor-intensive process. In this work, we present a fully automatic method that not only learns human design dental crowns, but also improves the consistency, functionality, and esthetic of the crowns. Following success in point cloud completion using the transformer-based network, we tackle the problem of the crown generation as a point-cloud completion around a prepared tooth. To this end, we use a geometry-aware transformer to generate dental crowns. Our main contribution is to add a margin line information to the network, as the accuracy of generating a precise margin line directly,determines whether the designed crown and prepared tooth are closely matched to allowappropriateadhesion.Using our ground truth crown, we can extract the margin line as a spline and sample the spline into 1000 points. We feed the obtained margin line along with two neighbor teeth of the prepared tooth and three closest teeth in the opposing jaw. We

also add the margin line points to our ground truth crown to increase the resolution at the margin line. Our experimental results show an improvement in the quality of the designed crown when considering the actual context composed of the prepared tooth along with the margin line compared with a crown generated in an empty space as was done by other studies in the literature.Our code is available at :

"https://github.com/Golriz-code/shellGeneration/"

Keywords: 3D dental model, Point Cloud completion, Surface reconstruction, Dental crown generation, Margin line

4.3 Introduction

If a tooth is missing, decayed, or fractured, it may need a dental crown. Professional dental practitioners need a significant amount of effort, skill, and time to design a personalized crown for each patient with current digital technologies for dentistry, such as computer-aided design software. Besides, there are functional and geometric constraints that put limits on possible solutions and desired precision. In a nutshell, the designed crown should not overlap existing teeth, and it must respect a space occlusion relationship with the opposing teeth to provide adequate chewing function. Each dental crown consists of a crown shell, crown bottom, and margin line. Here, our goal is to generate the crown shell using a fully automated algorithm, given a prep and its surrounding (context).

4.3.1 Related work

Few studies have been conducted on automating dental crown design. The author [1] downscaled 3D dental scans into 2D depth images and then created depth images of dental crowns in a way similar to pix2pix paper [87]. In this way, they may make use of improvements in image generation to achieve good performance. But information is lost during their conversion from 3D to 2D. Ping et al [83] presented SA-IFN, a voxel-based method for completing 3D tooth scans that integrate the functions of implicit representations with the self-attention mechanism. Voxel-based methods, on the other hand, need several unnecessary occupancy predictions, which reduces the output resolution. Zhu et al. [84] applied a new voxel-based reconstruction after completing the tooth point clouds using a transformer-based model. However, as opposed to reconstructing a thin surface with a boundary, their technique creates a watertight mesh. 3D deep learning has significantly advanced in recent years. Due to the frequent partial nature of 3D data coming from inaccurate scans, point cloud completion is essential for 3D computer vision. Predicting an object's complete shape from a

partial observation is a challenging task. This problem has received increasing attention in the computer vision community since the pioneering works of PointNet and FoldingNet [43]. As mentioned in Fei et al. [6], models can be categorized into point-based, convolution-based, graph-based, folding-based, GANbased, and transformer-based methods. Among these architectures, transformers [88] suggest a unique attention mechanism that uses many levels of self-attention and cross-multi-head attention to interchange data between input and output. Although the attention mechanism was initially intended for NLP tasks, it has recently been demonstrated to be beneficial for a variety of computer vision problems. Each layer of the transformer then consists of an attention mechanism, which allows for interaction between inputs at different positions, followed by a position-wise fully connected network, which is applied to all positions independently.

4.3.2 New work to be presented

There is currently no intelligent technology that can automatically assist technicians in generating dental crowns for prepared teeth, also referred to as 'preps'. In this paper, we propose a 3D shape completion transformer-based network to automatically reconstruct a patient-specific dental crown shape in which the dental crown fits onto the prepared tooth perfectly at the margin line. Our contributions are summarized as follows: 1) We propose to add the margin line information to our ground truth crown to increase the resolution at the margin line; 2) Our network is able to generate teeth in any position. In Contrast with [89] which only trained their network for the specific tooth, our network is able to generate teeth in any position.

4.4 Methods

Our approach uses an end-to-end deep learning architecture to generate the dental crown from given partial arches. We train our network on all incisors, canines, molars, and premolars crowns.

4.4.1 Data preprocessing

We first create a context from a given 3D scan. Context refers to the two neighbor teeth of a prepared tooth and the three teeth in the opposing jaw. Figure 4.1 shows the input data and created context. To create a context for a specifically prepared tooth, a 3D scan is semantically segmented into 14 classes (number of teeth in each jaw) using MeshSegNet method [75] trained on our dataset. Then two teeth in the prep's neighbor and the three

closest teeth in the opposing jaw were selected. We also add gingiva triangles close to the selected teeth. In the next step, we extract the margin line from a die (i.e., a scan of a prepared tooth). Based on the available crown cases for which we have the crown geometries provided by dental technicians using the commercial software of Exocad, we were able to extract the exact ground truth margin lines that were generated by the technicians. As opposed to other work published in the literature [89], we used the prepared tooth geometry and the ground truth margin line to enhance the fitting of the generated crown on the prepared tooth. By having more localized point density at the margin line we also teach the network to place a higher weight at the margin line. In production, we are using our pre-trained MeshSegNet to segment margin lines from prepared teeth [75, 90]. Figure 4.2 shows the ground truth shell with and without margin line.

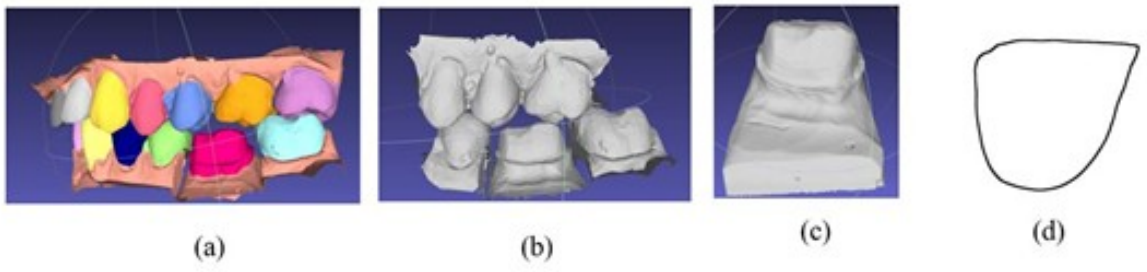


Figure 4.1 (a) Dental arch with a prep, (b) Generated context from segmentation, (c) die (prepped tooth) (d) Extracted margin line.

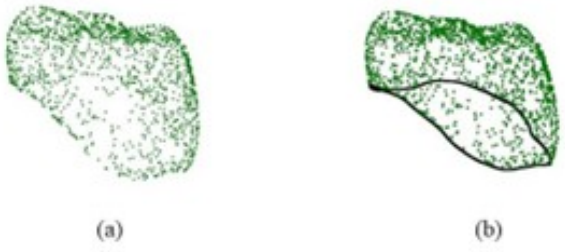


Figure 4.2 (a) Ground truth shell without margin line, (b) Ground truth shell with margin line.

4.4.2 Network Architecture

We use a transformer encoder-decoder architecture proposed by [9] to predict the point proxies for the dental crown. Figure 4.3illustrates our proposed framework for a point-based

dental crown generation. The transformer-based model by [9] tackles the point cloud completion problem as a set-to-set translation problem. The model is characterized by five main components: an encoder-decoder architecture to convert point cloud completion as a set-to-set translation problem, point proxies to represent the original point clouds as a set of feature vectors, geometry-aware transformer block to facilitate transformers to better leverage the inductive bias about 3D geometric structures, query generator to dynamically generate queries for the decoder, and multi-scale point cloud generation process to generate final prediction in a course to fine manner. One of the key challenges of applying transformers is that the self-attention mechanism in transformers lacks some inductive biases that explicitly model the structure of the point cloud. To enable transformers to better leverage the inductive bias about the 3D geometric structure of point clouds, [9] proposed a geometry-aware block that models the geometric relations which can be a plug-and-play module is designed to incorporate the attention blocks with any transformer architecture. The multi-scale point could generation is done by a small modification in the original FoldingNet [43] which takes predicted proxies, and centers from the transformer, and tries to map the 3D grid onto a proper tooth by mimicking the deformation of a 3D point cloud. We also used three consecutive folding operations, instead of two in the original FoldingNet paper, to help produce detailed structures.

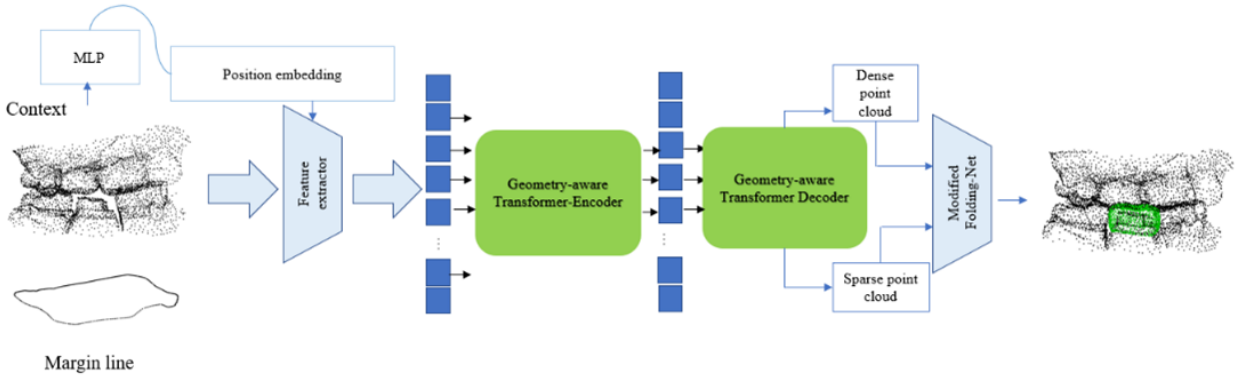


Figure 4.3 The pipeline of our proposed network.

We employed the mean chamfer distance as the evaluation metric, which can measure the distance between the prediction point cloud and ground truth at a set level. Fitting the crown with the neighboring teeth can be controlled by this metric. For each prediction, the chamfer distance between the prediction point set p and the ground truth point set g is calculated.

$$d_{\mathrm{CD}}(P,G) = \frac{1}{|P|}\sum_{p\in P}\min_{g\in G}\|p-g\|_2^2 + \frac{1}{|G|}\sum_{g\in G}\min_{p\in P}\|g-p\|_2^2 \tag{4.1}$$

where $P, G \in \mathbb{R}^{N\times 3}$ are point sets with N points, and $\|\cdot\|_2$ denotes the Euclidean (2-norm) distance.

4.5 Experimental Results

Our experimental dataset contains 90 training, 10 validation, and 25 test cases. The dataset included all type of teeth in the jaw, namely molars, canines, and incisors. We uniformly sampled 1568 points from the crown and 10240 points for the context and 1000 points for the margin line. Our training procedure is like the one used in [9]. The input point clouds were normalized by subtracting the mean of the points and dividing the outcome by the standard deviation. Regular augmentation methods such as rotation, translation, and scaling were applied to meshes. Figure 4.4 shows our results qualitatively. To evaluate our generation, we used the Chamfer distance (CD) between the prediction point set and the ground-truth point set. The mean CD of the different experiments is shown in Table 4.1.

Experiment	**Method**	**CD-L1**	**CD-L2**
All positions	Baseline + Margin line	56.39×10^{-3}	8.04×10^{-3}

Table 4.1 Chamfer Distance (L1 and L2) for the evaluated method.

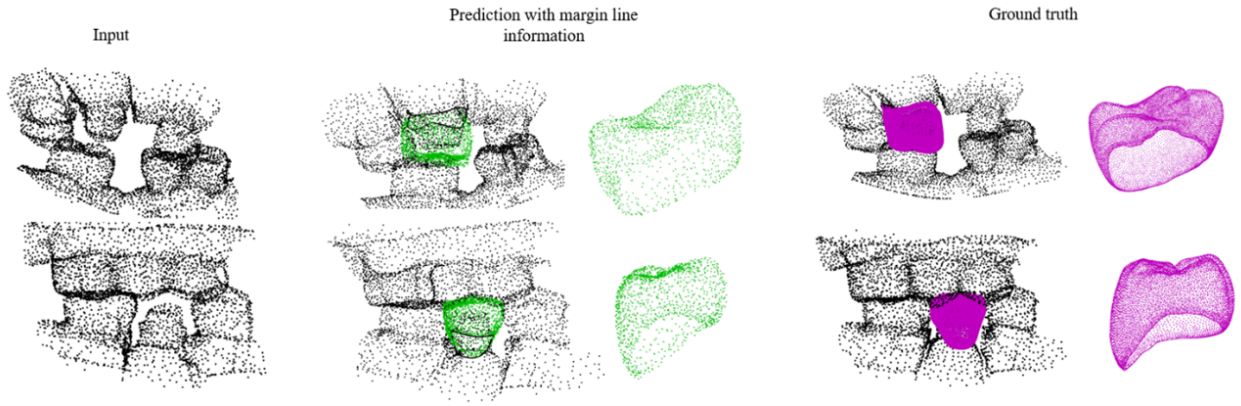
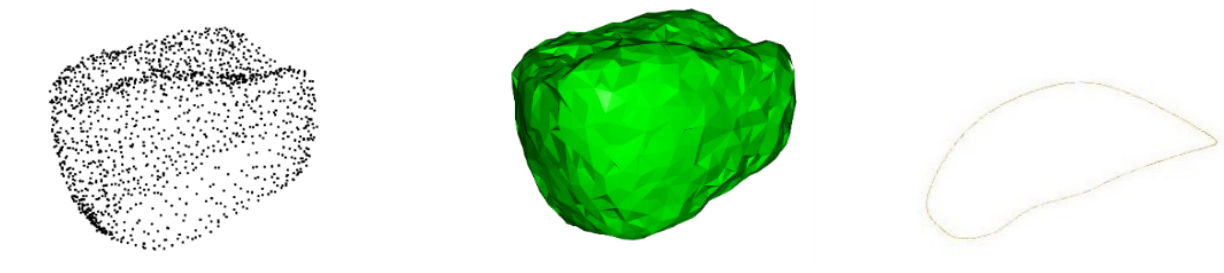


Figure 4.4 Prediction with margin line information.

In this stage, our goal is to evaluate the margin line that we can extract from the reconstructed surface of the predicted shell. We use a simple reconstruction method such as ball pivoting algorithm [91] (it can be substitute with other surface reconstruction methods such as Poisson surface reconstruction) to reconstruct the predicted shell. Then the margin line is extracted from the reconstructed surface . Figure 4.5 shows the predicted shell , reconstructed shell

and extracted marginline. In this experiment, we first predict the shell by the transformer-based network without considering margin line information. Second, we predict the shell by considering margin line information in the shell and context input. Finally, we evaluate the extracted margin line from reconstructed shell of the two experiments.



(a) (b) (c)

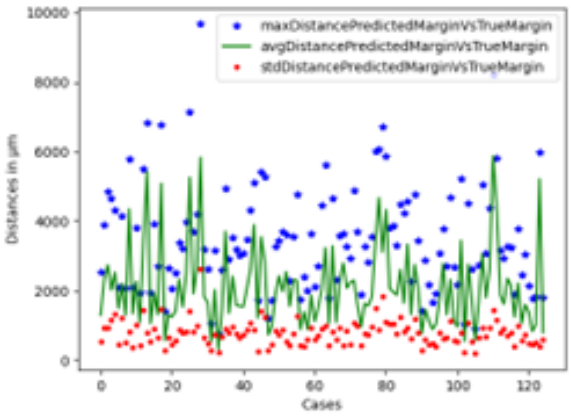
Figure 4.5 The point predicted shell (a), reconstructed surface (b), and extracted margin line (c).

The diagrams in figure 4.6 show the comparison of the distance between extracted margin line from the predicted shell and the ground truth margin line. We notice that when we consider the margin line information, there is much improvement in the predicted margin line in comparison to the ground truth margin line. There are some quantitative results in table 4.2.

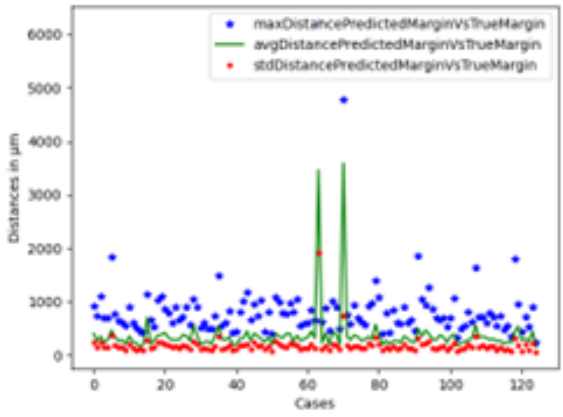
Metrics	**Without margin line information**	**With margin line information**
Max distance (μm)	3501.1311	847.5652
Min distance (μm)	695.0081	60.1825
Average distance (μm)	2105.3221	372.2782
Std distance (μm)	783.3138	184.4864

Table 4.2 The average distance between the predicted margin line and the ground truth margin line. Distances are in μm.

We evaluate the margin line extracted from the predicted shell with and without margin line information. Figure 4.7 shows one qualitative result in which one shell is predicted once by considering margin line information in the shell and context input, the green one, and the second without this information, the grey one. We can see the improvement in the scale of the predicted shell and the extracted margin line is so close to the ground truth margin line.



(a)



(b)

Figure 4.6 The diagram shows the (max, avg, std) distance of extracted margin line from the predicted shell and ground truth margin line. The left one is without considering margin line information in the input shell and context, the right one is by considering the margin line information.

4.6 Implimentation Details

We implement our networks on PyTorch. All modules in our completion network are trained with an ADAM optimizer, and the initial learning rate is set to 0.0005 with a decay rate of 0.9 every 20 epochs. We set our batch size to 16 and the number of workers to 8. We assign a radius in the range [0.6, 0.7, 0.8, 0.9, 1] for ball pivoting surface reconstruction.

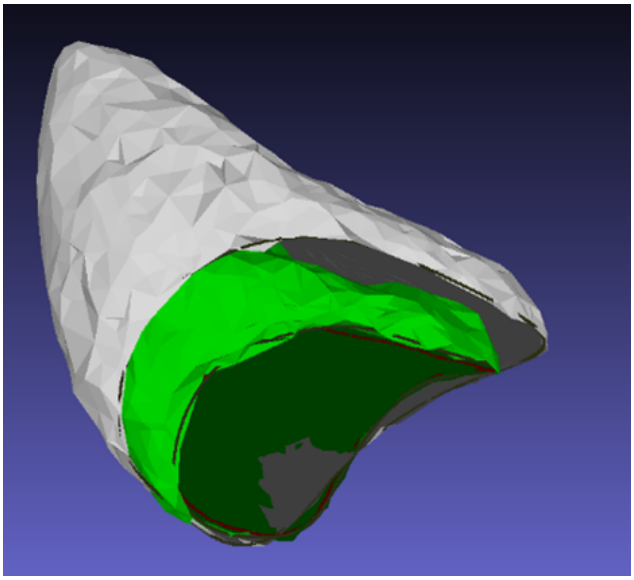


Figure 4.7 This figure shows the predicted crowns, the green one with margin line information and the grey one without margin line information. The black line of points is the prediction for the margin line and the red line point is the ground truth margin line.

4.7 Conclusions

Here, in contrast to recent works on dental restorations which focused only on restoring plausible missing teeth, we generated a crown for a realistic setting with the prepared tooth. We have shown that by adding margin line information obtained from the prep, our model can generate better-fit crowns. In the future, we are planning to generate the shell in a mesh format by using the graph convolutional network.

4.8 Acknowledgements

This work was funded by KerenOr, Intellident Dentaire Inc., iMD Research, the Natural Science and Engineering Research Council of Canada, Institut de valorisation des données (IVADO), and MEDTEQ. We thank Compute Canada for providing the computational resources used in this work. The authors acknowledge the help and support from JACOBB and Object Research Systems Inc. This work has not been submitted for publication or presentation elsewhere.

CHAPTER 5 ARTICLE 2: FROM MESH COMPLETION TO AI DESIGNED CROWN

Golriz Hosseinimanesh[a], Farnoosh Ghadiri[b], Julia Keren[c], Farida Cheriet[a], François Guibault[a]

[a]Polytechnique Montréal, Department of Computer and Software Engineering, P.O. Box 6079, Montréal, H3C 3A7, Québec, Canada

[b]Centre d'intelligence artificielle appliquée (JACOBB), Montréal, Québec, Canada

[c]Intellident Dentaire Inc., Montréal, Québec, Canada

Presentation: This work was published in the conference proceedings of Medical Image Computing and Computer Assisted Intervention Society (MICCAI), on May 25, 2023 [92].

5.1 CRediT authorship contribution statement

Golriz Hosseinimanesh: Methodology, Software, Validation, Investigation, Data Curation, Writing - Original Draft. **Farnoosh Ghadiri**: Writing - Review& Editing. **Julia Keren**: Conceptualization, Resources. **Farida Cheriet**: Supervision. Writing - Review & Editing. **François Guibault**: Supervision, Project administration, Funding acquisition, Resources, Writing - Review & Editing.

5.2 Abstract

Designing a dental crown is a time-consuming and labor-intensive process. Our goal is to simplify crown design and minimize the tediousness of making manual adjustments while still ensuring the highest level of accuracy and consistency. To this end, we present a new end-to-end deep learning approach, coined Dental Mesh Completion (DMC), to generate a crown mesh conditioned on a point cloud context. The dental context includes the tooth prepared to receive a crown and its surroundings, namely the two adjacent teeth and the three closest teeth in the opposing jaw. We formulate crown generation in terms of completing this point cloud context. A feature extractor first converts the input point cloud into a set of feature vectors that represent local regions in the point cloud. The set of feature vectors is then fed into a transformer to predict a new set of feature vectors for the missing region (crown). Subsequently, a point reconstruction head, followed by a multi-layer perceptron, is used to predict a dense set of points with normals. Finally, a differentiable point-to-mesh layer serves to reconstruct the crown surface mesh. We compare our DMC method to a graph-

based convolutional neural network which learns to deform a crown mesh from a generic crown shape to the target geometry. Extensive experiments on our dataset demonstrate the effectiveness of our method, which attains an average of 0.062 Chamfer Distance. The code is available at: https://github.com/Golriz-code/DMC.git

Keywords: Mesh completion, Transformer, 3D shape generation.

5.3 Introduction

If a tooth is missing, decayed, or fractured, its treatment may require a dental crown. Each crown must be customized to the individual patient in a process, as depicted in Figure 5.1. The manual design of these crowns is a time-consuming and labor-intensive task, even with the aid of computer-assisted design software. Designing natural grooves and ensuring proper contact points with the opposing jaw present significant challenges, often requiring technicians to rely on trial and error. As such, an automated approach capable of accelerating this process and generating crowns with comparable morphology and quality to that of a human expert would be a groundbreaking advancement for the dental industry. A limited number of studies

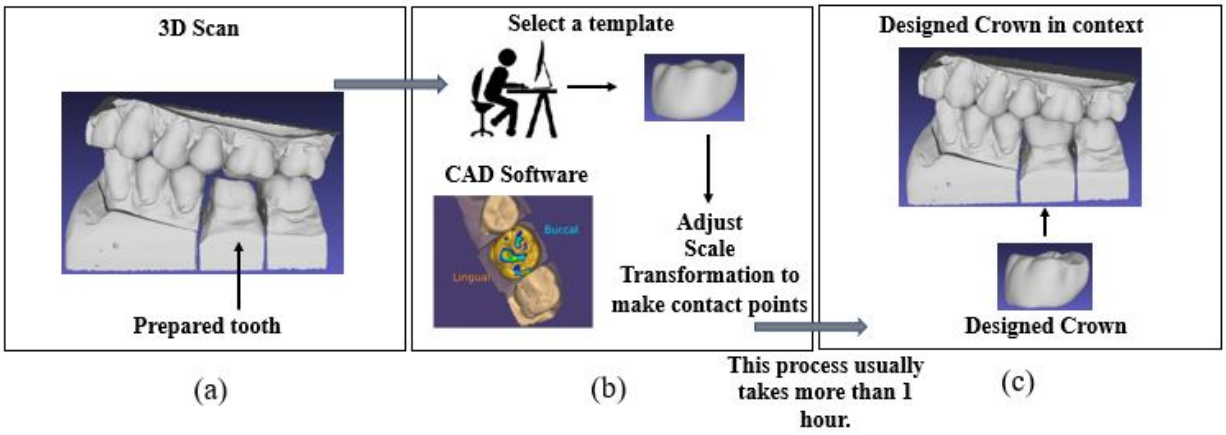


Figure 5.1 Dental crown design process: a) Dentist prepares the tooth; b) Technician designs the crown; c) Dentist places the crown on the prepared tooth.

have focused on how to automate dental crown designs. In [1, 81], a conditional generative adversarial based network (GAN) is applied to a 2D depth image obtained from a 3D scan to generate a crown for a prepared tooth. Depth images created from a 3D scan can be used directly with 2D convolutional neural networks (CNNs) such as pix2pix [87], which are well-established in computer vision. However, depth images are limited in their ability to capture fine-grained details and can suffer from noise and occlusion issues. By contrast, point clouds have the advantage of being able to represent arbitrary 3D shapes and can capture

fine-grained details such as surface textures and curvatures. [86, 89] use point cloud-based networks to generate crowns in the form of 3D point clouds. Input point clouds used by [57] are generated by randomly removing a tooth from a given jaw; then the network estimates the missing tooth by utilizing a feature points-based multi-scale generating network. [86] propose a more realistic setting by generating a crown for a prepared tooth instead of a missing tooth. They also incorporate margin line information extracted from the prepared tooth in their network to have a more accurate prediction in the margin line area. The crowns generated by both approaches are represented as point clouds, so another procedure must convert these point clouds into meshes. Creating a high-quality mesh that accurately represents the underlying point cloud data is a challenging task which is not addressed by these two works. [84] proposed a transformer-based network to generate a surface mesh of the crown for a missing tooth. They use two separate networks, one responsible for generating a point cloud and the other for reconstructing a mesh given the crown generated by the first network. Similar to [57, 86], the point completion network used by [84] only uses the Chamfer Distance (CD) loss to learn crown features. This metric's ability to capture shape details in point clouds is limited by the complexity and density of the data. Although all aforementioned methods are potentially applicable to the task of dental crown design, most of them fail to generate noise-free point clouds, which is critical for surface reconstruction. One way to alleviate this problem is to directly generate a crown mesh. In [68], the authors develop a deep learning-based network that directly generates personalized cardiac meshes from sparse contours by iteratively deforming a template mesh, mimicking the traditional 3D shape reconstruction method. To our knowledge, however, the approach in [68] has not been applied to 3D dental scans.

In this paper, we introduce Dental Mesh Completion (DMC), a novel end-to-end network for directly generating dental crowns without using generic templates. The network employs a transformer-based architecture with self-attention to predict features from a 3D scan of dental preparation and surrounding teeth. These features deform a 2D fixed grid into a 3D point cloud, and normals are computed using a simple MLP. A differentiable point-to-mesh module reconstructs the 3D surface. The process is supervised using an indicator grid function and Chamfer loss from the target crown mesh and point cloud. Extensive experiments validate the effectiveness of our approach, showing superior performance compared to existing methods as measured by the CD metric. In summary, our main contributions include proposing the first end-to-end network capable of generating crown meshes for all tooth positions, employing a non-template-based method for mesh deformation (unlike previous works), and showcasing the advantages of using a differentiable point-to-mesh component to achieve high-quality surface meshes.

5.4 Related work

In the field of 3D computer vision, completing missing regions of point clouds or meshes is a crucial task for many applications. Various methods have been proposed to tackle this problem. Since the introduction of PointNet [41, 42], numerous methods have been developed for point cloud completion [6]. The recent works PoinTr [9] and SnowflakeNet [60] leverage a transformer-based architecture with geometry-aware blocks to generate point clouds. It is hypothesized that using transformers preserves detailed information for point cloud completion. Nonetheless, the predicted point clouds lack connections between points, which complicates the creation of a smooth surface for mesh reconstruction.

Mesh completion methods are usually useful when there are small missing regions or large occlusions in the original mesh data. Common approaches based on geometric priors, self-similarity, or patch encoding can be used to fill small missing regions, as demonstrated in previous studies [30, 93], but are not suitable for large occlusions. [94] propose a model-based approach that can capture the variability of a particular shape category and enable the completion of large missing regions. However, the resulting meshes cannot achieve the necessary precision required by applications such as dental crown generation. Having a mesh prior template can also be a solution to generate a complete mesh given a sparse point cloud or a mesh with missing parts. In [68], cardiac meshes are reconstructed from sparse point clouds using several mesh deformation blocks. Their network can directly generate 3D meshes by deforming a template mesh under the guidance of learned features.

We combine the advantages of point cloud completion techniques with a differentiable surface reconstruction method to generate a dental mesh. Moreover, we used the approach in [68] to directly produce meshes from 3D dental points and compared those results with our proposed method.

5.5 Methods

5.5.1 Network Architecture

Our method is an end-to-end supervised framework to generate a crown mesh conditioned on a point cloud context. The overview of our network is illustrated in Figure 7.2. The network is characterized by two main components: a transformer encoder-decoder architecture and a mesh completion layer. The following sections explain each part of the network.

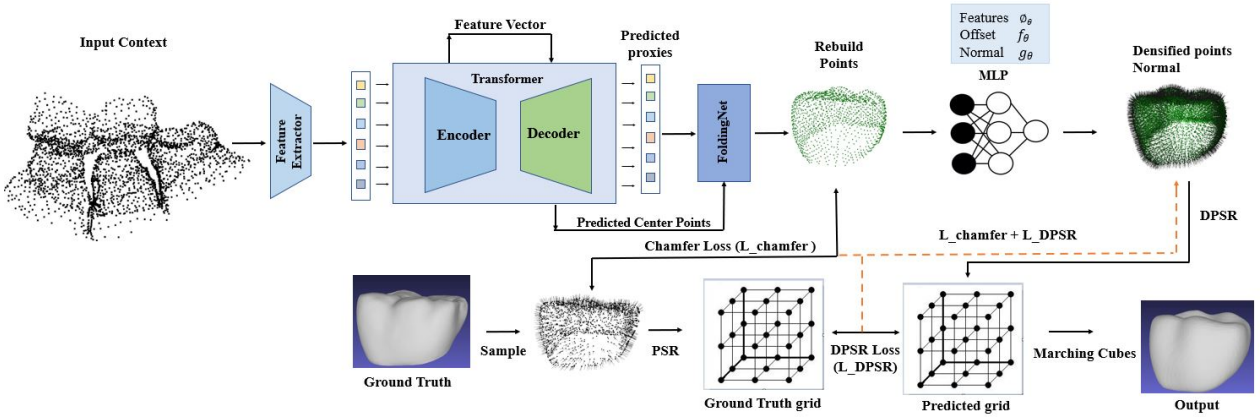


Figure 5.2 Pipeline of our proposed network.

Transformer encoder-decoder

We adapt the transformer encoder-decoder architecture from [9] to extract global and local 3D features from our input (context) using the encoder and generate crown points via the decoder. A dynamic graph convolution network (DGCNN) [50] is used to group the input points into a smaller set of feature vectors that represent local regions in the context. The generated feature vectors are then fed into the encoder with a geometry-aware block. This block is used to model the local geometric relationships explicitly. The self-attention layer in the encoder updates the feature vectors using both long-range and short-range information. The feature vectors are further updated by a multi-layer perceptron. The decoder's role is to reason about the crown based on the learnable pairwise interactions between features of the input context and the encoder output. The decoder incorporates a series of transformer layers with a self-attention and cross-attention mechanisms to learn structural knowledge. The output of the transformer decoder is fed into a folding-based decoder [43] to deform a canonical 2D grid onto the underlying 3D surface of the crown points.

Mesh completion layer

In this stage, to directly reconstruct the mesh from the crown points, we use a differentiable Poisson surface reconstruction (DPSR) method introduced by [10]. We reconstruct a 3D surface as the zero level set of an indicator function. The latter consists in a regular 3D point grid associated with values indicating whether a point is inside the underlying shape or not. To compute this function, We first densify the input unoriented crown points. This is done by predicting additional points and normals for each input point by means of an MLP network. After upsampling the point cloud and predicting normals, the network solves

a Poisson partial differential equation (PDE) to recover the indicator function from the densified oriented point cloud. We represent the indicator function as a discrete Fourier basis on a dense grid (of resolution 128^3) and solve the Poisson equation (PDE) with the spectral solver method in [10].

During training, we obtain the estimated indicator grid from the predicted point cloud by using the differentiable Poisson solver. We similarly acquire the ground truth indicator grid on a dense point cloud sampled from the ground truth mesh, together with the corresponding normals. The entire pipeline is differentiable, which enables the updating of various elements such as point offsets, oriented normals, and network parameters during the training process. At inference time, we leverage our trained model to predict normals and offsets using Differentiable Poisson Surface Reconstruction (DPSR) [10], solve for the indicator grid, and finally apply the Marching Cubes algorithm [95] to extract the final mesh.

Loss function

We use the mean Chamfer Distance (CD) [44] to constrain point locations. The CD measures the mean squared distance between two point clouds S_1 and S_2. Individual distances are measured between each point and its closest point in the other point set, as described in equation (5.1). In addition, we minimize the L_2 distance between the predicted indicator function x and a ground truth indicator function x_{gt}, each obtained by solving a Poisson PDE [10] on a dense set of points and normals. We can express the Mean Square Error (MSE) loss as equation (6.8), where $f_\theta(X)$ represents a neural network (MLP) with parameters θ conditioned on the input point cloud X, D is the training data distribution, along with indicator functions x_i and point samples X_i on the surface of shapes. The sum of the CD and MSE losses is used to train the overall network.

$$\mathrm{CD}(S_1, S_2) = \frac{1}{|S_1|} \sum_{x \in S_1} \min_{y \in S_2} |x - y|^2 + \frac{1}{|S_2|} \sum_{y \in S_2} \min_{x \in S_1} |y - x|^2 \tag{5.1}$$

$$L_{DPSR}(\theta) = E_{X_i, x_i \sim D} ||Poisson(f_\theta(X_i)) - x_i||_2^2 \tag{5.2}$$

5.6 Experimental results

5.6.1 Dataset and preprocessing

Our dataset consisted of 388 training, 97 validation, and 71 test cases, which included teeth in various positions in the jaw such as molars, canines, and incisors.

The first step in the preprocessing was to generate a context from a given 3D scan. To determine the context for a specific prepared tooth, we employed a pre-trained semantic segmentation model [75] to separate the 3D scan into 14 classes representing the tooth positions in each jaw. From the segmentations, we extracted the two adjacent and three opposing teeth of a given prepared tooth, as well as the surrounding gum tissue. To enhance the training data, we conducted data augmentation on the entire dental context, which included the master arch, opposing arch, and shell, treated as a single entity. Data augmentation involved applying 3D translation, scaling, and rotation, thereby increasing the training set by a factor of 10. To enrich our training set, we randomly sampled 10,240 cells from each context to form the input during training. We provide two types of ground truth: mesh and point cloud crowns. To supervise network training using the ground truth meshes, we calculate the gradient from a loss on an intermediate indicator grid. We use the spectral method from [10] to compute the indicator grid for our ground truth mesh.

5.6.2 Implementation details

We adapted the architecture of [9] for our transformer encoder-decoder module. For mesh reconstruction, we used Differentiable Poisson Surface Reconstruction (DPSR) from [10]. All models were implemented in PyTorch with the AdamW optimizer [96], using a learning rate of 5e-4 and a batch size of 16. Training the model took 400 epochs and 22 hours on an NVIDIA A100 GPU.

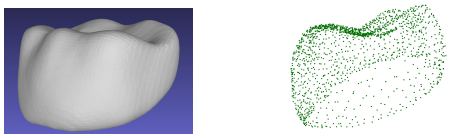
5.6.3 Evaluation and metrics

To evaluate the performance of our network and compare it with point cloud-based approaches, we used the Chamfer distance to measure the dissimilarity between the predicted and ground truth points. We employed two versions of CD: CD_{L_1} uses the L_1-norm, while CD_{L_2} uses the L_2-norm to calculate the distance between two sets of points. Additionally, we used the F-score [97] with a distance threshold of 0.3, chosen based on the distance between the predicted and ground truth point clouds. We also used the Mean Square Error (MSE) loss [10] to calculate the similarity between the predicted and ground truth indicator grids or meshes.

We conducted experiments to compare our approach with two distinct approaches from the literature, as shown in Table 6.1. The first such approach, PoinTr+margin line [86], uses the PoinTr [9] point completion method as a baseline and introduces margin line information to their network. To compare our work to [86], we used its official implementation provided by the author. In the second experiment, PoinTr+graph, we modified the work of [68] to

generate a dental crown mesh. To this end, we use deformation blocks in [68] to deform a generic template mesh to output a crown mesh under the guidance of the learned features from PoinTr. The deformation module included three Graph Convolutional Networks (GCNs) as in [66].

All experiments used the same dataset, which included all tooth positions, and were trained using the same methodology. To compare the results of the different experiments, we extracted points from the predicted meshes of our proposed network (DMC), as illustrated in Figure 5.3. Table 6.1 shows that DMC outperforms the two other networks in terms of both CD and F-score. PoinTr+graph achieves poor CD and F-score results compared to the other methods. While the idea of using graph convolutions seems interesting, features extracted from the point cloud completion network don't carry enough information to deform the template into an adequate final crown. Therefore, these methods are highly biased toward the template shape and need extensive pre-processing steps to scale and localize the template. In the initial two experiments, the MSE metric was not applicable as it was calculated on the output meshes. Figure 5.4 showcases the visual results obtained from our proposed network (DMC). Furthermore, Figure 5.5 presents a visual comparison of mesh surfaces generated by various methods for a sample molar tooth.

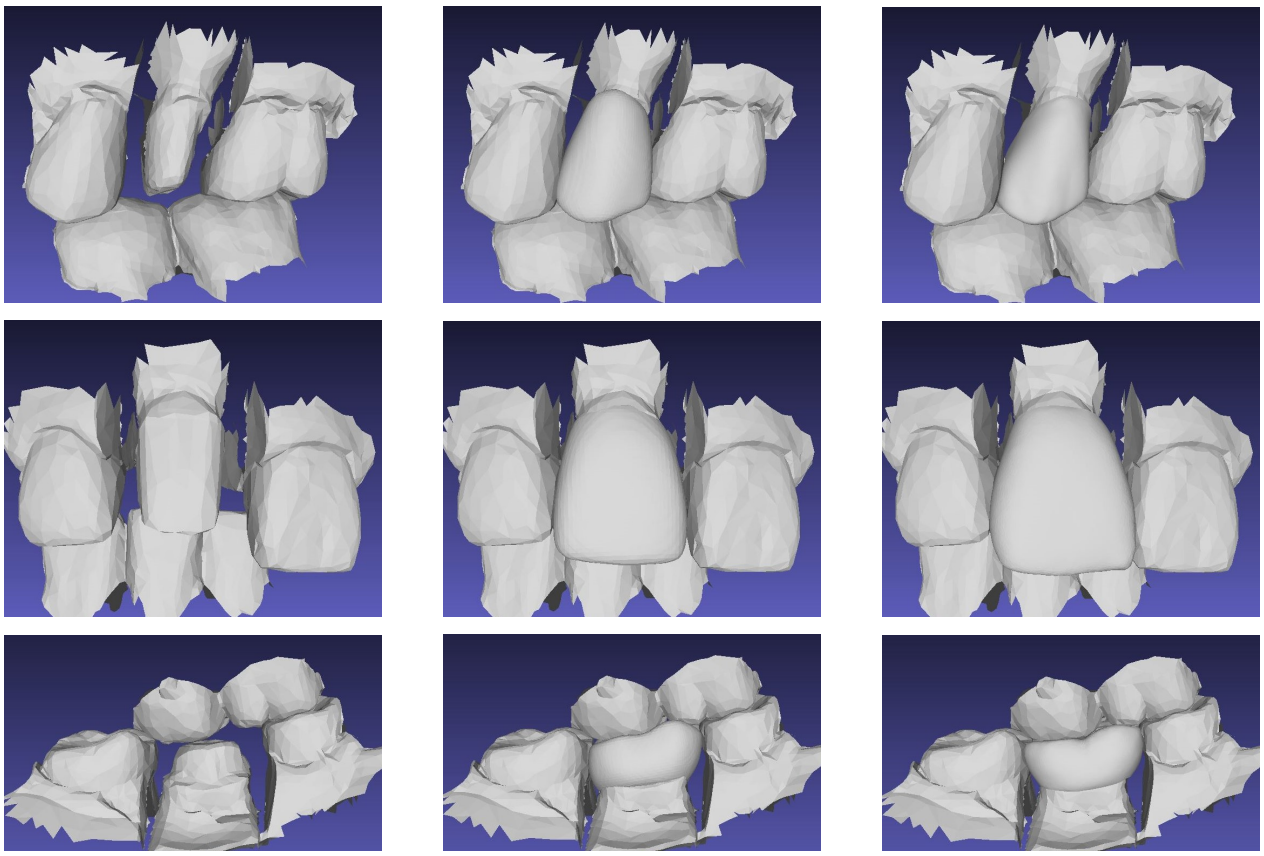


(a) (b)

Figure 5.3 Crown mesh predicted by DMC (a) and extracted point set (b).

Table 5.1 Comparison of proposed method (DMC) with two alternate methods. Evaluation metrics: CD_{L_1}, CD_{L_2} ; MSE on output meshes; F-Score@0.3 .

Method	CD-L1 (↓)	CD-L2 (↓)	MSE (↓)	$F1^{0.3}$ (↑)
PoinTr + margin line [86]	0.065	0.018		0.54
PoinTr + graph	1.99	1.51		0.08
DMC	**0.0623**	**0.011**	0.0028	**0.70**

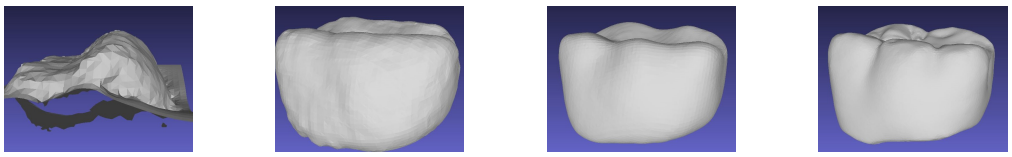


(a) (b) (c)

Figure 5.4 Examples of mesh completions by the proposed architecture (DMC). (a) Input context containing master arch, prepped tooth, and opposing arch; (b) Generated mesh in its context; (c) Ground truth mesh in its context.

5.6.4 Ablation study

We conducted an ablation study to evaluate the components of our architecture. We started with the baseline PoinTr [86], a point completion method. To enhance it, we integrated Shape as Points (SAP) [10] as a separate network for mesh reconstruction from the PoinTr-generated point cloud. Next, we tested our proposed method (DMC) by excluding the Mean Square Error (MSE) loss function. Finally, we assessed DMC's performance, including the MSE loss function. The results, shown in Table 5.2, demonstrate the consistent improvements achieved by our full model across all evaluation metrics.



(a) (b) (c) (d)

Figure 5.5 Qualitative comparison of crown mesh generation approaches: a) Standard Poisson surface reconstruction; b) PoinTr [9] for point cloud and Shape as points [10] for mesh; c) Proposed method (DMC); d) Ground truth shape.

Table 5.2 Results of ablation study. Metrics are the same as in Table 6.1.

Method	CD-L1 (↓)	CD-L2 (↓)	MSE (↓)	$F1^{0.3}$ (↑)
PoinTr [9]	0.070	0.023		0.24
PoinTr + SAP	0.067	0.021	0.031	0.50
DMC without MSE	0.0641	0.015		0.65
DMC (Full Model)	**0.0623**	**0.011**	**0.0028**	**0.70**

5.7 Conclusion

Existing deep learning-based dental crown design solutions require additional steps to reconstruct a surface mesh from the generated point cloud. In this study, we propose a new end-to-end approach that directly generates high-quality crown meshes for all tooth positions. By utilizing transformers and a differentiable Poisson surface reconstruction solver, we effectively reason about the crown points and convert them into mesh surfaces using Marching Cubes. Our experimental results demonstrate that our approach produces accurately fitting crown meshes with superior performance. In the future, incorporating statistical features into our deep learning method for chewing functionality, such as surface contacts with adjacent and opposing teeth, could be an interesting avenue to explore.

5.8 Acknowledgments

This work was funded by Kerenor Dental Studio, Intellident Dentaire Inc..

CHAPTER 6 ARTICLE 3: PERSONALIZED DENTAL CROWN DESIGN: A POINT-TO-MESH COMPLETION NETWORK

Golriz Hosseinimanesh[a], Ammar Alsheghri[b],Julia Keren[c], Farida Cheriet[a], François Guibault[a]

[a]Polytechnique Montréal, Department of Computer and Software Engineering, P.O. Box 6079, Montréal, H3C 3A7, Québec, Canada

[b]Interdisciplinary research center for Biosystems and Machines, King Fahd University of Petroleum and Minerals (KFUPM), Dhahran, KSA

[c]Intellident Dentaire Inc., Montréal, Québec, Canada

Presentation: This work was published in Medical Image Analysis journal, on December 10, 2024 [15].

6.1 CRediT authorship contribution statement

Golriz Hosseinimanesh: Methodology, Software, Validation, Investigation, Data Curation, Writing - Original Draft. **Ammar Alsheghri**: Writing - Review& Editing. **Julia Keren**: Conceptualization, Resources. **Farida Cheriet**: Supervision. Writing - Review & Editing. **François Guibault**: Supervision, Project administration, Funding acquisition, Resources, Writing - Review & Editing.

6.2 Abstract

Designing dental crowns with computer-aided design software in dental laboratories is complex and time-consuming. Using real clinical datasets, we developed an end-to-end deep learning model that automatically generates personalized dental crown meshes. The input context includes the prepared tooth, its adjacent teeth, and the two closest teeth in the opposing jaw. The training set contains this context, the ground truth crown, and the extracted margin line. Our model consists of two components: First, a feature extractor converts the input point cloud into a set of local feature vectors, which are then fed into a transformer-based model to predict the geometric features of the crown. Second, a point-to-mesh module generates a dense array of points with normal vectors, and a differentiable Poisson surface reconstruction method produces an accurate crown mesh. Training is conducted with three losses: (1) a customized margin line loss; (2) a contrastive-based Chamfer distance loss; and (3) a mean square error (MSE) loss to control mesh quality. We compare our method with our

previously published method, Dental Mesh Completion (DMC). Extensive testing confirms our method's superiority, achieving a 12.32% reduction in Chamfer distance and a 46.43% reduction in MSE compared to DMC. Margin line loss improves Chamfer distance by 5.59%. Keyword: Dental crown generation , Transformer , Mesh completion , Margin line

6.3 Introduction

Dental laboratories are responsible for designing hundreds of dental crowns annually tailored to each patient's unique tooth morphology and specific characteristics using computer-aided design (CAD) software. The process begins with a dentist preparing the damaged tooth to set a stable foundation for the crown. A 3D digital model of the prepared tooth and its neighboring teeth is then captured using an intraoral scanner, providing essential context for custom crown design. A dental technician employs a CAD system, equipped with specialized tools for dental applications, to meticulously design the crown. This step involves selecting a standard tooth template from a digital library and modifying it through careful scaling and positioning to meet functional and aesthetic requirements. The entire manual design process, including critical adjustments for contact points and the margin line, is depicted in the flowchart (Fig. 6.1). These adjustments ensure that the crown aligns perfectly with the dental preparation and is securely sealed at the prepped's margin line. Further advanced transformations are applied to ensure proper contact with adjacent teeth. The flowchart (Fig. 6.1) illustrates the detailed adjustments and customization involved in manually designing hundreds of crowns each year to meet individual patient needs.

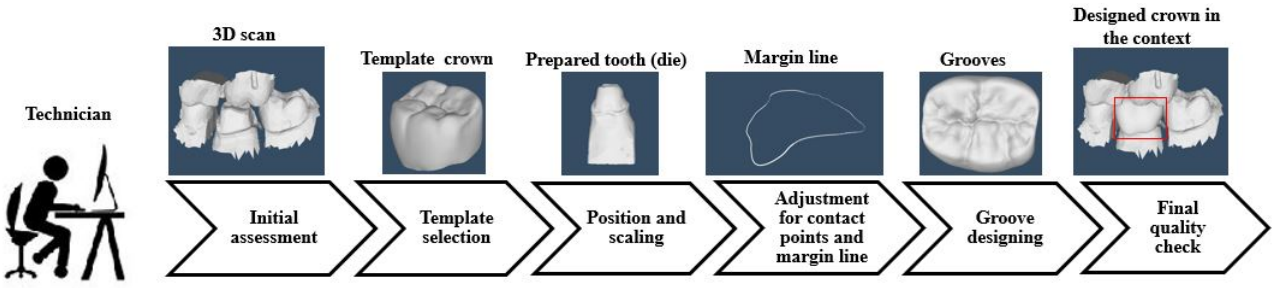


Figure 6.1 Flowchart of manually designing a dental crown by a technician. This process usually takes more than 1 hour.

Although CAD systems provide high precision, the manual design process is time-consuming and prone to human error. Each crown can take over an hour to design, and achieving consistency across designs is challenging. Therefore, an automated AI-based solution is needed to improve the accuracy, efficiency, and consistency of crown designs.

To address these challenges, emerging artificial intelligence (AI) and deep learning technologies are offering promising solutions. Recent studies show that AI has the potential to enhance diagnostic precision through techniques like fuzzy logic and convolutional networks ([98]; [99]; [100]). Significant advancements have also been made in image segmentation, employing dual attention modules, boundary-aware transformers, GAN-based models, and domain adaptation methods to improve segmentation quality and boundary detection across medical applications ([101]; [102]; [103]; [104]; [105]; [106]). In the field of 3D dental reconstruction, specific studies have advanced synthetic image reconstruction and prosthetic restoration, focusing on detailed crown morphology, gingival margin line generation, and occlusal surface reconstruction through innovative GAN and transformer-based models ([107]; [108]; [82]; [109]; [110]). Collectively, these innovations in AI address critical needs in automation, accuracy, and efficiency, although limitations in computational demands, data quality, and generalizability persist.

Building on these advancements, recent efforts in dental crown design automation have focused on simplifying 3D dental scans by transforming them into 2D depth images ([111,112]). The pix2pix technique ([113]) was used to generate dental crown images. While this approach allows for the visualization of the crown's occlusal surface from a single view, it inherently neglects the lateral portions that are crucial for a well-fitted crown. The lateral surfaces play a key role in ensuring the crown fits securely around the tooth and aligns properly with adjacent teeth. Additionally, the lateral margins are critical for maintaining proper bite alignment (occlusion) and for preventing issues like plaque buildup or gum irritation. Failure to capture these aspects can compromise the crown's stability, its aesthetic integration with neighboring teeth, and its overall longevity. The transition from 3D to 2D causes a significant loss of spatial information, particularly around the tooth margin line. Moreover, these 2D methods do not generate complete tooth surfaces directly, making the acquisition of detailed lateral surfaces from the occlusal view a non-trivial task.

In response to these limitations, point clouds have emerged as a superior alternative. Unlike 2D imaging, point clouds can faithfully represent any 3D shape; capturing the overall form and the intricate details. Within this context, ([114]) introduced a point cloud-based method that begins by randomly removing a tooth from a given jaw. Subsequently, the network estimates the missing tooth using a multi-scale network that generates feature points. However, this approach does not simulate the more realistic scenario, which is generating a crown for a prepared tooth. As such, inaccuracies occur in the positioning of their predictions. Additionally, because the method generates point clouds rather than meshes, it offers limited control over the smoothness of the resulting shapes. To this end, a subsequent study ([115]) developed a transformer-based network aimed at creating a surface mesh of the crown for a

missing tooth. Their approach involved two distinct networks: one that generates a point cloud, and a second that reconstructs a mesh based on the crown points produced by the first network. However, the method relied solely on point distance loss measures, such as Chamfer Distance (CD), to refine the point clouds. It lacked a specific loss function to manage the mesh reconstruction process, which significantly limited its ability to precisely control the quality and accuracy of the final mesh output. In a previous effort, our team published a pioneering study on the DMC model ([11]), proposing an end-to-end deep learning model designed to directly produce dental crown meshes from input point clouds. This model utilizes a transformer-based architecture with self-attention mechanisms to predict features from a 3D scan of dental preparation and surrounding teeth. These features are then employed to deform a 2D fixed grid into a 3D point cloud. Subsequently, a differentiable point-to-mesh module predicts the indicator grid, which facilitates the generation of the mesh. The training of this model incorporates Chamfer distance loss for the point cloud and mean square error (MSE) for the grid, aiming to enhance the learning of crown features. Despite these advancements, increasing the resolution of the output point cloud constitutes a challenge for the transformer to generate noise-free results. Although introducing the MSE loss alongside the Chamfer distance loss improves point cloud distribution, the method still struggles to produce high-resolution point clouds. This limitation hinders its potential to predict high-quality meshes. Moreover, it does not effectively solve the critical issue of accurately fitting the crown at the margin line position of the prepared tooth. Another study on generating dental crowns for prepared teeth improves accuracy at the critical margin line by integrating margin line data from the prepared tooth ([86]). This approach significantly enhances precision in the margin line region. However, their crowns, represented as point clouds, required an additional procedure for conversion into meshes. To address these issues, we introduce a new end-to-end network designed to directly generate high-quality dental crowns for all tooth positions. Our model utilizes a transformer encoder-decoder architecture, enriched with a self-attention mechanism, adaptive query generation, and a denoising task strategy, to process points from 3D scans of dental preparations and adjacent teeth. This approach enables our network to dynamically predict a set of features that accurately encapsulate the dental crown's geometry. These features are then passed through a fully connected layer to reconstruct the crown's point cloud, with the transformer's advanced mechanisms ensuring the prediction of high-resolution and noise-free point clouds. Normals for these points are derived using a straightforward multi-layer perceptron (MLP). A differentiable point-to-mesh module integrated into our network facilitates 3D surface reconstruction. The entire process is supervised by a contrastive learning chamfer loss (InfoCD) ([116]), optimizing point distributions between the prediction and the ground truth while capturing surface similarities

through mutual information estimation. Additionally, a margin line loss ensures the accurate alignment of predicted points around the margin line, crucial for the crown's proper fit on the prepared tooth. An indicator grid function, computed from the target crown mesh, further guides the reconstruction accuracy. Extensive experimentation validates our methodology's effectiveness and its innovative contributions, showcasing advancements in dental crown generation. To sum up, the contributions of this paper are as follows:

1. Proposed an automated end-to-end network for personalized dental crown generation across all tooth positions.

2. Developed a fully differentiable point-to-mesh architecture that directly generates dental meshes from 3D point clouds through a transformer-based completion mechanism.

3. Developed custom loss functions aimed at improved crown alignment and accuracy, including a margin line loss for precise positioning and a state-of-the-art contrastive learning Chamfer loss (InfoCD) for refined surface matching. These innovations enhance the overall reliability and accuracy of the crown generation process.

6.4 Related work

This section reviews the essential domains of AI approaches for dental crown design, point completion, mesh completion and reconstruction, and the metrics used for their evaluation; laying the groundwork for our proposed advancements in dental crown design automation.

6.4.1 AI approaches for dental crown design

AI-driven approaches for dental crown design represent a rapidly growing area of research, directly informing the advancements proposed in this study. [107] introduced a dual discriminator adversarial learning approach for occlusal surface reconstruction, integrating a dilated convolution-based generative model with global-local discriminators to automate the design process while preserving the natural morphology of crowns. This method takes 2D depth images as input and outputs 3D reconstructed dental surfaces. While effective, it focuses on specific tooth types and requires large datasets, limiting its broader application. Building on this, ([108]) employed a true 3D deep learning model (3D-DCGAN), using 3D scans of teeth as input and generating accurate, biomechanically sound dental crowns as output. However, the high computational requirements and limited validation across diverse datasets remain significant challenges. The DAIS framework, proposed by [106], takes a different approach by incorporating depth map generation and a deep-learning-based restoration network, using

depth maps and 3D dental scans as input and producing anatomically accurate inlay prostheses. This method effectively addresses cases with large, irregularly shaped missing teeth, yet its dependence on large datasets and difficulty in handling diverse tooth shapes present obstacles to its generalizability. Further advancing the field, the DCPR-GAN method by [82] utilizes a two-stage GAN framework specifically for dental crown prosthesis restoration, using 3D scans and depth maps as input and reconstructing dental crowns as output. While this approach offers improvements in occlusal surface reconstruction, it faces limitations due to its narrow focus on specific teeth and potential gaps in depth information. Additionally, ([109]) demonstrated the effectiveness of deep adversarial networks in reconstructing gingival margin lines, particularly in cases with multiple missing molars. This method processes 3D scans of gingival contours as input and outputs reconstructed gingival margin lines. Despite its accuracy, the reliance on depth images and the need for post-processing steps impact both efficiency and generalizability. Collectively, these AI-driven methods demonstrate significant progress in dental crown design but also present challenges, including data requirements, computational complexity, and generalizability across diverse cases. Addressing these issues remains critical for the practical application of AI in dental crown design.

Building upon these innovations, diffusion models ([117]) have emerged as a promising approach in medical image synthesis, with potential applications in dental crown design. Their strengths include generating high-resolution synthetic images that facilitate model training and reduce dependence on extensive datasets. The iterative noise-removal mechanism enables detailed reconstructions, maintaining the natural morphology of teeth. Additionally, these models are highly adaptable, allowing customization to specific dental imaging characteristics, such as intricate textures and precise occlusal alignment. However, their significant computational demands and the need for substantial processing power present challenges for clinical adoption. Enhancing computational efficiency and optimizing algorithms are essential for integrating these models effectively into dental workflows.

6.4.2 Point completion

The task of reconstructing incomplete parts of point clouds or meshes is essential in 3D computer vision, supporting a wide range of applications. This field has seen significant innovation, particularly since the foundational work of PointNet and FoldingNet [41, 43, 118]. According to [119], point cloud completion strategies are categorized into point-based, convolution-based, graph-based, GAN-based, diffusion-based, and transformer-based methods.

Point-based methods directly manipulate individual points but may overlook complex spa-

tial relationships. Convolution-based approaches provide strong feature extraction but risk losing detail during voxelization. Graph-based strategies effectively capture intricate point connections, while GANs excel at generating realistic point clouds despite stability challenges. Transformer-based models bring a novel perspective by leveraging self-attention mechanisms to capture global dependencies and detailed structures within point clouds [119]. Diffusion models, meanwhile, methodically convert noise into structured point clouds, enabling the creation of complex geometries in a controlled manner [120].

Focusing on techniques that address missing regions, [57] utilized a GAN-based approach with a multi-resolution encoder and point pyramid decoder for reconstruction, though this method has limitations in capturing finer details. Transformer-based models like PoinTr [9], SnowflakeNet [60], and AdapoinTr [12] have pushed the boundaries of point cloud generation by employing geometry-aware blocks and adaptive query generation, alongside denoising tasks. These models excel in preserving detailed features but face challenges in achieving point connectivity, which is vital for seamless mesh reconstruction.

The Few-point Shape Completion (FSC) model by [121] addresses the challenge of completing point clouds using as few as 64 input points. The model employs a dual-branch feature extractor to maximize sparse input utilization and dynamically prioritize key points, followed by a two-stage revision network that refines the output to preserve both local details and global structure. The authors report that experiments on the ShapeNet dataset revealed that with only 64 input points, the FSC model retained nearly 60% of the original shape information and achieved an average Chamfer Distance (CD-L_1) of 7.89. This performance surpassed prior methods, such as GRNet (CD-L_1 of 17.61) and PCN (CD-L_1 of 12.11). Despite its strong results in sparse data scenarios, the FSC's architecture—featuring dual-branch extraction with stacked layers and two-stage adversarial refinement—adds substantial computational complexity. This design increases the operational load relative to simpler architectures, posing challenges due to higher memory and processing demands for point feature extraction and refinement stages.

6.4.3 Mesh completion and reconstruction

Mesh completion in 3D computer vision, which is crucial for various applications, involves reconstructing missing parts of meshes, regardless of whether the gaps are small or the occlusions are large. Traditional methods utilizing geometric priors, self-similarity, or patch encoding have proven effective for filling small gaps ([93, 122]) yet they struggle when faced with larger occlusions. Innovative model-based approaches have been developed to tackle these significant gaps by capturing the variability within specific shape categories, offering a

solution for extensive reconstructions ([94]). However, these methods often lack the precision required for intricate tasks such as dental crown design. A promising direction is the use of mesh prior templates, which, guided by learned features, deform to reconstruct complete meshes from sparse data or meshes with missing sections, as demonstrated in cardiac mesh reconstructions ([123]). Nevertheless, these approaches exhibit a strong dependency on the template shape and necessitate extensive pre-processing to adjust and position the template accurately for dental crown generation. Recent methods like DiffComplete ([124]) offer a generative diffusion-based approach for completing 3D shapes from partial scans. It leverages hierarchical feature aggregation and occupancy-aware fusion to iteratively reconstruct meshes. The authors reported that on the 3D-EPN benchmark, DiffComplete demonstrated a 40% reduction in L_1 error compared to previous state-of-the-art methods, achieving an L_1 error of 0.053. Despite its advantages, DiffComplete has several limitations. The model struggles to complete highly irregular or noisy shapes. Additionally, the dense 3D CNN architecture limits the model's ability to handle high-resolution 3D shapes due to the cubic increase in computational costs with volume size. Although the model shows robust generalizability to unseen object classes, its performance may be adversely affected by the quality and diversity of the training data, necessitating careful selection to boost completion robustness. The multi-step inference process introduces substantial computational complexity, and it is only able to complete a small portion of the incomplete shape, which is not appropriate for dental crown generation. Direct mesh completion methods, including autoencoders and variational autoencoders (VAEs), have been proposed to learn a latent space of complete shapes for potential completions [94, 125]. However, the iterative optimization required at inference by these techniques, particularly by Point2Mesh ([13]), imposes considerable time constraints. This process becomes notably impractical in scenarios demanding quick processing, such as the individualized meshing of teeth for dental crown design. Recently, Neural Kernel Surface Reconstruction (NKSR) ([14]) introduced a framework utilizing neural kernel fields with compactly supported kernels and gradient fitting to reconstruct 3D surfaces from noisy point clouds. While NKSR effectively addresses challenges related to scalability and noise sensitivity, the method's requirement to handle a potentially large number of basis functions across the voxel hierarchy leads to heavy memory consumption. Additionally, the complexity of solving a linear system during the forward pass, particularly for large-scale inputs or those with high levels of noise, imposes a significant computational burden.

6.4.4 Evaluation metrics

In the field of point cloud completion, the Chamfer Distance (CD) and Earth Mover's Distance (EMD) [44] are standard metrics used to evaluate shape differences and point cloud

similarity. CD assesses shape disparities by calculating the average distance between nearest point pairs across two point clouds, whereas EMD seeks an optimal point-to-point mapping to minimize overall distances, striving for uniform distribution at the cost of higher computational demand and the need for equal-sized point sets. These metrics, however, can not fully capture true shape nuances due to their reliance on predefined matching rules, which can lead to inaccurate reconstructions. Addressing these limitations, the Density-aware Chamfer Distance (DCD) [126] introduces a refinement by integrating density considerations, offering a more nuanced evaluation that accounts for both global structure and local geometric details. Similarly, Hyperbolic CD (HyperCD) [127] modifies the Chamfer distance calculation by operating in hyperbolic space and applying a position-aware weighting in backpropagation to prioritize the preservation of accurate matches. Despite its innovative approach, HyperCD's computational complexity and the intricacies of working in non-Euclidean space pose practical challenges.

CALoss [128] emerges as another framework, blending contrastive and adversarial techniques within a non-linear representation space to distinguish shape differences more dynamically. While it enhances reconstruction quality, the computational overhead and complexity inherent in its dual approach raise considerations regarding model training efficiency. Furthermore, InfoCD [116] proposes a contrastive Chamfer distance loss that incorporates contrastive learning to refine point cloud completion. By maximizing the mutual information between geometric surfaces and distributing matched points more effectively, InfoCD improves upon traditional metrics by offering a robust and computationally efficient solution tailored for deep learning applications.

The methodology introduced in this work is inspired by the comprehensive review of point completion, mesh completion, and mesh reconstruction strategies. Given the complexity and information density of dental scans, we leverage the analytical capabilities of point cloud completion techniques, particularly those employing transformers, for their efficiency in processing 3D data. Despite their effectiveness, point cloud representations alone are not directly applicable in dental practice, necessitating a transition towards mesh outputs. To this end, our approach is also informed by mesh completion methods, which are adept at generating the necessary mesh structures but often lack the precision required for dental applications. Addressing this, we incorporate advanced mesh reconstruction techniques, such as differentiable Poisson surface reconstruction, to meet the high standards of accuracy essential for dental crown design. Furthermore, informed by the nuances of evaluation metrics, our methodology devises a loss function specifically aimed at enhancing the fidelity of margin line points, a critical aspect in the fitting of dental crowns. This integrated methodology aims to combine the strengths of each domain to achieve a precise, efficient solution tailored

for the automation of dental crown design.

6.5 Methods

Our model is an end-to-end supervised framework that takes a partial point cloud context as input and generates the missing region in a mesh format. The overview of our framework is illustrated in Fig. 6.2. It comprises two key components: a transformer model for encoding and decoding the data, and a specialized module designed to complete the mesh. Distinctively, our framework applies contrastive learning to fine-tune the network, utilizing a Chamfer distance metric. Furthermore, the framework is enhanced by a custom function specifically designed to meticulously refine the margin line of the crown. Moreover, we utilize a mean square error function to achieve precise control over the crown's indicator grid. The following sections provide a detailed analysis of each component within our network, along with the loss functions we have implemented.

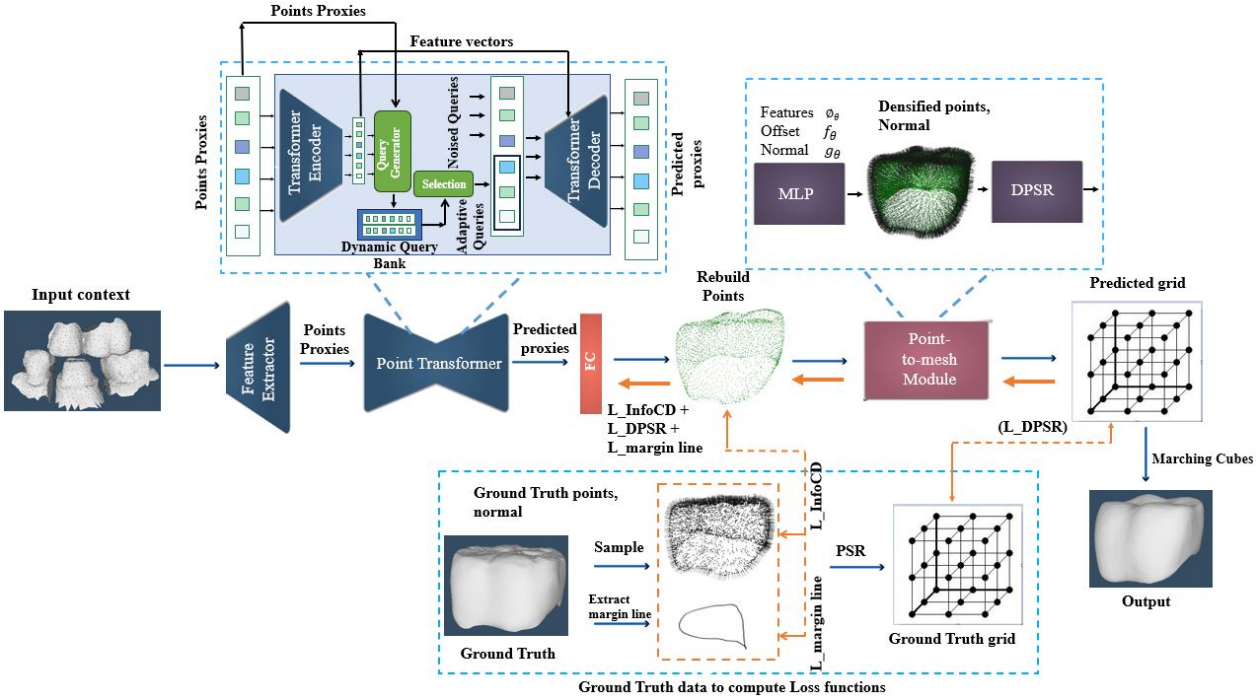


Figure 6.2 Architecture flow of our proposed model, highlighting key components such as DPSR (Differentiable Poisson Surface Reconstruction), FC (Fully Connected Layer), and MLP (Multi-Layer Perceptron).

6.5.1 Point Transformer Encoder-Decoder

We adapt the transformer encoder-decoder architecture ([12]) to extract both global and local 3D features from our input (context). The encoder processes the given context, while the decoder generates crown points, enabling our model to fill missing areas in point clouds by creating proxies for the absent segments. Our approach leverages the sequential generation capability of transformers, employing a set-to-set translation strategy to effectively map incomplete point cloud data to a corresponding completed version. This process begins by grouping the input points into a smaller set of features, converting the point cloud into a series of vectors through a point proxy method. This method utilizes farthest point sampling (FPS) ([41]) and a lightweight dynamic graph convolution network (DGCNN) ([129]) to capture local geometric structures in the context. The geometry-aware transformer encoder refines these features and captures intricate 3D spatial relationships. It does this through a modified self-attention mechanism, which incorporates k-nearest neighbors (KNN) within the attention blocks to output feature vectors. The self-attention layer in the encoder updates the feature vectors using both long-range and short-range information. To generate queries, which serve as the starting state of predicted proxies, we employ a multi-layer perceptron (MLP) query generator by [12]. This MLP dynamically adjusts queries, crafting them from the encoder output and the input point proxies. It establishes a dynamic query bank to ensure that the queries accurately reflect the desired output's sketch. Following this, the selection process uses a neural network scoring module to assess each query in the query bank, choosing a subset of the top-rated queries, referred to as 'adaptive queries', based on their informational content, without being bound by their categorization into abstracted or original sets. Integral to this process, the transformer decoder leverages self-attention and cross-attention mechanisms to interpret and refine the point cloud structure. It determines the internal configuration of input data and the connection between queries and the output feature from the encoder, translating this relationship into the accurate prediction of missing point proxies. Accompanying this, an auxiliary denoising task is essential for enhancing point precision; it introduces noised queries derived from the input's ground truth centers and infused with scaled random noise. These noised queries, along with the adaptive queries, are fed into the decoder, ensuring a consistent supply of high-quality queries for optimal decoding performance. To mitigate knowledge leakage from noised to adaptive queries, we employ an attention mask ([12]) within the self-attention layers of the decoder transformer. This mask selectively cancels out attention towards noised queries, safeguarding the integrity of adaptive queries and ensuring the reconstructed model's accuracy and reliability. The decoder's output finally passes through a fully connected layer to generate the prediction for crown point clouds.

6.5.2 Mesh Completion

Building directly upon the crown points generated by the transformer decoder, we progress to the mesh completion module. Here, we integrate a spectral method for Differentiable Poisson Surface Reconstruction (DPSR) as described by [10], enabling the transformation of discrete point clouds into continuous 3D surfaces. This technique discerns the zero level set of an indicator function within a structured 3D grid, differentiating points inside the shape from those outside. To refine the input unoriented crown points, a Multi-Layer Perceptron (MLP) network is employed. The MLP network extrapolates additional points and normals, thus densifying and orienting the point cloud for subsequent processing.

Next, we formulate the indicator function for the crown's shape as the solution to a Poisson partial differential equation (PDE) ([10]). This function is represented within a Fourier basis on a dense grid (at a resolution of 256^3). The spectral method computes the Fast Fourier Transform (FFT) of the gradient field of the indicator function, applies a Gaussian smoothing kernel to alleviate the ringing effects due to the Gibbs phenomenon, and solves the equation within the frequency domain for efficient computation ([10]). The solution to the PDE is then derived by manipulating the unnormalized indicator function χ' with an inverse FFT as equation 6.1. where $\hat{\chi}$ is the product of the spectral representation of the gradient field and the Gaussian kernel in the frequency domain. Subsequently, we refine χ', to obtain the final indicator function χ, scaling its absolute value at the zero level set and adjusting it with respect to the set C , representing the crown predicted points, as depicted in equation 6.2:

$$\chi' = \mathrm{IFFT}\left(\hat{\chi}\right), \tag{6.1}$$

$$\chi = \frac{m}{\mathrm{abs}\left(\chi'_{|x=0}\right)} \left(\chi' - \frac{1}{|\{C\}|} \sum_{c \in \{C\}} \chi'_{|x=c} \right) \tag{6.2}$$

This process ensures that the indicator function conforms to the exact geometry depicted by the predicted crown point cloud.

During training, our methodology employs a differentiable Poisson solver (DPSR) ([10]) to derive the estimated indicator grid from the predicted point cloud. Concurrently, we generate the ground truth indicator grid by running Poisson Surface Reconstruction (PSR) on a densely sampled point cloud obtained from the ground truth mesh, incorporating the exact corresponding normals. This setup ensures that the entire pipeline is fully differentiable, facilitating the simultaneous updating of multiple variables. Specifically, we refine point offsets, adjust oriented normals, and optimize network parameters. In the inference stage, the trained model uses the spectral solver method ([10]) to predict the indicator function

χ. The extracted scalar field from this function serves as the basis for the Marching Cubes algorithm to generate the final crown mesh. The specifics of this process are detailed in Algorithm 1, which outlines the step-by-step methodology for mesh reconstruction.

6.5.3 Loss Function

We use three different losses to constrain the output shape and predict a high-quality mesh. We utilize the InfoCD loss ([116]), a contrastive Chamfer distance loss designed to enhance point cloud completion tasks. Building upon the concepts of mutual information and contrastive learning, the InfoCD loss aims to improve the alignment of the completed point cloud with the ground truth. The overall loss function combines the traditional Chamfer distance loss with a novel regularization term, encouraging a uniform point distribution across the completed surface. This loss function is articulated as follows in equations 6.3, 6.4, 6.5 , 6.6:

$$L_{\text{InfoCD}}(x_i, y_i) = \ell_{\text{InfoCD}}(x_i, y_i) + \ell_{\text{InfoCD}}(y_i, x_i), \tag{6.3}$$

where each term ℓ_{InfoCD} is calculated by:

$$\ell_{\text{InfoCD}}(x_i, y_i) = -\frac{1}{|y_i|} \sum_k \log \left(\frac{\exp\left\{-\frac{1}{\tau} \min_j d(x_{ij}, y_{ik})\right\}}{\left(\sum_k \exp\left\{-\frac{1}{\tau} \min_j d(x_{ij}, y_{ik})\right\}\right)^{\lambda}} \right), \tag{6.4}$$

$$L_{\text{InfoCD}}(x_i, y_i) \propto \frac{1}{\tau} L_{\text{CD}}(x_i, y_i) + R(x_i, y_i) \tag{6.5}$$

To prevent the formation of clusters and encourage a more even spread of points, we use a regularization term $R(x_i, y_i)$, given by:

$$R(x_i, y_i) = \log \left\{ \sum_{m,n} \exp\left(-\frac{1}{\tau} \left[\min_j d(x_{ij}, y_{in}) + \min_k d(x_{im}, y_{ik}) \right] \right) \right\}, \tag{6.6}$$

where $x_i = \{x_{ij}\}$ and $y_i = \{y_{ik}\}$ denote the ground truth and predicted sets of 3D points, respectively. The m and n in the regularization term $R(x_i, y_i)$ assess the structural alignment and distribution uniformity by considering distances between all point pairs, enhancing the loss's sensitivity to the spatial arrangement and density of the point clouds. The function $d(\cdot, \cdot)$ signifies the chosen distance metric, such as the L1 or L2 norm ([9]). The parameters τ and λ (where $\lambda = \frac{\tau'}{\tau} \in (0, 1]$) control the trade-off between the fidelity of point matching and distributional uniformity, with smaller values of $L_{\text{CD}}(x_i, y_i)$ inducing larger $R(x_i, y_i)$. By tuning τ and τ', one can control the strictness of point matching and the extent of distribution

Algorithm 1 Mesh Completion algorithm

Require: Unoriented crown point cloud P, Neural network f_χ with parameters θ
Ensure: Reconstructed 3D surface mesh M
Forward Pass:
for each point p in point cloud P **do**
  Predict k offsets and normals N' for point p using f_χ
end for
Upsample point cloud P to create densified point cloud P'
Predict indicator function χ over P' using f_χ
Smooth normal vector field V by Gaussian filtering
Solve for indicator function χ by minimizing difference between $\nabla\chi$ and V'
Subtract mean value of χ at input points from χ
Reconstruct surface mesh M from χ using Marching Cubes
Backward Pass (Training Loop):
Require: Point samples $\{X_i\}$ and their corresponding indicator functions $\{x_i\}$
Ensure: Trained parameters θ of neural network f_χ
Define loss function $L(\theta)$, as given in equation 6.8
Initialize parameters θ of neural network f_χ
while not convergence **do**
  for each sample X_i in training data **do**
    Compute loss $L(\theta)$ for current X_i and x_i
  end for
  Update parameters θ by AdamW optimizer to minimize $L(\theta)$
end while
return M Reconstructed 3D surface dental crown mesh

alignment, allowing for a flexible approach to point cloud completion. The InfoCD loss, as developed by [116], serves as a bridge between the traditional CD loss and a more structured, distribution-sensitive metric.

We define a custom loss function to manage the margin line point clouds in dental crowns, which is crucial for ensuring an optimal fit on the prepared tooth (die). This function calculates the distances between points on the predicted crown, inclusive of the margin line, and the corresponding margin line points extracted from the ground truth at the base of the crown. The objective is to identify the nearest predicted points—approximately 300 in number—to the actual margin line. We refer to this function as compute_closest_points_function, which is utilized in the subsequent algorithm. Subsequently, we apply the mean Chamfer Distance (CD) ([44]) to constrain the locations of these points. The CD assesses the mean squared distance between two point clouds: S_1,representing the ground truth margin line,and S_2 consisting of points from the predicted crown nearest to the ground truth margin line. The measurement of individual distances is carried out between each point in one set and its

closest point in the other set, as delineated in equation 6.7 and Algorithm 2.

Let $L_{\text{margin line}}$ be the loss function for the margin line, defined as:

$$L_{\text{margin line}}(S_1, S_2) = \frac{1}{|S_1|} \sum_{x \in S_1} \min_{y \in S_2} |x - y|^2 + \frac{1}{|S_2|} \sum_{y \in S_2} \min_{x \in S_1} |y - x|^2 \tag{6.7}$$

Furthermore, we minimize the L_2 distance between the predicted and ground truth indicator functions, both derived from Poisson PDE solutions ([10]). This Mean Square Error (MSE) loss is computed as equation 6.8, where $f_\theta(X)$ represents a neural network (MLP) with parameters θ conditioned on the input point cloud X, D is the training data distribution, along with indicator functions x_i and point samples X_i on the surface of shapes.

$$L_{DPSR}(\theta) = E_{X_i, x_i \sim D} ||Poisson(f_\theta(X_i)) - x_i||_2^2 \tag{6.8}$$

The final loss function, represented as equal weighted sum of the individual loss functions, is described by equation 6.9:

$$L_{\text{total}} = L_{\text{InfoCD}} + L_{\text{margin line}} + L_{\text{DPSR}} \tag{6.9}$$

In our model, we use equal weighting for the three components of our loss function to ensure balanced influence during training. This method simplifies the optimization process and prevents bias toward any particular loss component, supporting well-rounded development.

6.6 Experiments and results

In this section, we assess our model's performance on a dataset encompassing all types of teeth within the jaw, including molars, premolars, canines, and incisors.

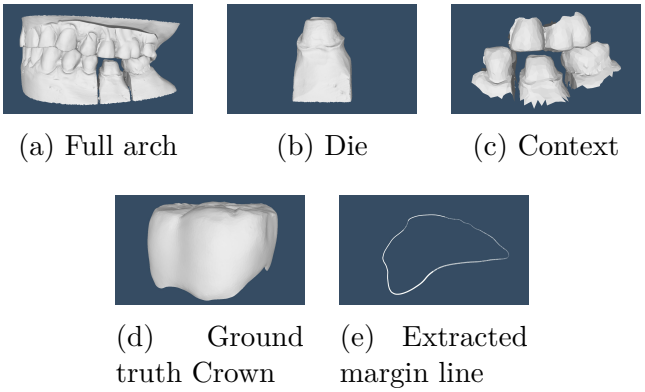
6.6.1 Dataset and preprocessing

In our initial phase, we generate a specific context from a provided 3D dental scan. This context includes the neighboring teeth adjacent to the prepared tooth and the two or three nearest teeth in the opposite jaw. Fig. 6.3 illustrates this process, showing the original scan alongside the generated context, which highlights the die and the margin line extracted from the ground truth crown.

To create this context, which is cropped from registered master and antagonist arches, we use the die file to precisely trim the master arch, aligning it with the die, which represents the prepared tooth, while preserving the two adjacent teeth. Additionally, in the opposing

Algorithm 2 Margin Line Loss Function.

1: **Input**: Predicted crown point cloud $\{P_i\}$, Ground truth margin line $\{M_i\}$, Number of closest points n
2: **Output**: Loss L_m
3: **for** each batch i in B **do**
4: Compute n closest points in P_i to M_i using *compute_closest_points_function*
5: Store indices of the closest points
6: Extract the closest points from P_i
7: Compute margin line loss $L_{m,i}$ between closest points and M_i using mean Chamfer Distance
8: **end for**



(a) Full arch (b) Die (c) Context

(d) Ground truth Crown (e) Extracted margin line

Figure 6.3 (a) Dental arch with a prep, (b) generated context, (c) die (prepared tooth), (d) Ground truth crown, (e) extracted margin line.

arch, we retain the two teeth closest to the die position. Using our ground truth crown, we extract the margin line as a spline and sample it into 300 points. This feature is crucial for ensuring the crown's proper fit, thereby enhancing its durability and the patient's comfort. Our research employs an experimental dataset comprising 388 training, 98 validation, and 71 test cases, enriched through data augmentation techniques including 3D translation, scaling, and rotation. For translation, each dental setup is randomly shifted in the 3D space within a range of -0.2 to 0.2 millimeters (mm) along each axis. In scaling, the size of each setup is randomly adjusted by a factor of 0.8 to 1.2, centered around the geometric center of the structure. Rotation is applied by generating a random rotation matrix for each setup, with angles varying between -0.35 and 0.35 radians around each principal axis. This process, applied to the entire dental setup (master arch, opposing arch, and crown), amplifies our training data tenfold. We selected 10,240 points to represent the dental context, providing two forms of ground truth for network training: mesh and point cloud crowns. To supervise network training using the ground truth meshes, we calculate the gradient from a loss on

an intermediate indicator grid. We use the spectral method from ([10]) to compute the indicator grid for our ground truth mesh.

6.6.2 Implementation details

Our model incorporates a transformer encoder-decoder module that aligns with the architecture detailed in [12]. For the network's mesh reconstruction phase, we employ the differentiable Poisson surface reconstruction (DPSR) method, as introduced in [10]. We implemented all models using PyTorch and optimized them with the AdamW optimizer, setting the learning rate to 5e-4 and the batch size to 4. The training process spanned 250 epochs, necessitating 22 hours of computing time on an NVIDIA A100 GPU.

6.6.3 Performance evaluation

To evaluate our network's performance and benchmark it against point cloud-based methodologies, we measure the dissimilarity between predicted outputs and ground truth data using the Chamfer distance. This measurement is implemented in two variants: CD_{L_1}, applying the L_1-norm, which sums the absolute differences along each dimension, and CD_{L_2}, utilizing the L_2-norm, which calculates the square root of the sum of squared differences between points. This allows us to compute the disparity between two point sets using both Manhattan and Euclidean metrics respectively. Furthermore, the Earth Mover's Distance (EMD) is employed for additional depth in our evaluation. EMD is particularly used in our ablation study to analyze the impact of our proposed model's components and validate its robustness in different scenarios. For assessing the similarity between the predicted indicator grids or meshes and their ground truth counterparts, we calculate the Mean Square Error (MSE) loss. We show the improvements achieved by our new model in Table 6.1, comparing it with our previous methods. Pointr+ Margin Line ([86]) builds on the point transformer approach ([9]) for point completion by integrating margin line data, which enhances accuracy at the critical margin line. However, this method requires additional post-processing to convert point clouds into meshes, adding complexity to the workflow. DMC ([11]) introduced a significant advancement by directly generating dental crown meshes from point clouds, streamlining the pipeline. Despite these advances, limitations in mesh quality and accuracy at the margin line persisted. Our new model addresses these limitations with several innovations, including a contrastive-based Chamfer distance (InfoCD) and a dedicated margin line loss function. These enhancements ensure more accurate alignment of crown meshes around the margin line and significantly improve mesh quality. As shown in Table 6.1, our model outperforms previous methods across all metrics, achieving a CD_{L_1} of 54.39 (a 12.3% improvement over

DMC) and a CD_{L_2} of 8.41 (a 23.9% improvement over DMC). Additionally, the MSE for output meshes is reduced to 0.0015, highlighting the model's ability to generate high-resolution, high-quality meshes with minimal error.

A key factor in creating an accurate dental crown mesh is the ability to predict point clouds with high resolution, devoid of noise. Our data are measured in millimeters, ensuring high precision in our evaluation metrics. Fig. 6.4 showcases the predictive capabilities of the DMC method where subfigure (a) demonstrates a prediction with 1536 point clouds, capturing lower resolution details, while subfigure (b), using the same method but at a higher resolution of 3072 points, exhibits increased noise and a lack of uniformity in point distribution. Conversely, our model, equipped with adaptive query generation and denoising capabilities, produces predictions with a more uniform resolution, as seen in subfigure (c), closely mirroring the high fidelity of the ground truth presented in subfigure (d). This comparison underscores the advancements our model introduces in the generation of high-quality dental crown meshes from noise-minimized point clouds. Furthermore, as indicated by Table 6.1, the evaluation metrics for our new model—specifically, Chamfer distance and MSE—show improvements compared to the previous method ([11]). This demonstrates that the predicted points are more closely aligned with the ground truth, validating the enhanced accuracy of our approach.

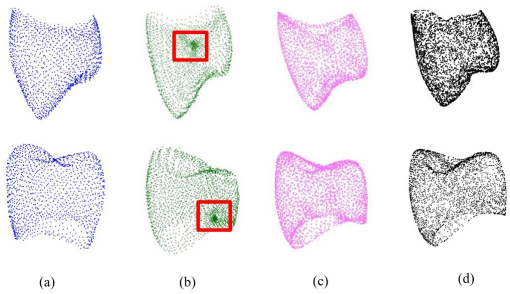


Figure 6.4 Comparison of predicted point clouds using our proposed method with previous work. (a) Illustrates the predicted point cloud using the method by [11] with 1536 points. (b) Depicts the same method by [11], but with a higher resolution of 3072 points, resulting in a noisier output. (c) Displays the prediction using our proposed model, and (d) represents the ground truth, with both (c) and (d) at the same resolution of 3072 points.

To provide a comprehensive assessment of our model, we compare it against several state-of-the-art methods. To ensure a fair comparison with the work detailed in [12, 115, 121], we utilized the official implementations provided by the authors. This approach ensured that all experiments were conducted under consistent conditions, using a unified dataset that includes every tooth position and adhering to a standardized training protocol. We

Table 6.1 Comparison of the new model with our previous methods. Evaluation metrics are CD_{L_1} and CD_{L_2} (multiplied by 1000; lower is better) and MSE (calculated on output meshes; lower is better). Bold values indicate improvements achieved by the new model. MSE is unavailable for the Pointr+ Margin Line experiment as it is calculated on output meshes.

Method	**CD-L1 (↓)**	**CD-L2 (↓)**	**MSE (↓)**
Pointr+ Margin Line ([86])	65.04	18.31	–
DMC ([11])	62.03	11.06	0.0028
Our Model	**54.39**	**8.41**	**0.0015**

compare our model against state-of-the-art methods in the domain, as shown in Table 6.2. ToothCR ([115]) employs a dual-network strategy, first predicting crown point clouds and then reconstructing surfaces from these points. AdaPoinTr ([12]), a point transformer-based model, completes the dental context by leveraging transformers. Few-point Shape Completion (FSC) ([121]) is designed to complete dental arch contexts using dental crown point clouds, with both master and antagonist arches as input contexts. While FSC's dual-branch feature extractor captures sparse details effectively, it underperforms compared to the AdaPoinTr model, which predicts crown point clouds with greater accuracy. As shown in Table 6.2, FSC achieves CD-L_1 and CD-L_2 scores of 72.86 and 17.62, respectively, while AdaPoinTr outperforms it with scores of 68.82 (CD-L_1) and 15.14 (CD-L_2), indicating closer alignment with the ground truth. However, our model surpasses both, achieving significant improvements with CD-L_1 and CD-L_2 scores of 54.39 and 8.41, respectively, and an MSE of 0.0015. These results demonstrate the superior accuracy and consistency of our method in generating dental crown shapes. To further illustrate these performance differences, (Fig. 6.5) visually compares the predictions of FSC and AdaPoinTr across four cases. In these examples, black points represent the ground truth crown points, red points indicate FSC predictions, and blue points represent AdaPoinTr predictions. As seen in the (Fig. 6.5) , AdaPoinTr predictions align more closely with the ground truth, with blue points consistently overlapping black points, while FSC predictions (red points) deviate noticeably from the actual crown shape. Fig. 6.6 provides visual results produced by our model, highlighting its ability to generate accurate, high-quality dental crowns compared to other methods.

In Fig. 6.7 , we present a color map analysis for four tooth types: Canine, Incisor, Molar, and Premolar, each visualized across instances with the lowest, median, and highest deviations from the ground truth. This visualization quantitatively represents the signed distance between the predicted 3D meshes and the actual ground-truth meshes, with colors indicating regions of overestimation (positive values) and underestimation (negative values). For

Table 6.2 Comparison of our model with the sate of the art methods in the domain. Metrics are the same as in Table 6.1.

Method	**CD-L1 (↓)**	**CD-L2 (↓)**	**MSE (↓)**
ToothCR ([115])	70.01	23.42	0.0024
AdapoinTr [12]	68.82	15.14	–
FSC ([121])	72.86	17.62	–
Our Model	**54.39**	**8.41**	**0.0015**

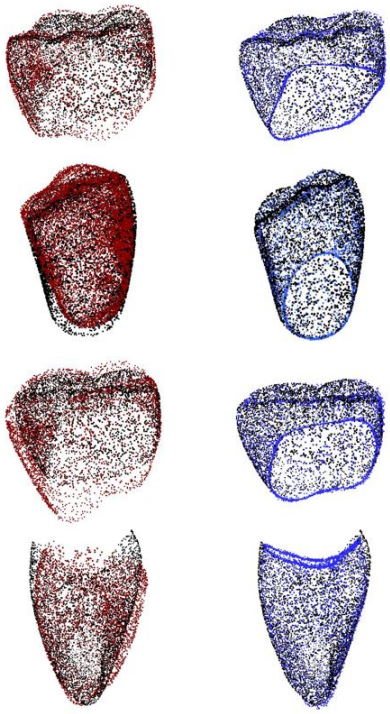
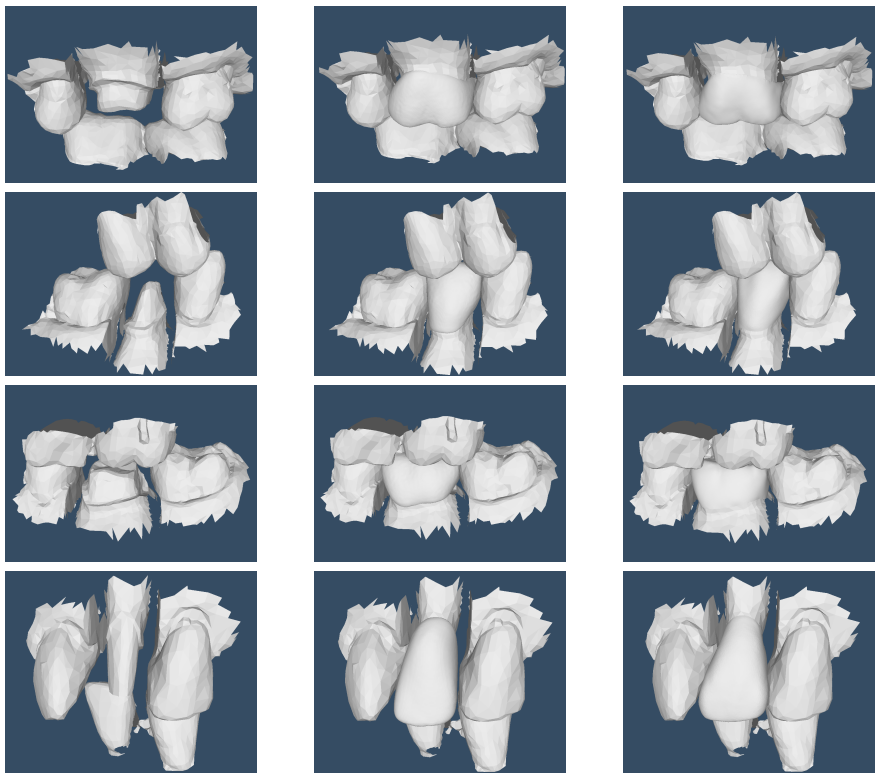


Figure 6.5 Visual comparison of dental crown predictions: Predictions from the FSC model (left column) and AdaPoinTr model (right column) are shown for four cases. In each visualization, black points represent the ground truth, red points indicate predictions by the FSC model, and blue points correspond to predictions by AdaPoinTr model. As depicted, the blue points from AdaPoinTr model consistently align closely with the black ground truth points, while red points from the FSC model exhibit noticeable deviations, indicating a lower accuracy in shape completion compared to AdaPoinTr approach.

Canines, the lowest deviation case (tooth 43) shows minor discrepancies, reflecting a strong alignment with the ground truth, while the highest deviation case (tooth 13) highlights pronounced misalignment, especially in regions of structural complexity. For Incisors, the lowest deviation case (tooth 31) displays a generally close alignment with some underestimation, whereas the highest deviation case (tooth 11) reveals a greater variability with both positive



(a) **Context** (b) **Prediction** (c) **Ground truth**

Figure 6.6 Examples of mesh completions by the proposed architecture. a) Input context containing master arch, prepped tooth and opposing arch; b) Generated mesh in its context; c) Ground truth mesh in its context.

and negative errors around the edges. In the Molar set, the lowest deviation case (tooth 46) aligns closely with the ground truth, but the highest deviation case (tooth 16) exhibits significant areas of discrepancy, suggesting challenges in capturing molar complexities. Finally, in the Premolar set, the lowest deviation case (tooth 14) shows near-uniform alignment, while the highest deviation case (tooth 15) has notable areas of overestimation, indicating regions where the predicted mesh extends beyond the ground truth. This color map analysis, with deviations ranging from -3 to +3 mm, provides critical insights into the model's performance, highlighting areas of both success and limitation across different tooth types and variation levels. To provide a thorough evaluation of the generated crown's functionality, our analysis includes both morphological comparison and spatial positioning relative to the neighboring teeth. While morphology is crucial for the crown's structural accuracy, spatial positioning is equally essential for assessing its fit and function within the context. We defined the find_intersection_measures function to compute the spatial relationship between a predicted crown and its neighboring teeth. By generating a 3D grid around the crown and

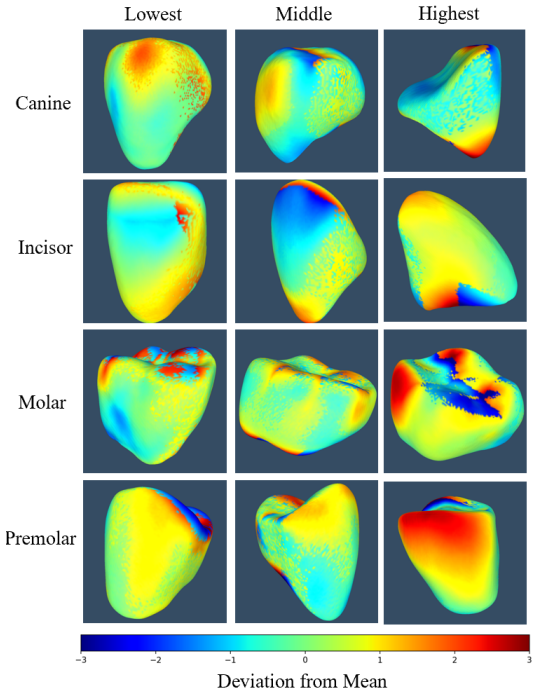


Figure 6.7 Color maps showing the signed distance (in mm) between predicted and ground-truth 3D meshes for four tooth types—Canine, Incisor, Molar, and Premolar—each visualized at low, median, and high deviation levels. Positive values indicate mesh overestimation, and negative values indicate underestimation. Canines (tooth 43, 13), Incisors (tooth 31, 11), Molars (tooth 46, 16), and Premolars (tooth 14, 15) reveal varying degrees of alignment accuracy, with notable discrepancies in complex regions. The analysis ranges from -3 to +3 mm, highlighting model performance across diverse dental structures.

context points and using Poisson Surface Reconstruction (PSR) fields, the function identifies regions where the crown overlaps or intersects with adjacent teeth. A threshold of 0.01 is applied to the PSR fields to detect surface points within this defined proximity, ensuring precise identification of overlapping areas. To ensure consistency across measurements, we converted the distance metrics, originally in 1D, to 2D values to make them comparable with the area-based measurements.

We applied find_intersection_measures function to 71 unseen cases, both for the predicted crown and the ground truth crown in the same context. The results, presented in the table 6.3, show the average, median, and standard deviation for left and right interpenetration distances and areas, all calculated using Mean Squared Error (MSE). The metrics include the left distance MSE^2, right distance MSE^2, left area MSE, and right area MSE. For example, the average left distance MSE^2 is 9.30×10^{-5}, while the average right distance MSE^2 is 1.54×10^{-4}, indicating larger spatial discrepancies on the right side. The left area MSE

(0.4169) and right area MSE (0.6478) offer further insight into surface overlap discrepancies, with the right side having a higher average error.

To further understand these spatial relationships, we computed intersection heatmap (Fig. 6.8) showing the distance to the nearest point between the crown and context (neighboring teeth) in millimeters. These cases were randomly selected from different tooth types among our unseen cases. In each case, subfigures (1) and (2) represent the ground truth and predicted distances, respectively, for both the crown and its context. In these subfigures, the right side shows higher values due to the calculated distances from context points, while the left side remains relatively consistent. Subfigures (3) and (4) show only the crown shell without context, with (3) being the ground truth and (4) the prediction. In these, the right-side values are lower as they represent distances solely between crown points and their nearest context points, without any overlap with context. These visualizations reveal that discrepancies are more prominent on the right side, correlating with the MSE results for interpenetration distances.

These metrics are essential for assessing and adjusting the fit of the crown, as the higher average and standard deviation on the right side indicate areas of larger spatial discrepancies that need further correction. The variability in errors, particularly on the left side where the standard deviation is high, underscores the need for case-specific adjustments. By analyzing these interpenetration and surface overlap measures, we gain a comprehensive understanding of spatial positioning and how well the generated crown aligns within the context—a critical factor for both functionality and overall dental alignment.

6.6.4 Ablation analysis

We conducted a comprehensive ablation study to assess the impact of various components within our newly developed model on its overall performance.

Impact of margin loss function

In our earlier discussion, we highlighted the impact of integrating a margin line loss into our model, as detailed in Table 6.4. Subsequently, Fig. 6.9 offers a visual comparison between our model's performance with and without the margin line loss, utilizing point clouds from the predicted meshes for a clearer visualization. This analysis highlights how the margin line loss enhances the precision and quality of the model's crown generation capabilities, particularly at critical margin line positions on the prepared tooth surface. As depicted in Fig. 6.9 panel (c), predictions without the margin line loss (represented in blue) do not fit

Table 6.3 Comparison of metrics for left and right distance MSE^2 and area MSE, scaled to 10^{-3}, across 71 unseen test samples. Each sample consists of a predicted crown and a corresponding ground truth crown, analyzed for interpenetration distances and surface areas on both the left and right sides. The metrics shown include the average, median, and standard deviation (STD) for each measurement, offering insight into spatial discrepancies and surface overlaps across samples.

Metrics	**Left distance MSE^2** ($\times 10^{-3}$)	**Right distance MSE^2** ($\times 10^{-3}$)	**Left area MSE** ($\times 10^{-3}$)	**Right area MSE** ($\times 10^{-3}$)
Average	0.0930	0.1544	416.8690	647.7620
Median	0.0033	0.0105	63.6040	156.7600
STD	0.4442	0.5356	1183.2580	930.8120

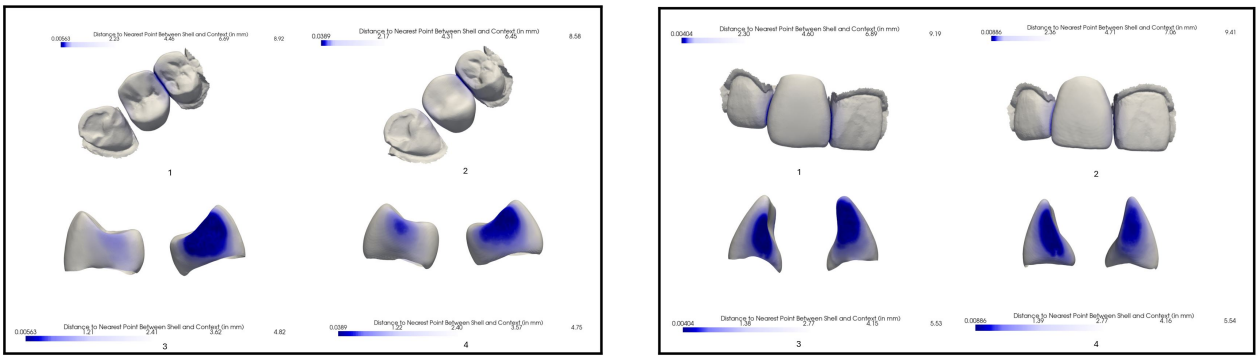
perfectly at the margin line, and sometimes exceed the ground truth dimensions. In contrast, predictions with the margin line loss (represented in green) align perfectly with the margin line and maintain the correct size compared to the ground truth.

Impact of point distance loss function

In this section, we compare our novel model with established loss functions commonly utilized in point cloud computation tasks, as detailed in Table 6.4. The loss functions under comparison include Chamfer Distance (CD) ([43]), Density-aware Chamfer Distance (DCD) ([126]), and Hyperbolic Chamfer Distance (HyperCD) ([127]). Our evaluation criteria encompass the average CD and Earth Mover's Distance (EMD) ([43]) across unseen datasets. The results demonstrate that our model, employing the InfoCD loss function ([116]), achieves a CD-L1 score of 57.61, a CD-L2 score of 9.45, and an EMD score of 77.67 (values multiplied by 1000, where lower is better), outperforming the alternatives. This superior performance indicates that InfoCD more effectively aligns the point distribution between the prediction and the ground truth, thus better measuring the underlying geometric surfaces of the point clouds through mutual information estimation.

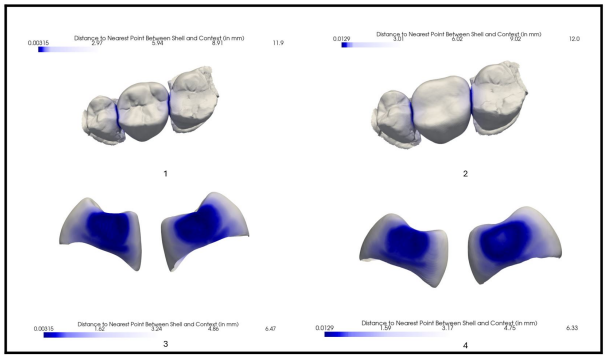
Impact of mesh completion module

In this section, we undertake a comprehensive evaluation, both visually and quantitatively, of several mesh reconstruction methodologies compared with our proposed framework. The evaluation begins with the point transformer ([12]) model, a method for completing points, as the baseline to measure the quality of the meshes. Subsequently, we incorporate the Shape as Points (SAP) methodology ([10]) to investigate its synergistic effect on mesh reconstruction, leveraging the point cloud output generated by the transformer. Following this, we

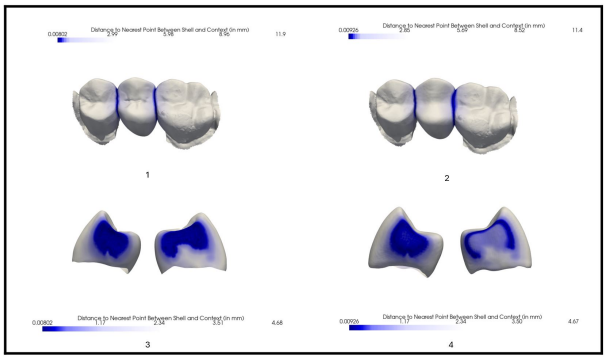


(a)

(b)



(c)

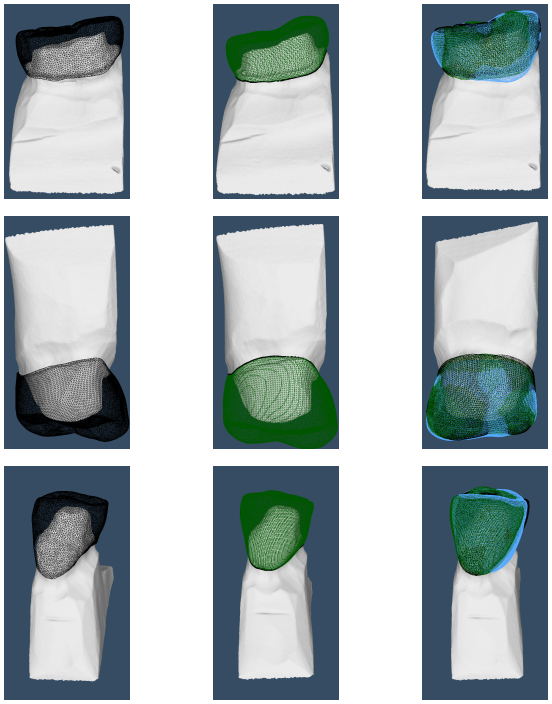


(d)

Figure 6.8 Heatmap visualizations of the intersection function, displaying multiple cases for different tooth positions: (*a*) position 24, (*b*) position 11, (*c*) position 16, and (*d*) position 25. In each case, subfigures (1) and (2) represent the ground truth and predicted distances, respectively, for both the crown and its context (neighboring teeth), with higher values on the right side due to distances calculated from context points. Subfigure (3) represents the ground truth and subfigure (4) the prediction, showing only the crown without context, resulting in lower right-side values as these distances are measured solely between the crown and its nearest context points.

Table 6.4 Comparison results of our model trained with some popular losses. Evaluation metrics are: CD_{L_1}, CD_{L_2}, EMD (multiplied by 1000, where lower is better.)

Loss Functions	**CD-L1 (↓)**	**CD-L2 (↓)**	**EMD (↓)**
CD ([43])	69.12	18.3	99.613
DCD ([126])	67.8	15.4	97.61
HyperCD ([127])	66.27	12.83	86.38
InfoCD ([116])	**57.61**	**9.45**	**77.67**

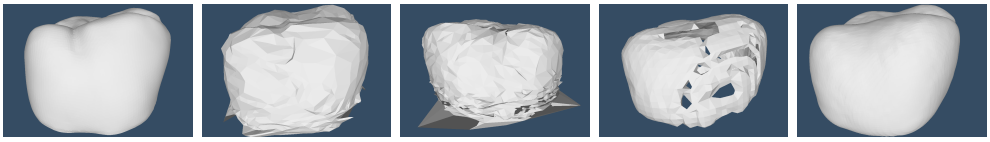


(a) (b) (c)

Figure 6.9 Illustration of the impact of incorporating a margin line loss function on our model's ability to generate dental crown meshes. Panel (a) showcases the ground truth crown for the prepared tooth surface. Panel (b) displays the model's generated crown utilizing the margin line loss, illustrating enhanced accuracy. Panel (c) combines these visualizations, juxtaposing the generated crowns (with and without margin line loss) against the ground truth on the prep. For clarity, black denotes the ground truth, green represents predictions with margin line loss, and blue signifies predictions without margin line loss, highlighting the qualitative differences in accuracy and detail.

integrate Neural Kernel Surface Reconstruction (NKSR) ([14]) into our evaluation, exploring its efficacy in refining meshes derived from the identical point cloud dataset. PoinT2Mesh ([13]) is then employed as an additional comparative measure, providing a comprehensive analysis relative to the aforementioned methods. The comparisons are quantitatively, employing metrics such as the CD, shown in table 6.5 and selectively visualized in Fig. 6.10. All experiments used the same dataset, which included all tooth positions, and were trained using the same methodology. For NKSR we provide point normals as an extra input channel. This ablation study shows the significant role of the differentiable point-to-mesh module within our transformer framework, illustrating its importance in enhancing the fidelity of fine mesh details. This enhancement substantially improves the overall quality of the dental crown meshes. The progressive improvements observed through this systematic examination

affirm the robustness and effectiveness of our integrated model, underscoring its superiority in the domain of high-fidelity mesh reconstruction.



(a) (b) (c) (d) (e)

Figure 6.10 Qualitative comparison of different approaches for crown mesh generation: (a) Ground truth shape; (b) point cloud processed using Point Transformer ([12]) and mesh generated by Shape as Points ([10]); (c) point cloud processed using Point Transformer ([12]) and mesh generated by Point2Mesh ([13]); (d) point cloud processed using Point Transformer ([12]) with mesh reconstruction by NKSR ([14]); (e) proposed method.

6.7 Discussion

This study introduces an innovative method for automating the design of dental crowns by integrating a transformer encoder-decoder with a differentiable Poisson surface reconstruction method, enhancing both the efficiency and accuracy of dental crown generation. Our approach employs a transformer encoder-decoder architecture equipped with self-attention and cross-attention mechanisms for advanced spatial feature extraction, coupled with dynamic query generation and a denoising task. This enables the generation of more accurate points compared with previous methods [11,86,115]. Additionally, the integration with differentiable Poisson surface reconstruction generates accurate meshes. DPSR transitions point clouds to continuous 3D surface meshes by solving a Poisson partial differential equation (PDE), providing precise mesh continuity and quality. To balance quality and computational efficiency, we incorporated Fourier-based spectral methods to streamline these operations. This optimization supports an overall processing time of less than five seconds, making the approach practical for real-world dental workflows where both speed and precision are paramount.

Crucially, our method uses a contrastive-based Chamfer distance loss function—InfoCD loss—which accurately aligns point cloud distributions and surpasses traditional Chamfer distance measures in matching precision. This enhanced loss function contributes to the method's precision in capturing the complex topography of tooth surfaces. Additionally, the integration of margin line loss targets crucial fit areas, ensuring precise alignment of crown meshes along the margin line, thus enhancing both the functional and aesthetic outcomes of the restorations. The integration of our method into clinical practice could revolutionize the

Table 6.5 Comparative analysis of mesh Reconstruction methods. Metrics are the same as in Table 6.4.

Method	**CD-L1 (↓)**	**CD-L2 (↓)**	**EMD (↓)**
Point Transformer (PT) ([12])	70.08	23.05	99.42
PT + SAP ([10])	67.01	21.12	96.03
PT + NKSR ([14])	61.43	18.34	86.23
PT + Point2Mesh ([13])	59.62	15.27	80.01
Our model	**54.39**	**8.41**	**75.31**

speed and accuracy of dental crown production. Our approach reduces the crown generation process to less than five seconds, compared to the traditional methods that typically take about an hour 6.1. This significant reduction in production time not only makes high-quality dental care more accessible but also drastically cuts down the turnaround time for patient treatment, thereby enhancing the efficiency of dental practices.

Our model shows improvements over existing methods in the domain, as detailed in Tables 6.1 and 6.2. The performance evaluation demonstrates that, when trained with InfoCD loss and MSE, it achieves superior accuracy and alignment with the ground truth. Specifically, the incorporation of the margin line loss has improved the fit and adjustment of the generated crown within its context. Compared to ToothCR ([115]), the Pointr+margin line approach ([86]), DMC ([11]), the AdapoinTr model ([12]), and FSC ([121]), it consistently outperforms these methods in terms of Chamfer distance and MSE metrics. These results validate the robustness of our approach in generating high-quality dental crown meshes.

Despite these advancements, our model has some limitations. The global receptive field of transformer models, while offering extensive context, can be a double-edged sword. The transformer's ability to capture global features comes at the expense of losing intricate details. This characteristic can limit the model's capacity to predict and generate the fine, nuanced structures essential for dental applications, such as the subtle grooves and bumps on a tooth's surface. Furthermore, the current representation approach, which involves learning features from the input point cloud and predicting the crown points, faces challenges in capturing the necessary detailed features. Because point cloud representation does not include the information between points, such as faces or cells, learning features from them alone is not enough to predict high-quality meshes with intricate details. Although our method strives to seamlessly predict the indicator grid by leveraging point cloud context, it falls short in learning the intricate details required for accurate 3D shape generation. This limitation is particularly evident when dealing with complex dental geometries. Additionally, Fig. 6.7

highlights occasional discrepancies where the predicted model is either slightly larger or smaller than the ground truth, which are crucial for the functional aspects of crowns. Such discrepancies can impact the fit and effectiveness of the dental crowns, necessitating further refinement of the model to improve its accuracy in these critical areas. These limitations underscore the need for future research to refine our model's algorithms for better detail capturing and to explore its adaptability across a broader range of dental reconstruction tasks. The specifics of this future work are outlined in the next section.

6.8 Conclusion and future work

Our research proposes a novel model for creating detailed mesh models of dental crowns from incomplete point cloud data. Leveraging an end-to-end model, our approach seamlessly generates high-quality meshes, accurately representing every tooth position. This model harnesses the power of transformers, enhanced by dynamic query generation and an adaptive denoising task, to precisely interpret the complex geometry of dental structures. Central to our methodology is the application of a differentiable Poisson surface reconstruction solver, which intelligently infers the positioning of crown points. These are then transformed into definitive mesh surfaces through the Marching Cubes algorithm, ensuring a smooth and accurate representation of dental crowns. The entire network operates under the guidance of a contrastive learning chamfer distance and an innovative function designed to precisely manage the margin line points, in conjunction with an MSE for the output mesh. The efficacy of our technique is confirmed by experimental results, which show our model's ability to produce high-quality crown meshes that fit precisely along the margin line.

An interesting direction for future research could involve addressing the challenges posed by the global receptive field of transformer models in capturing intricate details. Diffusion models are emerging as a promising alternative, potentially offering better detail capture and accuracy in 3D shape completion. Additionally, to tackle the representation problem, exploring the application of spherical harmonics in dental reconstruction offers a promising avenue. Spherical harmonics ([130]) are a set of solutions to Laplace's equation defined on the sphere, often used in physics and engineering, particularly for analyzing functions defined on the surface of a sphere or for applications involving spherical symmetry. In the context of dental restoration, a neural network trained to understand and reconstruct these coefficients could potentially enhance the smoothness and precision of the mesh surfaces by capturing both the detailed and overall shape of individual teeth. Furthermore, defining a functionality-aware loss function could help ensure that the predicted tooth has better contact points with adjacent and opposing teeth, considering the practical aspects of chewing functionality. By

exploring these avenues, we can address the current limitations and continue to advance the field of automated dental crown design, ensuring both the aesthetic and functional success of dental restorations.

6.9 Declaration of Competing Interest

Golriz Hosseinimanesh reports financial support from Intellident Dentaire Inc., iMD Research Inc., NSERC, and MEDTEQ. All other 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.

6.10 Acknowledgments

We acknowledge the help and funding received from Intellident Dentaire Inc., iMD Research Inc., the Natural Science and Engineering Research Council of Canada (NSERC), and MEDTEQ. A. Alsheghri acknowledges the support of King Fahd University of Petroleum and Minerals (KFUPM) [Ref: EC241009]. The authors would also like to thank the Digital Research Alliance of Canada for their continued support of our research and access to computational resources. Additionally, we thank Victoria-Mae Carrière for her assistance with editing the paper. The work presented has not been published or presented elsewhere.

CHAPTER 7 ARTICLE 4: AUTOMATIC DENTAL CROWN GENERATION WITH SPATIAL CONSTRAINT MODELING AND ADVERSARIAL SURFACE REFINEMENT

Golriz Hosseinimanesh[a], Farida Cheriet[a], Ammar Alsheghri[b], Victoria-Mae Carrière[a], Julia Keren[c], François Guibault[a]

[a]Polytechnique Montréal, Department of Computer and Software Engineering, P.O. Box 6079, Montréal, H3C 3A7, Québec, Canada

[b]Interdisciplinary research center for Biosystems and Machines, King Fahd University of Petroleum and Minerals (KFUPM), Dhahran, KSA

[c]Intellident Dentaire Inc., Montréal, Québec, Canada

Presentation: This work was submitted to the SPIE Medical Imaging journal on September 11, 2025.

7.1 CRediT authorship contribution statement

Golriz Hosseinimanesh: Methodology, Software, Validation, Investigation, Data Curation, Writing - Original Draft. **Ammar Alsheghri**: Writing - Review& Editing. **Victoria Mae Carrière** : Experiment - Review& Editing. **Julia Keren**: Conceptualization, Resources. **Farida Cheriet**: Supervision. Writing—Review & Editing. **François Guibault**: Supervision, Project administration, Funding acquisition, Resources, Writing - Review & Editing.

7.2 Abstract

Purpose: Deep learning algorithms offer potential to automate dental crown generation, reducing time-intensive manual design in dental laboratories. However, achieving crowns suitable for direct clinical use requires both geometric precision and functional accuracy to minimize post-generation adjustments. Current approaches focus primarily on shape completion without explicitly modeling critical spatial relationships, including margin line boundaries, occlusal contact patterns, and adjacent tooth interactions. These limitations result in generated crowns lacking spatial accuracy necessary for direct clinical application.

Approach: We present a comprehensive framework employing transformer encoder-decoder architecture integrated with differentiable Poisson surface reconstruction for direct dental

crown mesh generation. The framework incorporates two key innovations to address clinical limitations. First, margin line data is integrated as direct network input, concatenated with master and antagonist arch geometries, providing explicit boundary constraints during crown generation. Second, spatial constraint losses ensure anatomically valid relationships through antagonist interaction loss for proper occlusal contact patterns and intersection loss to prevent crown penetration into adjacent teeth.

Results: The framework demonstrates substantial improvements over existing state-of-the-art methods, achieving 35.9-40.6% improvements in geometric accuracy metrics. Margin line integration yields 31.2% improvement in geometric precision with maximum boundary errors reduced from 1.37 to 0.74 mm and 58.4% reduction in variability. Antagonist interaction loss provides 9.51% improvement in occlusal alignment, while intersection loss substantially reduces crown penetration into adjacent teeth.

Conclusions: Substantial performance improvements validate effectiveness of integrating spatial constraint modeling and direct margin line input into the generation process, establishing a foundation for clinical deployment of automated dental crown design systems.

Keyword: Dental crown design, Digital dentistry, Spatial constraint modeling, Geometric deep learning, Automated restoration design, Margin line integration

7.3 Introduction

Computer vision and artificial intelligence have advanced significantly in recent years, with discriminative approaches such as object detection, semantic segmentation, and pattern recognition now successfully deployed across numerous real-world applications [131, 132]. Generative approaches, particularly those focused on creating realistic 3D models, have shown remarkable promise in specialized domains requiring high precision and customization [133]. In medical imaging and computer-aided manufacturing, these technologies are increasingly being applied to automate complex design tasks that traditionally require extensive human expertise [134, 135].

Dental restorations represent a highly challenging application of automated design due to the precise geometric relationships required between the restoration and surrounding anatomical structures [136]. The conventional dental crown fabrication process begins with tooth preparation, where the dentist removes damaged sections and shapes the tooth to receive a crown. A digital impression is typically captured using an intraoral scanner, providing 3D data of the prepared tooth and the surrounding dentition. This data must then be processed to design a crown that satisfies multiple critical requirements: (1) perfect fit to the patient's

prepared tooth, (2) appropriate contact relationships with adjacent and opposing teeth, (3) functional occlusal morphology for effective chewing function, and (4) aesthetically pleasing appearance [137, 138]. Figure 7.1 illustrates the dental crown design workflow, showing the input context and target crown specifications.

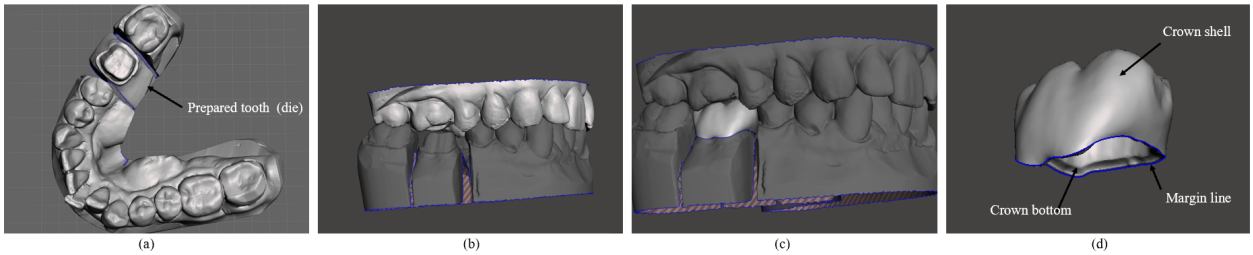


Figure 7.1 Dental crown design workflow: (a-c) Input dental context showing prepared tooth and surrounding spatial relationships, (d) resulting target crown with precise margin line boundaries. The complexity of achieving proper geometric fit and spatial constraints drives the development of automated design solutions.

Current computer-aided design (CAD) technologies in dentistry rely heavily on predefined template libraries and manual adjustment by skilled dental technicians [139]. The typical workflow involves selecting an appropriate tooth template, positioning it on the prepared site, and manually modifying the geometry to achieve proper fit and function. This process is time-demanding and requires extensive expertise to navigate the complex geometric constraints involved in crown design [140]. While CAD-CAM technologies have streamlined manufacturing processes, the design phase remains a significant bottleneck requiring substantial human intervention, customization, and expertise. These limitations have motivated researchers to explore machine learning approaches that can automate the design process by learning from existing crown designs.

Recent advances in deep learning have enabled automated crown design systems, yet fundamental challenges remain unaddressed for clinical deployment. Most automated methods lack explicit control over spatial interactions between the generated crown and surrounding dental structures [84, 89, 141], being optimized primarily for geometric similarity rather than incorporating the critical spatial constraints essential for proper crown function. This limitation potentially results in designs that require significant manual correction, undermining the efficiency gains that motivate automated approaches.

Our previous efforts exemplify this limitation. We explored point-based crown generation by integrating margin line data from the prepared tooth, which enhanced precision in the critical margin line region [86]. Subsequently, we developed an end-to-end framework that directly generates crown meshes through transformer-based architecture with differentiable surface

reconstruction [15, 92]. While these approaches demonstrated improved geometric accuracy, they too lacked explicit spatial constraint modeling essential for clinical functionality.

This widespread limitation across automated crown design reflects the primary challenge in the field: ensuring proper functional relationships alongside morphological accuracy. Functionality requirements involve complex spatial constraints including appropriate contact patterns with opposing teeth, optimal occlusal engagement for effective chewing function, and prevention of interference with adjacent dental structures. These requirements demand systematic modeling of 3D spatial relationships that traditional template-based approaches struggle to consistently achieve, with complex 3D spatial interactions between the crown and surrounding dental structures presenting particular challenges for consistent manual design.

While most automated approaches have focused solely on geometric reconstruction, only limited prior work has attempted to address functionality constraints explicitly. Among these efforts, early approaches formulated crown generation as 2D image synthesis problems. Hwang et al. [1] developed conditional GANs that convert 3D dental scans to 2D depth images, introducing both hard constraints (preventing crown penetration into opposing teeth) and soft constraints (capturing natural gap statistics for proper occlusal function). Their approach incorporates space information through gap distance maps and uses statistical histogram matching to enforce functionality constraints, demonstrating systematic evaluation through penetration analysis and contact point distribution measurements. Yuan et al. [81] extended this 2D paradigm by introducing perceptual loss functions and occlusal groove filter networks, conditioning generation on antagonist geometry through curriculum learning to progressively incorporate functional constraints. While these approaches pioneered functionality constraint integration and demonstrated that automated methods could address relationships typically managed by technicians, their reliance on 2D depth representations necessitates CAD-based post-processing to recover 3D models, limiting clinical applicability. Furthermore, the 2D formulation inherently constrains the modeling of complex 3D spatial interactions essential for comprehensive functionality assessment.

To address these limitations and advance automated crown design with explicit functionality modeling, we extend our end-to-end dental crown generation framework with spatial constraint modeling. We introduce (1) margin line data integration as additional input along with master and antagonist arches for enhanced crown fitting, (2) intersection loss functions that explicitly prevent crown penetration into adjacent teeth, (3) antagonist interaction loss that optimizes occlusal contact relationships with opposing dentition.

The primary contributions of this work include

- Margin line data integration for enhanced boundary precision: Extension of our proven approach of incorporating margin line information as additional input data alongside master and antagonist arch from point-based crown generation [86] to end-to-end mesh generation, ensuring critical boundary accuracy throughout the differentiable reconstruction pipeline.
- Spatial constraint loss functions for adjacent teeth interaction: Novel intersection loss functions that enforce geometric constraints preventing crown penetration into adjacent teeth while maintaining proper proximal contact relationships, ensuring anatomically valid spatial arrangements critical for preventing clinical complications [142].
- Antagonist interaction optimization: Development of antagonist interaction loss that optimizes natural contact relationships with opposing dentition, preventing interference patterns that could compromise restoration functionality.

7.3.1 Related work

Our approach draws on advances in 3D shape generation, spatial constraint modeling, and automated dental design. This section examines prior work in two major areas that inform our methodology: (1) the evolution of AI-driven dental crown generation and (2) spatial constraint modeling techniques for 3D shape generation.

AI-Driven Dental Crown Generation

The evolution of automated crown generation has progressed through distinct technological phases, each seeking to improve geometric fidelity while moving closer to clinical applicability. However, the integration of explicit functionality modeling has remained limited throughout this progression.

Early approaches formulated crown generation as 2D image synthesis problems, with some pioneering attempts at functionality constraint integration. As discussed earlier, Hwang et al. [1] and Yuan et al. [81] developed the first systematic approaches to incorporate functional constraints through statistical modeling of gap distances and occlusal contact patterns. Subsequent 2D methods attempted to retain more geometric information through enhanced modeling approaches. Multistage pipelines such as DCPR-GAN [82] and DAIS [106] combined depth prediction with surface reconstruction, while dual-discriminator architectures [107] refined local and global structures simultaneously. However, these methods remained fundamentally constrained by 2D representations, unable to recover complete crown geometry or capture full 3D spatial interactions necessary for comprehensive functionality assessment.

To overcome dimensional limitations, the field shifted toward 3D approaches that preserve geometric information throughout the generation process. Point cloud methods emerged as the dominant 3D paradigm, with Lessard et al. [89] proposing hierarchical point cloud completion and Zhu et al. [84] employing transformer-based contextual feature extraction for crown generation. While these methods improved geometric fidelity over 2D approaches, they largely abandoned the explicit functionality modeling that characterized earlier 2D approaches, instead relying on implicit geometric similarity to achieve functional adequacy. Moreover, point cloud approaches produce outputs requiring subsequent mesh reconstruction, creating multi-stage pipelines prone to error accumulation and limiting end-to-end optimization.

Recent frameworks have emphasized direct mesh generation to simplify pipelines and enable complete end-to-end training. DCrownFormer [141] applies morphology-aware transformers for direct mesh prediction from preparation and antagonist scans, eliminating intermediate reconstruction steps while maintaining geometric accuracy. Similarly, Ding et al. [108] demonstrated that 3D-DCGAN frameworks achieve high morphological fidelity and biomechanical performance comparable to natural teeth. However, these direct mesh generation approaches, while achieving high geometric accuracy, do not incorporate explicit mechanisms to enforce functional relationships with adjacent or opposing teeth, representing a step backward from the functionality modeling pioneered in earlier 2D work.

This progression reveals a fundamental limitation in the field: while early 2D methods successfully incorporated explicit functionality constraints, they were constrained by dimensional limitations that prevented accurate 3D spatial modeling. Conversely, modern 3D approaches achieve superior geometric fidelity but lack the systematic functionality modeling that made earlier methods clinically relevant. This creates a critical gap where no existing approach combines the dimensional advantages of 3D representation with explicit spatial constraint enforcement necessary for clinical deployment.

Spatial Constraint Modeling in 3D Shape Generation

The enforcement of spatial constraints in 3D shape generation represents a critical yet underexplored area, particularly for domains where collision-free geometries and anatomically valid arrangements are essential for functional viability. Unlike traditional generation tasks that prioritize visual similarity, constraint-aware generation must ensure that synthesized objects satisfy specific spatial relationships relevant to their intended function.

In computer graphics and general 3D modeling, spatial consistency has typically been addressed through geometric penalty functions. Signed distance fields and occupancy grids are commonly used to enforce object separation via proximity penalties [143], while mesh defor-

mation methods rely on vertex-face penalty terms or pairwise distance approximations [144]. These techniques effectively prevent unintended intersections but are designed for general collision avoidance rather than task-specific functional requirements.

Advanced spatial analysis technologies have been developed for intersection detection and quantification in specialized applications. Wang et al. [145] introduced efficient voxel-based intersection methods for complex geometries, while Chen et al. [146] applied clustering-based boundary detection specifically for 3D anatomical data. However, these approaches operate as post-processing tools for validation or measurement rather than enforcing constraints during the generation process itself, limiting their utility for end-to-end optimization.

In clinical and anatomical contexts, spatial accuracy directly determines functional outcomes. Orthodontic modeling, implant planning, and craniofacial reconstruction all require precise spacing and clearance to preserve anatomical validity and function [147–149]. Within dentistry specifically, occlusal functionality depends critically on accurately modeled contact relationships. Computer-aided systems have been developed to identify occlusal contacts and prevent interferences [150], digital models have been used to quantify contact areas across different malocclusion types [151], and validation studies confirm that virtual occlusion models can accurately reproduce real-world contact patterns [152]. More recently, digital force mapping techniques have enabled three-dimensional functional representations of occlusal pressure distribution [153].

Despite these advances in spatial analysis and clinical understanding of functional requirements, integrating spatial constraint modeling directly into generative processes remains largely unexplored, particularly for applications requiring both geometric fidelity and functional accuracy. This gap represents a critical barrier to clinical deployment of automated design systems across medical and dental applications.

7.4 Material and methods

7.4.1 Dataset

The dataset consists of 557 clinical dental crown cases, split into 388 for training, 98 for validation, and 71 for testing. Cases span all teeth types, including incisors, canines, premolars, and molars, from both maxillary and mandibular arches.

For each case, the dental context is constructed by cropping the registered master and antagonist arches using the die file as a reference. The die, which marks the prepared tooth, is used to trim the master arch while preserving the two adjacent teeth on either side. In the opposing arch, the two nearest teeth to the die are retained to maintain anatomical

context [15].

Each dental setup is uniformly sampled to 10,240 points, capturing the geometric structure needed for training. Ground truth crowns are provided in both mesh and point cloud formats [15]. For mesh-based supervision, indicator grids are computed using the spectral method from DPSR [10], allowing differentiable training with continuous shape representations.

Each case also includes a margin line extracted from the crown through boundary edge detection, interpolation, and optimal point ordering [110]. This process produces uniformly sampled points per crown that serve as geometric constraints during training to enforce boundary level accuracy.

To improve model generalization, the dataset is augmented ten-fold using rigid 3D transformations applied to the entire setup. These include random translations of ± 0.2 mm along each axis, isotropic scaling from $0.8\times$ to $1.2\times$, and rotations of ± 0.35 radians around each principal axis. All augmentations preserve anatomical relationships to ensure consistency during training [15].

7.4.2 Methods

Network Architecture

Our approach builds upon the end-to-end supervised framework previously published in [15], which takes partial point cloud context as input and generates missing dental regions in mesh format using a transformer-based encoder-decoder architecture integrated with differentiable Poisson surface reconstruction (DPSR). The baseline architecture details are provided in our previous work [15].

Margin Line Integration

We extend this framework to incorporate margin line data as additional input, building upon our previous demonstration of margin line integration benefits in point-based crown generation [86]. As shown in section (a) of Figure 7.2, the enhanced input specification includes: (1) master arch context containing the prepared tooth (die) and two adjacent teeth, (2) antagonist arch information providing opposing teeth geometry for spatial relationship modeling, and (3) margin line geometry extracted as a spline curve and uniformly sampled points representing the critical crown-tooth interface.

The margin line integration requires preprocessing to ensure geometric compatibility with the transformer architecture. Margin line points are concatenated with the master arch context

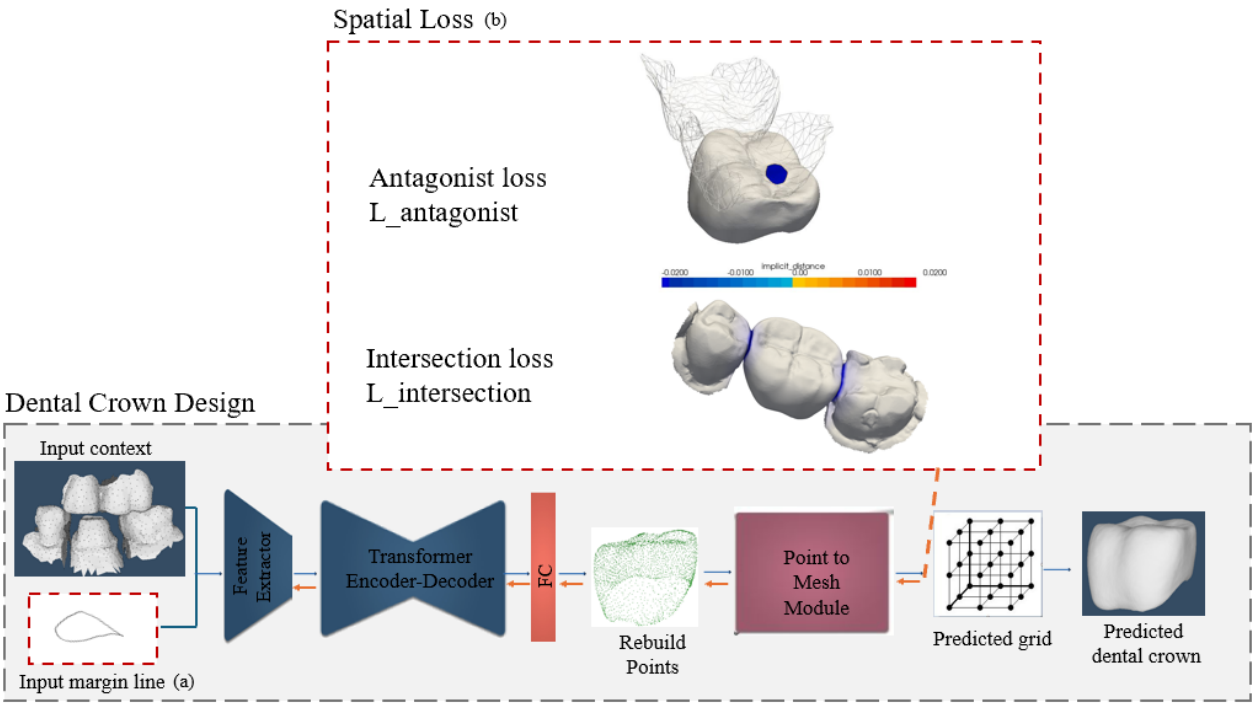


Figure 7.2 Dental crown generation framework. Building upon our published architecture [15], the framework incorporates two novel enhancements: (a) margin line input integration and (b) spatial constraint loss functions comprising antagonist interaction loss and intersection loss for anatomically valid generation. Red dashed boxes highlight the novel contributions, while gray components represent the baseline architecture.

points, creating a unified input representation that preserves both spatial context and critical boundary constraints. For training supervision, the margin line points extracted from the ground truth crown are concatenated with the ground truth crown point cloud, ensuring the network learns to predict both crown geometry and accurate margin line positioning. This multi-modal input and supervision strategy enables the network to explicitly consider margin line constraints during crown generation, improving fitting accuracy and clinical viability. The enhanced input processing accommodates the increased input dimensionality while ensuring that margin line constraints are propagated through the feature representations during encoding and subsequent crown prediction.

While the input specification improves geometric constraints, ensuring anatomically valid crown generation requires additional spatial constraint losses, as described in the following sections.

Loss Function

Our framework employs morphology losses from our published work [15] enhanced with spatial constraint terms for proper occlusal relationships and collision avoidance.

Morphology Loss The morphology reconstruction employs two loss components from our previous work [15]. The InfoCD loss [154] is a contrastive Chamfer distance designed to enhance point cloud completion by improving alignment between predicted and ground truth point clouds through mutual information and contrastive learning principles. The second component minimizes the L_2 distance between predicted and ground truth indicator functions derived from Poisson PDE solutions [10], ensuring accurate surface reconstruction through a Mean Square Error (MSE) loss on the indicator grid. Detailed mathematical formulations for these morphology losses are provided in our previous publication [15].

Beyond these morphology constraints, our framework introduces two spatial constraint losses to address clinical requirements for proper occlusal relationships and collision avoidance, as illustrated in Section (b) of Figure 7.2.

Spatial Losses

Antagonist Interaction Loss The antagonist interaction loss ensures optimal occlusal contact relationships between the generated crown and opposing dentition by minimizing the difference in spatial proximity patterns between predicted and ground truth crowns relative to the antagonist teeth. Let $P_{pred} \in \mathbb{R}^{N\times 3}$ denote the predicted crown point cloud, $P_{gt} \in \mathbb{R}^{N\times 3}$ represent the ground truth crown point cloud, and $A \in \mathbb{R}^{N\times 3}$ denote the antagonist tooth point cloud sampled to size N. The symmetric Chamfer L1 distance between point sets X and Y is defined as:

$$\mathrm{CD}_{L1}(X,Y) = \frac{1}{|X|}\sum_{x\in X}\min_{y\in Y}||x-y||_1 + \frac{1}{|Y|}\sum_{y\in Y}\min_{x\in X}||y-x||_1 \tag{7.1}$$

The antagonist interaction loss is then formulated as:

$$L_{antagonist} = |\mathrm{CD}_{L1}(P_{pred}, A) - \mathrm{CD}_{L1}(P_{gt}, A)| \tag{7.2}$$

This loss function encourages the predicted crown to maintain the same spatial relationship with the antagonist teeth as observed in the ground truth, ensuring clinically appropriate occlusal engagement. The computational procedure is detailed in Algorithm 1.

Algorithm 1 Antagonist Interaction Loss Function.

1. **Input:** Predicted crown point cloud P_i, Ground truth crown point cloud G_i, Antagonist tooth point cloud A_i, Number of points N
2. **Output:** Loss L_{inter}
3. **for** each batch i in B **do**
4. Sample N points from A_i using farthest point sampling: $A_{i,N} = \text{FPS}(A_i, N)$
5. Compute Chamfer L1 distance: $D_{pred} = \text{CD}_{L1}(P_i, A_{i,N})$
6. Compute Chamfer L1 distance: $D_{gt} = \text{CD}_{L1}(G_i, A_{i,N})$
7. Compute interaction loss: $L_{inter,i} = |D_{pred} - D_{gt}|$
8. **end for**

Intersection Loss Function The intersection loss prevents crown penetration into adjacent teeth by enforcing consistent interpenetration patterns between predicted and ground truth crowns relative to the master arch context. This loss leverages pre-computed intersection measurements that quantify the spatial relationship between crowns and adjacent dental structures.

During preprocessing, we compute intersection measurements for each ground truth case by analyzing the interpenetration between the crown and master arch using PSR field analysis. The crown and context are normalized to a common coordinate space, and their respective PSR fields are computed using differentiable Poisson surface reconstruction [10]. Intersection regions are identified where both PSR fields indicate solid material ($\psi_{crown} \leq \tau$ and $\psi_{context} \leq \tau$, where τ is the threshold). These intersection points are separated into left and right regions using principal component analysis (PCA) along the master arch's oriented bounding box axis. For each side, we compute the interpenetration distance (smallest extent of the oriented bounding box (OBB)) and the intersection area (product of the two larger extents). During training, corresponding intersection measurements are computed for predicted crowns using the same methodology. The intersection loss employs Mean Square Error (MSE) between ground truth and predicted measurements:

$$L_{intersection} = \text{MSE}(d_L^{pred}, d_L^{gt}) + \text{MSE}(d_R^{pred}, d_R^{gt}) + \alpha[\text{MSE}(A_L^{pred}, A_L^{gt}) + \text{MSE}(A_R^{pred}, A_R^{gt})] \quad (7.3)$$

where d_L, d_R represent left and right interpenetration distances, A_L, A_R represent left and right intersection areas, and $\alpha = 0.01$ is an area weighting factor. The computational procedure is detailed in Algorithm 2.

Algorithm 2 Intersection Loss Function.

1. **Input:** Predicted crown points P_{pred}, crown normals N_{crown}, context points $P_{context}$, context normals $N_{context}$, pre-computed context PSR field $\psi_{context}$, ground truth (GT) intersection measurements $(d_L^{gt}, d_R^{gt}, A_L^{gt}, A_R^{gt})$
2. **Output:** Loss $L_{intersection}$
3. **for** each case i in batch **do**
4. Compute normalization bounds from crown and context points
5. Normalize crown points: $P_{norm} = \frac{P_{pred} - \text{min bound}}{\text{max bound} - \text{min bound} + 1}$
6. Compute crown PSR field: $\psi_{crown} = \text{DPSR}(P_{norm}, N_{crown})$
7. Build the 3D surface reconstruction grid
8. Find intersection mask: $\mathcal{M} = (\psi_{crown} \leq \tau) \wedge (\psi_{context} \leq \tau)$
9. Extract intersection points and denormalize to original coordinates
10. Separate intersection points into left and right regions using PCA along context OBB axis
11. Compute oriented bounding boxes for left and right intersection regions
12. Extract interpenetration distances: d_L^{pred}, d_R^{pred} (smallest OBB extents)
13. Extract intersection areas: A_L^{pred}, A_R^{pred} (product of two larger OBB extents)
14. Compute MSE losses for each measurement component
15. **end for**
16. Return weighted sum: $L_{intersection} = \text{MSE}(d_L) + \text{MSE}(d_R) + 0.01[\text{MSE}(A_L) + \text{MSE}(A_R)]$

Each spatial constraint loss addresses specific anatomical requirements in crown generation. The antagonist interaction loss ensures proper occlusal relationships with opposing dentition, while the intersection loss prevents crown penetration into adjacent structures. Combined with margin line input integration, these individual components demonstrate the potential for addressing fundamental limitations in automated dental crown generation, establishing building blocks for clinically viable crown design systems.

7.5 Experimental Setup

Implementation : Our framework is implemented in PyTorch using the AdamW optimizer (learning rate: 5e-4, batch size: 4) on an NVIDIA A100 GPU. Training requires 250 epochs (22 hours) with intersection loss activated after epoch 170 to allow basic geometry establishment before spatial constraint enforcement.

Training strategy : Due to computational constraints, we evaluate our contributions using two training configurations: (1) baseline with margin line input and antagonist interaction loss, and (2) baseline with intersection loss. All loss components use equal weighting within each configuration.

Parameters : Key parameters include PSR grid resolution (128^3), intersection threshold ($\tau = 0.1$), and area weighting factor ($\alpha = 0.01$).

Evaluation metrics : We assess performance using Chamfer distances (CD_{L1}, CD_{L2}) between predicted and ground truth point clouds, and Mean Square Error (MSE) on indicator grids for mesh reconstruction accuracy. Spatial constraints are validated through distance and area MSE metrics corresponding to the terms in Equation 7.3.

7.6 Performance evaluation

Our evaluation demonstrates the effectiveness of the enhanced framework through comprehensive analysis across three key areas: systematic performance improvements, margin line integration analysis, and spatial constraint loss validation.

7.6.1 Overall framework performance

We evaluate our framework through systematic component analysis, demonstrating consistent improvements from our initial point-based approaches to the current enhanced framework. Our evaluation demonstrates the geometric accuracy improvements achieved through each contribution.

Table 7.1 presents our research progression in geometric accuracy, highlighting how each methodological advancement builds upon previous work while addressing specific limitations. The evolution from point-based to mesh generation approaches demonstrates consistent improvement across all evaluation metrics.

Table 7.1 Geometric accuracy progression and margin line integration impact. Evaluation metrics are CD_{L_1}, CD_{L_2} (multiplied by 1000; lower is better), L1 distance, and MSE (lower is better).

Method	CD-L1 (↓)	CD-L2 (↓)	L1 (↓)	MSE (↓)
Pointr + Margin Line [86]	67.01	19.32	0.0151	-
DMC [92]	62.41	17.06	0.0127	0.0028
Personalized Framework [15]	61.34	15.35	0.0103	0.0025
Personalized Framework + Margin as Input	**42.01**	**5.60**	**0.0078**	**0.0015**

The most substantial advancement comes from reformulating margin line data as direct network input rather than loss supervision, yielding a 31.5% improvement in CD-L1 (61.34 $\rightarrow$ 42.01). This architectural decision fundamentally changes how the transformer processes critical boundary constraints, enabling immediate geometric constraint incorporation during feature extraction rather than relying solely on gradient-based supervision.

Building upon this geometric foundation, we evaluate our spatial constraint innovations through focused experimental approaches that isolate individual component effects due to computational constraints. Table 7.2 presents systematic analysis of our spatial constraint losses, demonstrating their impact on geometric accuracy when applied to the Personalized Framework (CD-L1: 61.34) as our baseline.

Table 7.2 Spatial constraint modeling evaluation. Individual assessment of intersection and antagonist interaction losses compared to baseline framework. Evaluation metrics are CD_{L_1}, CD_{L_2} (multiplied by 1000; lower is better), L1 distance, and MSE (lower is better).

Method	CD-L1 (↓)	CD-L2 (↓)	L1 (↓)	MSE (↓)
Baseline + Intersection Loss	58.80	14.41	0.0095	0.0023
Baseline + Antagonist Loss	**41.18**	**5.40**	**0.0070**	**0.0011**

The intersection loss provides measurable enhancement (61.34 → 58.80 CD-L1), demonstrating its effectiveness in preventing crown penetration into adjacent teeth. The antagonist interaction loss achieves more substantial improvement (61.34 → 41.18 CD-L1) by optimizing occlusal contact relationships with opposing dentition.

Our optimal achievable combination integrates the most effective components from both morphological and spatial constraint enhancements. Table 7.3 presents our best performance alongside comparisons with state-of-the-art methods.

Table 7.3 Final framework performance and comparison with state-of-the-art methods. All metrics: lower is better.

Method	CD-L1 (↓)	CD-L2 (↓)	L1 (↓)	MSE (↓)
ToothCR [84]	64.01	17.41	0.0120	0.0024
DCrownFormer [141]	59.38	15.22	0.0109	0.0022
Enhanced Framework (Margin + Antagonist)	**38.04**	**4.67**	**0.0068**	**0.0010**

Our enhanced framework significantly outperforms both state-of-the-art methods across all metrics, achieving 35.9% and 40.6% improvements in CD-L1 over DCrownFormer and ToothCR respectively. CD-L2 shows 69.3% and 73.2% improvements, while MSE demonstrates 54.5% and 58.3% improvements. The complete progression from our initial point-based approach (CD-L1: 67.01) to our enhanced framework (CD-L1: 38.04) represents a 43.3% improvement, validating the effectiveness of systematic component integration and spatial constraint modeling for automated dental crown generation.

7.6.2 Margin line integration analysis

The margin line represents the critical interface between the crown and prepared tooth, directly impacting restoration fit, durability, and patient comfort. To evaluate the impact of margin line integration, we extract margin lines from predicted crown meshes under two conditions: with margin line input (where margin line data was provided to the model alongside master and antagonist arches) and without margin line input (where only master and antagonist arch data was provided). The margin line extraction follows the same procedure used for ground truth data, ensuring consistent comparison methodology.

Distance analysis employs Chamfer distance computation between extracted predicted margin lines and ground truth margin lines. For each test case, we construct k-d trees for both ground truth and predicted margin line points, computing bidirectional nearest neighbor distances. The comprehensive distance metrics include maximum, minimum, average, and standard deviation values, providing complete characterization of margin line accuracy across

the entire crown boundary.

Figure 7.3(a) demonstrates the substantial impact of margin line integration across all test cases. The case-by-case analysis reveals that margin line input consistently reduces prediction errors, with particularly notable improvements in maximum distance errors. Without margin input, the framework exhibits significant variability in performance, with maximum distances reaching up to 7.326 mm and several cases showing substantial outliers. In contrast, the framework with margin input maintains more consistent performance with maximum distances typically below 1.161 mm.

The distribution analysis (Figure 7.3(b)) further confirms the effectiveness of margin line integration, showing a clear shift toward lower error values with reduced variance. This improvement translates directly to clinical benefits: better crown fit reduces the need for manual adjustments, improves restoration longevity, and enhances patient comfort through more precise margin adaptation.

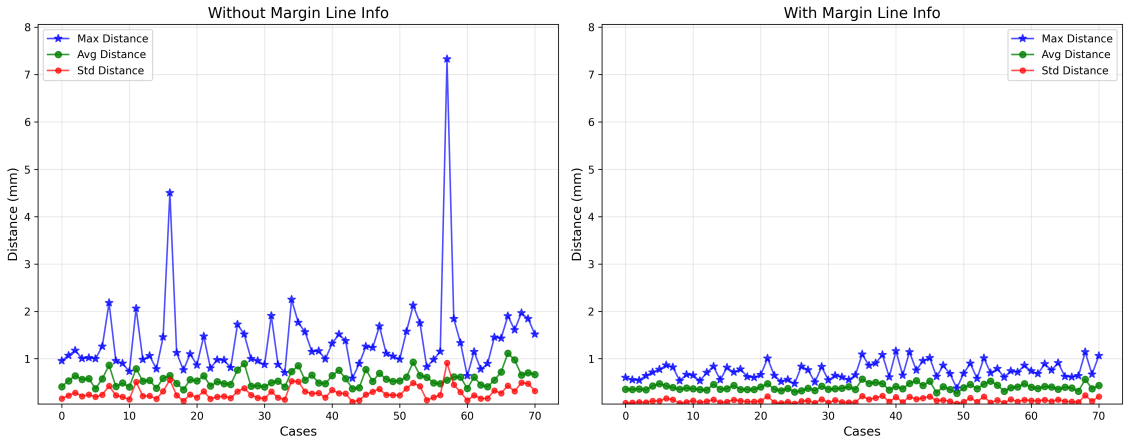
The quantitative analysis in Table 7.4 reveals significant improvements across all metrics. The average distance error decreases from 0.5759 mm to 0.3964 mm (31.2% improvement), while the standard deviation reduces dramatically from 0.2770 to 0.1152 mm (58.4% reduction), indicating substantially improved consistency. Most importantly, the maximum distance error decreases by 46.3%, demonstrating enhanced precision in critical boundary regions where crown-tooth interface accuracy is paramount.

Table 7.4 Impact of margin line integration on crown generation accuracy. All distance measurements in millimeters (mm).

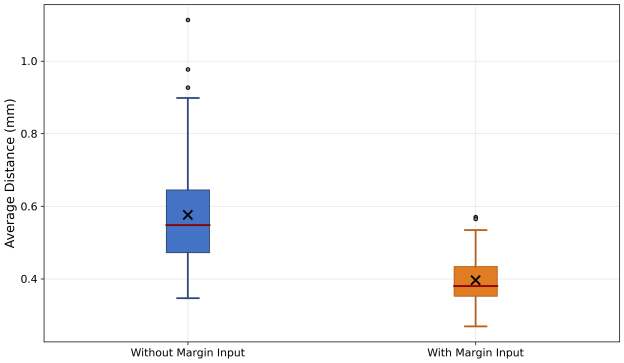
Metrics	Without Margin Input	With Margin Input
Max Distance	1.3729	**0.7376**
Min Distance	0.2527	**0.2368**
Average Distance	0.5759	**0.3964**
Std Distance	0.2770	**0.1152**
Median Distance	0.5121	**0.3716**

To further illustrate the qualitative impact of margin line integration, Figure 7.4 presents visual comparisons of crown mesh predictions in three representative cases.

These quantitative and qualitative results demonstrate that direct margin line integration significantly enhances crown generation accuracy and consistency, providing the boundary precision essential for clinical deployment.

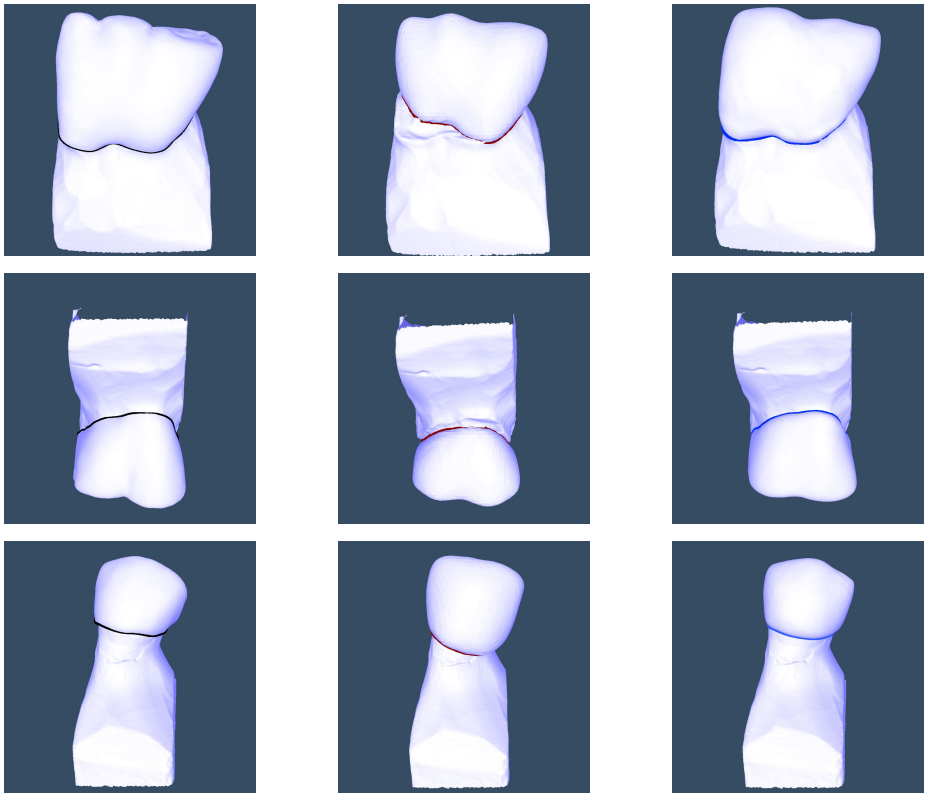


(a) Case-by-case comparison with and without margin line input



(b) Distribution comparison

Figure 7.3 Impact of margin line integration on crown generation accuracy. (a) Case-by-case comparison showing consistent improvement with margin line input across all test cases, and (b) distribution analysis demonstrating improved consistency and reduced error variance.



(a) Ground Truth (b) Without Margin (c) With Margin

Figure 7.4 Qualitative comparison of crown mesh predictions across three representative cases. Each row shows ground truth, prediction without margin line input, and prediction with margin line input. The visual comparison demonstrates improved crown morphology, boundary accuracy, and overall geometric fidelity when margin line information is included in the model's input specification.

7.6.3 Spatial constraint loss evaluation

Antagonist interaction loss

Geometric performance evaluation The antagonist interaction loss ensures optimal occlusal contact relationships between the generated crown and opposing dentition by enforcing consistent spatial proximity patterns. To evaluate occlusal accuracy, we computed point-wise distances between predicted crown point clouds and corresponding antagonist mesh surfaces using Open3D, uniformly sampling 10,000 points on each antagonist mesh. This evaluation was conducted for crowns generated with and without the proposed interaction loss, as well as for ground truth crowns.

Figure 7.5a illustrates the distribution of average crown-to-antagonist distances across three

groups. The model trained without the interaction constraint predicts crowns with larger occlusal distances compared to ground truth. In contrast, predictions with interaction loss exhibit improved antagonist alignment, resulting in a 9.51% average reduction in occlusal distance and enhanced geometric stability.

The per-case analysis in Figure 7.5b demonstrates consistent improvements across the dataset, with predictions using interaction loss consistently closer to ground truth than baseline predictions. Figure 7.5c shows the top ten cases with improvements exceeding 15%, with the best case achieving over 21% reduction.

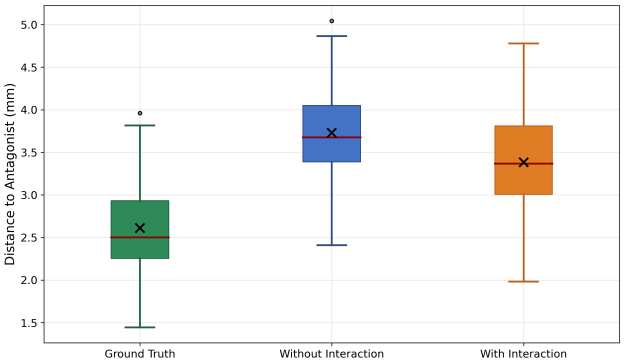
To evaluate clinical relevance, we performed occlusal contact classification using implicit distance field computation. Following established clinical protocols [155], we used a three-class system: Class 1 (+0.02 to 0 mm) indicates planned contacts, Class 2 (0 to -0.02 mm) represents light premature contact, and Class 3 ($<$ -0.02 mm) signifies heavy premature contact requiring adjustment.

Analysis across the complete dataset reveals substantial improvements with interaction loss integration. Baseline predictions showed concerning outcomes with the majority falling into Class 3 (heavy premature contact), while interaction loss integration demonstrated measurable improvement with increased Class 1 contacts and reduced Class 3 cases. Figure 7.6 illustrates this impact, showing that baseline predictions exhibit extensive premature contact areas, while predictions with antagonist interaction loss demonstrate substantially reduced premature contact zones that more closely approximate ground truth distributions.

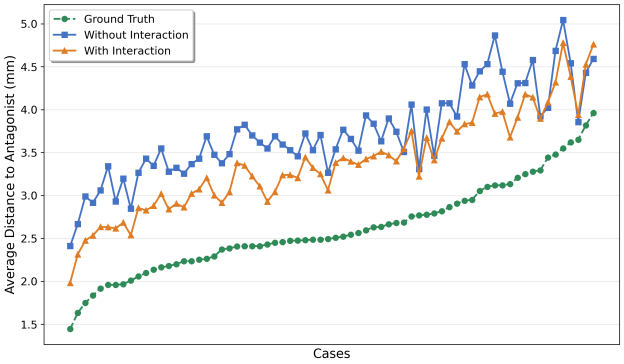
These findings demonstrate that the antagonist interaction loss provides both geometric accuracy improvements and clinically relevant contact optimization, reducing the likelihood of premature contacts requiring post-generation adjustments in clinical practice.

Functional performance evaluation To address the clinical relevance of our spatial constraints beyond geometric accuracy, we employ the functional evaluation methodology established by Hwang et al. [1]. This approach evaluates crown functionality through penetration analysis and contact point assessment, measuring clinical viability rather than geometric similarity. While the original work used 2D depth images with gap distance reconstruction, we apply these penetration analysis principles to our 3D mesh data by computing crown-to-antagonist distances using PyVista's implicit distance calculation between crown point clouds and antagonist mesh surfaces.

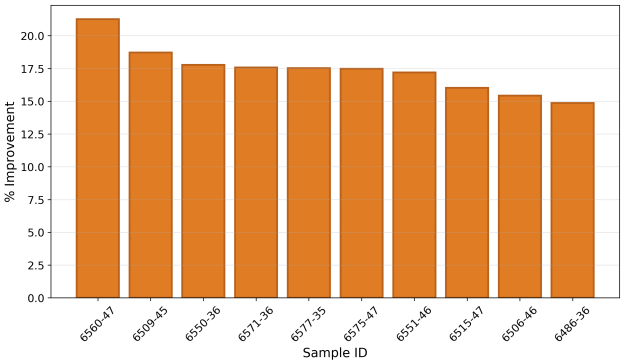
We implement penetration analysis by identifying crown points with negative distances relative to antagonist surfaces, indicating functional failures requiring clinical intervention. This follows the established principle that crowns penetrating opposing teeth represent clinical



(a) Distribution of average distances

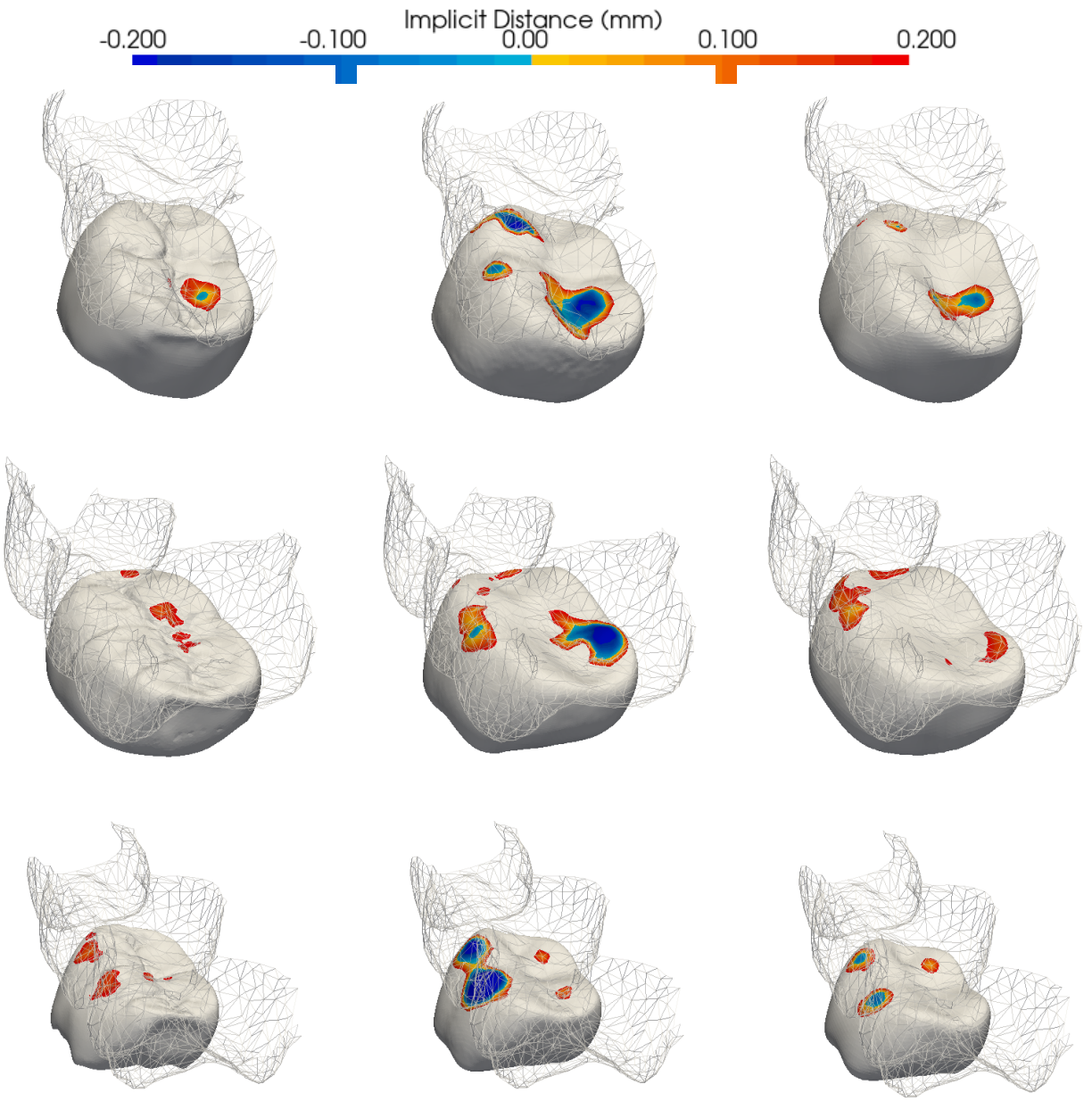


(b) Per-case distance comparison



(c) Top 10 improvement cases

Figure 7.5 Antagonist interaction loss evaluation showing improved occlusal relationships. (a) Distribution comparison across all test cases, (b) detailed per-case analysis, and (c) cases with highest improvement percentages.



(a) Ground Truth (b) Without Interaction (c) With Interaction

Figure 7.6 Occlusal contact analysis across three representative cases showing implicit distance visualization between crown and antagonist surfaces. Color deviation maps were generated, ranging from red (heavy premature contact, Class 3) to blue (loss of contact, Class 1) for convenient visual inspection. The color bar indicates the full visualization range (-0.2 to $+0.2$ mm), while classification is based on clinically relevant thresholds (± 0.02 mm). Blue regions indicate loss of contact (Class 1), green/yellow regions show light premature contact (Class 2), and red regions indicate heavy premature contact requiring adjustment (Class 3). Exact numerical values were also examined for quantitative analysis.

unusability. For contact assessment, we quantify crown points within 0.5mm of antagonist surfaces, representing functional contact zones based on established dental proximity thresholds from clinical practice.

We analyze three conditions: ground truth crowns, predictions without antagonist interaction loss (Without Interaction), and predictions with antagonist interaction loss (With Interaction). Each crown is evaluated against its corresponding antagonist tooth using the binary failure criterion that any penetrating crown is clinically unusable.

Table 7.5 Functional performance evaluation using Hwang et al. [1] methodology adapted for 3D crown generation. Contact area represents percentage of crown points within 0.5mm functional threshold.

Method	Penetration Rate (%)	Avg Contact Points	Contact Area (%)
Ground Truth	25.0	1,357	14.24
Without Interaction	62.5	2,980	12.81
With Interaction	37.5	630	10.93

The functional evaluation reveals significant challenges in automated crown generation that geometric metrics alone cannot capture. Ground truth crowns demonstrate a 25.0% penetration rate, reflecting the inherent difficulty in achieving optimal crown-antagonist relationships even in professionally designed restorations. Our baseline approach (Without Interaction) exhibits 62.5% penetration rate, indicating substantial functional deficiencies despite geometric accuracy. The enhanced framework with antagonist interaction loss achieves 37.5% penetration rate, representing a 40% reduction in functional failures compared to the baseline.

Contact point analysis demonstrates the trade-offs inherent in functional optimization. Ground truth crowns maintain 1,357 average contact points (14.24% contact area), establishing the reference standard for functional contact patterns. The baseline method generates excessive contact with 2,980 points (12.81% area), contributing to higher penetration rates. Our enhanced framework produces more conservative geometries with 630 contact points (10.93% area), prioritizing penetration avoidance while maintaining functional contact zones.

This functional evaluation demonstrates the necessity of clinical validation beyond geometric accuracy metrics. While our geometric improvements show enhanced spatial relationships (9.51% distance reduction), the functional analysis reveals that clinical viability requires explicit constraint modeling. The antagonist interaction loss creates more conservative crown geometries that reduce penetration risk, achieving a 40% reduction in functional failures compared to baseline methods. The trade-off between contact area and penetration avoidance is evident in our results, where reduced contact points (630 vs 2,980) correlate with lower penetration rates (37.5% vs 62.5%).

Direct comparison with Hwang et al. [1] is not feasible since they evaluated hard test cases with 100% human failure rates on single tooth position (36), while our evaluation covers standard clinical cases across multiple tooth positions. The substantial penetration reduction in our study validates that spatial constraint integration provides measurable functional benefits, establishing a foundation for clinically viable automated crown design systems.

Intersection loss analysis

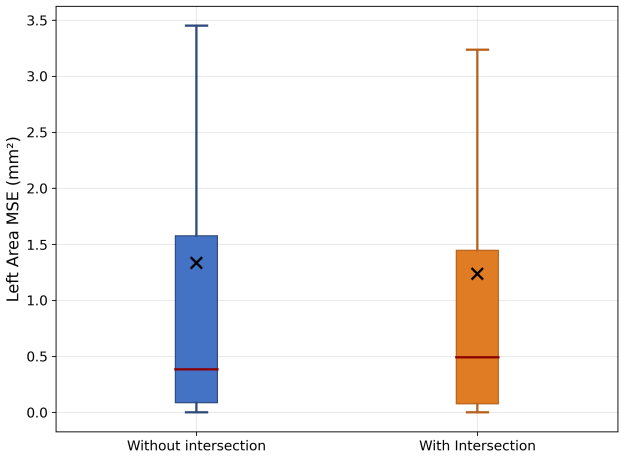
The intersection loss function addresses the critical challenge of crown penetration into adjacent dental structures by enforcing consistent spatial relationships between predicted crowns and the surrounding master arch context. Unlike the antagonist interaction loss, the goal of the intersection loss function is not to minimize the interpenetration between the shell and its context. In the case of adjacent intersection, loss of contact would be considered a failure, as adjacent contact is required to prevent various complications [142]. The intersection loss therefore aims to ensure an appropriate contact between the crown and the adjacent teeth, similar to the one observed for the ground truth shell. We evaluate three key metrics corresponding to Equation 7.3: left area MSE and right area MSE quantify intersection area estimation precision on respective sides, while total intersection MSE provides overall spatial constraint effectiveness. All MSE values are computed in area units (mm^2).

Figure 7.7 compares baseline training (250 epochs without intersection loss) and training with intersection loss activation (after epoch 170). Both left and right area MSE demonstrate reduced median values when the constraint is active, while the total intersection MSE validates the effectiveness of the weighted combination approach.

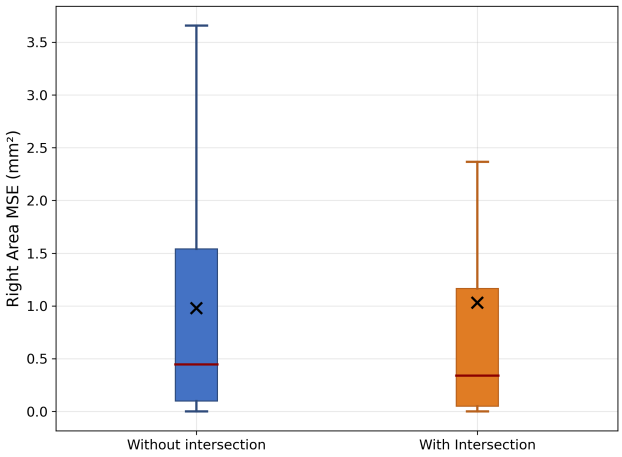
To assess constraint effectiveness, we analyze spatial relationships between crowns and adjacent teeth using implicit distance computation with a three-class proximal contact classification: Class 1 (> 0 mm) indicates loss of contact, Class 2 (0 to -0.02 mm) represents optimal contact, and Class 3 ($<$ -0.02 mm) indicates heavy contact requiring clinical adjustment.

Figure 7.8 reveals that ground truth crowns exhibit complex, spatially varied contact patterns reflecting natural anatomical relationships. Baseline predictions show simplified contact regions, while crowns generated with active constraints demonstrate more controlled patterns that better approximate ground truth complexity.

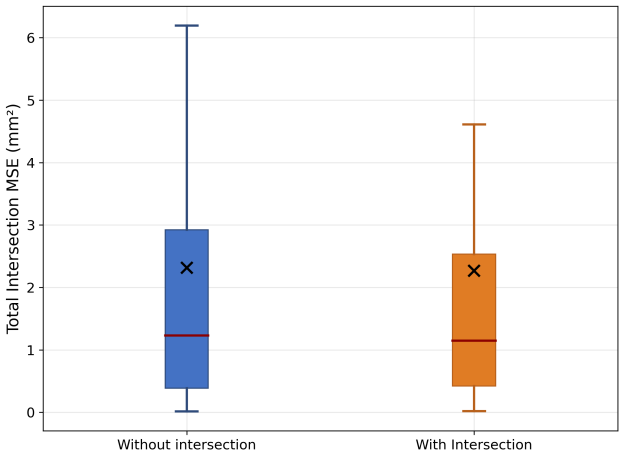
Distance heat mapping in Figure 7.9 shows that ground truth crowns exhibit complex distance patterns with ranges varying from 1.08–4.29 mm, 1.87–7.42 mm, and 1.39–5.39 mm across the three cases. Baseline predictions demonstrate uniform spacing patterns, while crowns generated with active constraints exhibit distance patterns that better approximate



(a) Left area MSE



(b) Right area MSE



(c) Total intersection MSE

Figure 7.7 Intersection loss component analysis showing left and right area MSE components and their weighted combination. (a) Left-side interpenetration area analysis, (b) right-side interpenetration area analysis, and (c) overall intersection loss combining area components with $\alpha = 0.01$ weighting for area terms.

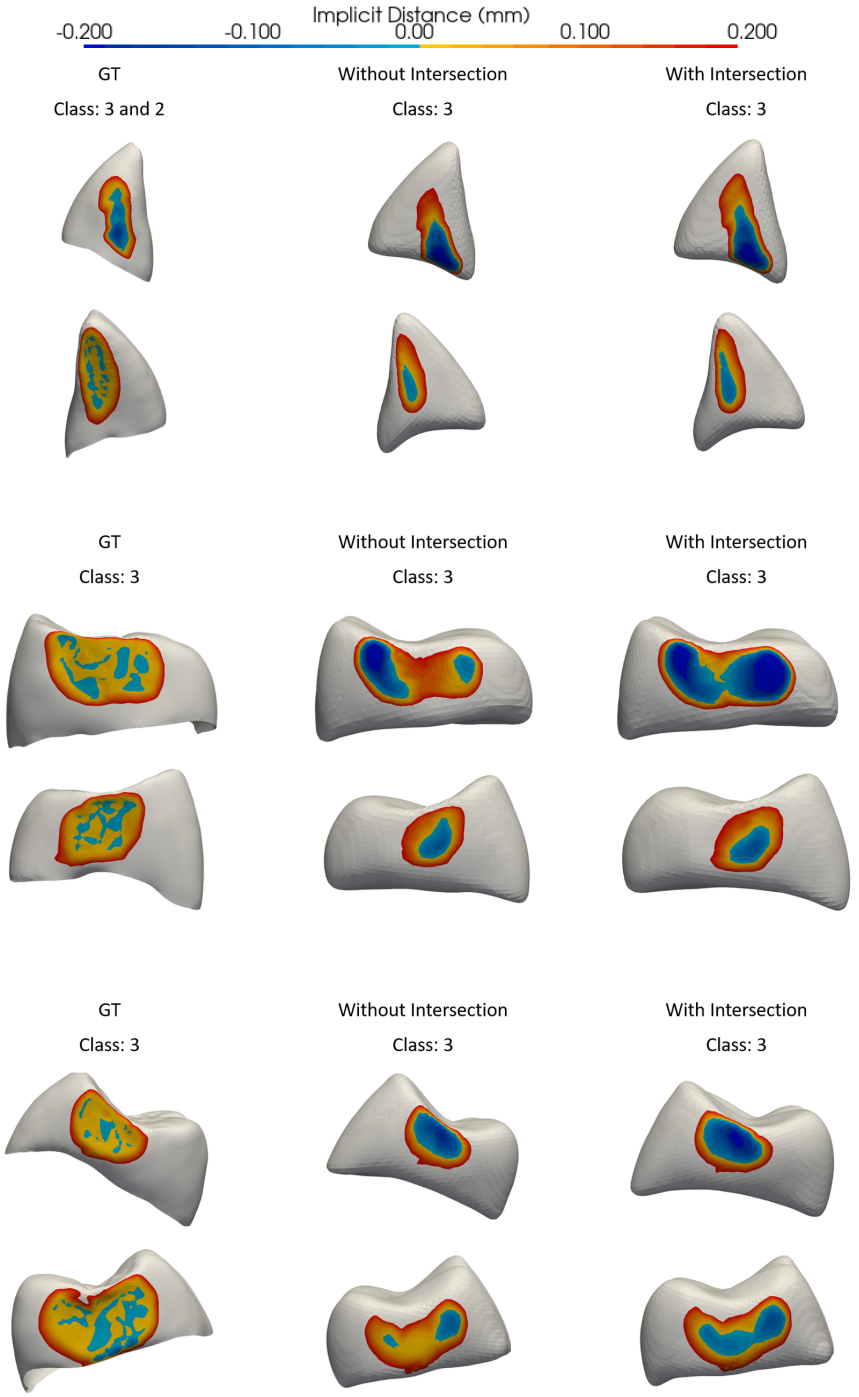


Figure 7.8 Implicit distance analysis between crown and master arch context across three representative cases. Each row shows ground truth, prediction without constraint, and prediction with constraint active. The color map indicates contact intensity classification with visualization range (-0.2 to $+0.2$ mm): Class 1 (> 0 mm) shows loss of contact, Class 2 (0 to -0.02 mm) represents contact as planned, and Class 3 ($<$ -0.02 mm) indicates heavy intensity contact.

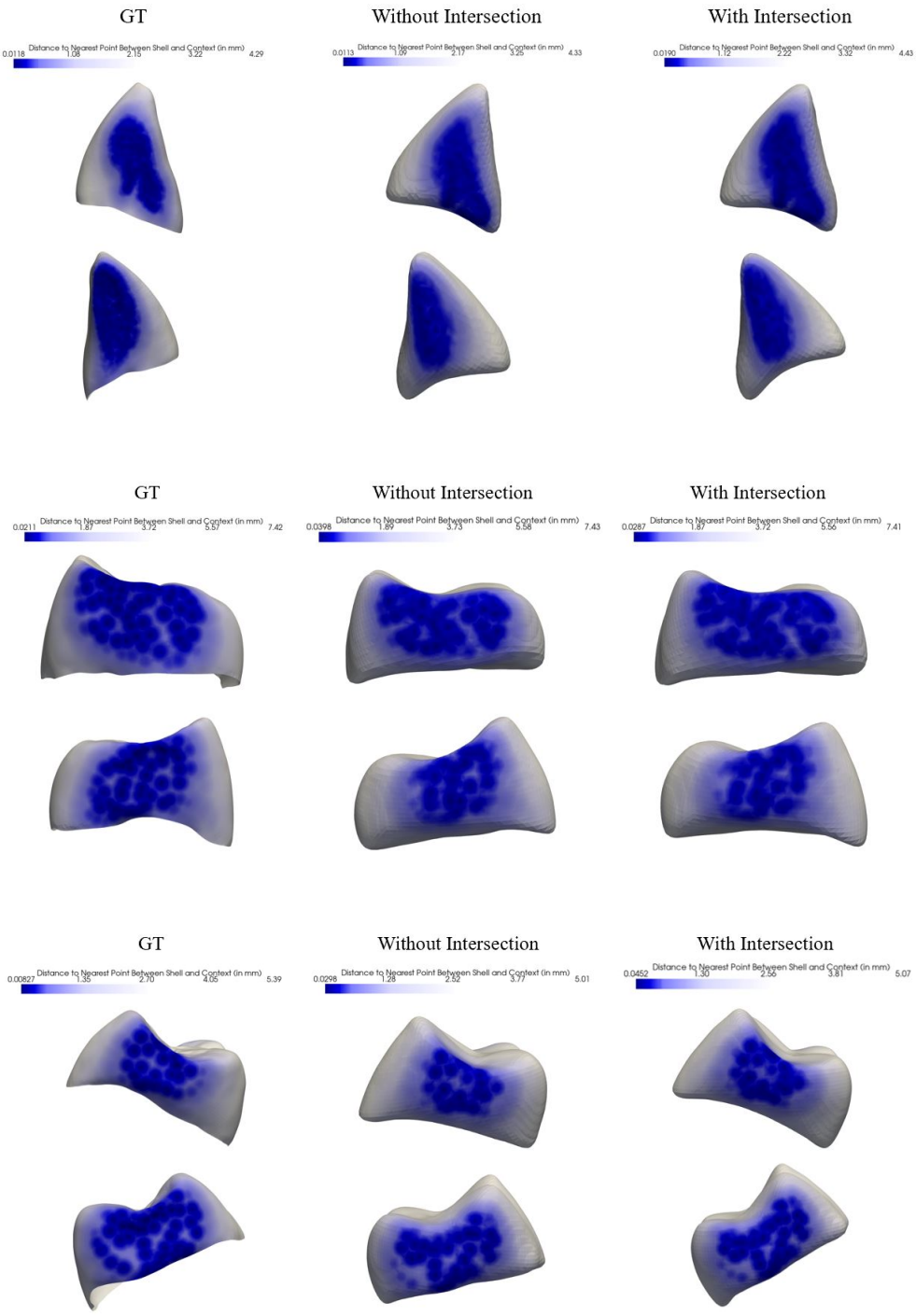


Figure 7.9 Absolute distance heat map analysis between crown and master arch context. Each row shows ground truth, prediction without constraint, and prediction with constraint active. Darker blue regions indicate closer proximity between the crown and adjacent tooth surfaces, with distance ranges varying across cases to reflect natural anatomical variation.

ground truth complexity, showing improved spatial relationship control with adjacent dental structures.

The intersection loss function effectively establishes anatomically appropriate spatial relationships between predicted crowns and adjacent dental structures. The quantitative MSE improvements combined with qualitative enhancements demonstrate the clinical viability of the proposed constraint-based learning approach, with consistent reduction in intersection area errors and improved contact classification patterns supporting the effectiveness of integrating spatial constraints into the training process.

7.7 Discussion

Our dental crown generation framework demonstrates significant advancement through the systematic integration of margin line input and spatial constraint losses. We achieve substantial improvements over state-of-the-art methods, validating our hypothesis that explicit spatial relationship modeling is essential for automated crown generation approaching clinical standards. These improvements reflect limitations in existing approaches that lack explicit constraint mechanisms and rely solely on geometric reconstruction without considering critical boundary relationships or inter-tooth dependencies.

The most transformative contribution emerges from integrating margin line data as direct network input rather than loss supervision. This architectural decision fundamentally changes how the transformer processes critical boundary constraints. While loss-based supervision provides only indirect guidance through gradient backpropagation, input integration enables immediate geometric constraint incorporation during feature extraction and spatial attention mechanisms. This enhanced attention to boundary regions translates to substantially improved geometric precision and reliability across diverse anatomical configurations, with significant reductions in maximum distance errors (46.3% reduction) and variability (58.4% reduction) at the crown-tooth interface.

Building upon this foundation, our constraint losses address the challenge of establishing proper 3D relationships within the dental arch. The antagonist interaction loss corrects spatial positioning relative to opposing dentition through Chamfer distance comparison, achieving improved occlusal distance alignment (9.51% geometric improvement). However, our functional evaluation reveals that geometric improvements alone are insufficient for clinical deployment. The functional analysis demonstrates that the antagonist interaction loss provides clinically meaningful benefits, reducing penetration rates from 62.5% to 37.5%—a 40% reduction in functional failures. This substantial improvement validates that spatial

constraints target clinically relevant relationships rather than merely optimizing geometric similarity metrics.

The functional evaluation illuminates the fundamental challenge in automated crown generation: balancing geometric accuracy with clinical viability. While our enhanced framework achieves superior geometric performance (35.9-40.6% improvements over state-of-the-art methods), the functional analysis reveals that even ground truth crowns demonstrate 25% penetration rates, highlighting the inherent complexity of crown-antagonist relationships. The trade-off between contact area and penetration avoidance is evident in our results, where the antagonist interaction loss creates more conservative geometries (630 vs 2,980 contact points) that prioritize functional viability over extensive occlusal contact.

The intersection loss function prevents crown penetration into adjacent structures through PSR field-based modeling. Critical design decisions include balancing PSR grid resolution between precision and computational tractability, and selecting intersection thresholds that avoid false positives while ensuring reliable detection. The progressive activation strategy after initial training epochs reflects a key insight: early predicted geometries lack sufficient tooth-like characteristics for meaningful constraint optimization, making delayed activation essential for stable learning.

The choice to employ different methodologies for adjacent versus antagonist interactions reflects distinct geometric characteristics of these relationships. Adjacent intersection analysis benefits from oriented bounding box computation providing interpretable measurements and consistent results less sensitive to point sampling randomness. In contrast, antagonist interactions involve multiple dispersed contact areas requiring Chamfer distance metrics to capture complex distributed contact patterns. This dual-methodology approach optimizes each constraint type for its specific geometric context while demonstrating that both geometric and functional improvements can be achieved simultaneously.

Several implementation considerations shaped our methodology. Loss function weights were balanced to achieve similar scales across components, ensuring stable training convergence. The PSR-based intersection approach demonstrates potential for broader constraint applications beyond dental crown generation. While computational constraints required separate evaluation of loss components, results show that margin line integration combined with spatial constraints achieves performance that addresses both geometric accuracy and functional requirements.

The systematic integration of these innovations establishes a foundation for clinically viable automated crown generation. The framework addresses limitations in existing methods by explicitly modeling relationships during generation rather than as post-processing correc-

tions. The functional evaluation methodology demonstrates the necessity of clinical validation beyond geometric accuracy metrics, establishing that automated crown design must be evaluated for clinical viability rather than solely geometric similarity. While current predictions exhibit characteristic neural network smoothness lacking fine-scale morphological details, the substantial improvements in both geometric precision and functional performance establish a foundation for automated dental crown design systems approaching clinical deployment standards.

Future work should focus on addressing the remaining functional challenges while maintaining geometric accuracy. The trade-off between penetration avoidance and adequate contact distribution suggests that multi-objective optimization approaches may be necessary to balance these competing functional requirements. Additionally, computational efficiency improvements through octree-based spatial data structures could reduce processing time while maintaining accuracy. Integration of fine-scale morphological details through alternative surface refinement approaches could enhance clinical acceptance while preserving the functional benefits demonstrated by our spatial constraint modeling.

7.8 Conclusion and Future Work

This work demonstrates that integrating spatial constraints significantly improves automated dental crown generation. Our key innovations, margin line input integration and spatial constraint losses, successfully address major limitations in existing approaches. The results show substantial improvements in geometric accuracy, validating that explicit spatial modeling is essential for clinical viability.

The framework achieves both precise geometric fit and proper spatial relationships while maintaining computational efficiency. These improvements, combined with reduced manual adjustment requirements, demonstrate practical viability for deployment in dental laboratories.

Beyond technical contributions, this approach could transform dental restoration workflows by reducing manual design time, improving consistency across technicians, and enabling more accessible high-quality dental care. The spatial constraint modeling principles extend to broader medical device design and anatomical modeling applications.

Future work will focus on computational efficiency improvements through octree-based spatial data structures [145], which can reduce processing time while maintaining precision. Additionally, investigating methods to incorporate fine-scale morphological details such as grooves and cusps remains important. This could involve learning-based approaches that

capture surface patterns from training data or analytical methods such as spherical harmonics [130] that provide precise morphological control through mathematical formulations. These developments would enhance surface realism and advance geometric modeling in computer vision applications.

Disclosures Golriz Hosseinimanesh reports financial support from Intellident Dentaire Inc., iMD Research Inc., NSERC, and MEDTEQ. All other 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.

Code and Data Availability The clinical dental data are confidential and cannot be publicly shared. The implementation code is not currently in a public repository but can be requested from the corresponding author.

Acknowledgments We acknowledge the help and funding received from Intellident Dentaire Inc., iMD Research Inc., the Natural Science and Engineering Research Council of Canada (NSERC), and MEDTEQ. A. Alsheghri acknowledges the support of King Fahd University of Petroleum and Minerals (KFUPM) [Ref: EC241009]. The authors would also like to thank the Digital Research Alliance of Canada for their continued support of our research and access to computational resources.The work presented has not been published or presented elsewhere.

Golriz Hosseinimanesh earned her bachelor's degree in computer engineering in 2015 in Iran, followed by a master's degree in the same field in 2018. She is currently a PhD candidate in computer engineering at Polytechnique Montréal, Canada. Her research focuses on deep learning, artificial intelligence, data science, and applied machine learning in the medical domain.

Farida Cheriet received the Ph.D. degree in computer science from the University of Montreal, QC, Canada, in 1996. She held a postdoctoral position at the Biomedical Engineering Institute, École Polytechnique de Montréal, from 1997 to 1999. Since 1999, she has been appointed in the Department of Computer and Software Engineering, École Polytechnique de Montréal, where she is currently a full Professor. She is an IEEE Senior member, a member of the research center of Sainte Justine hospital in Montreal and a member of Ordre des Ingénieurs du Québec. She is the director of two laboratories: LIV4D (Laboratory of Imaging and Vision in 4 Dimensions) at Polytechnique and LAVIANI (Laboratory of Non-Invasive Active Vision) at Sainte Justine research center. She has made contributions in

medical imaging and has been collaborating with clinicians at Sainte-Justine Hospital since her appointment at Polytechnique in 1999. She also has recently established a partnership with Mila Institute and ophthalmologists from Maisonneuve Rosemont hospital to develop artificial intelligence platforms for diagnosis of ocular diseases. Her research interests include three-dimensional (3-D) reconstruction of bone structures from x-rays, calibration of x-ray imaging systems, non-invasive 3-D modeling of scoliosis deformities, 3-D augmented reality systems for minimally invasive surgery, retinal image analysis, 3-D reconstruction and visualization of vascular structures from multimodal images, and spatiotemporal registration of medical images. Her research activities have generated three patents, four licenses, and more than 150 journal papers and 200 full manuscripts and abstracts in conference proceedings.

François Guibault is a full professor in the CSED at Polytechnique since 1999. His research and expertise include the development of methods for 3D surface and volume representation of computational domains, spatial discretization, error estimation and control, and shape optimization for simulation-based engineering processes. His research in geometry and meshing, combined with his work in artificial intelligence, has had a significant impact in the healthcare sector, firstly in the modeling and simulation in orthopedics and more importantly in the modeling and generation of dental arches and crowns, in partnership with Intellident Dentaire Inc. and iMD Research Inc.

Ammar Alsheghri is an Assistant Professor in the Mechanical Engineering Department at King Fahd University of Petroleum and Minerals (KFUPM) since September 2023, and serves as a visiting professor at Polytechnique Montréal during summers 2024-2025. He earned his PhD in Materials Engineering from McGill University in 2020 with collaborative research in the Faculty of Dentistry, and previously worked as a technology developer at Comet Technologies Canada Inc. His research focuses on artificial intelligence applications, bioinformatics, and bioengineering, with over 20 peer-reviewed publications and 2 U.S. patents.

Victoria-Mae Carrière is currently an undergraduate student in software engineering at Polytechnique Montréal, Canada. Her research interests include applied artificial intelligence and medical image analysis.

Julia Keren BEd, is the president of KerenOr Dental Studio in Montreal, Canada, a leading all-ceramic laboratory specializing in Zirconia. She has focused her career on studying dental technologies, particularly CAD/CAM systems. Alongside her husband Haim, Julia developed protocols for the digital manufacturing and design of Zirconia restorations. She also contributes to dental education through lectures, technical courses, and publications, including Quintessence dental textbooks. Recently, Julia co-founded Intellident with Haim

Keren and Nathaniel Lasry. The company is dedicated to developing a cutting-edge 3D AI app that automates the creation of high-quality dental restorations in minutes, using scanned data. Intellident collaborates with École Polytechnique Montréal on this innovative project..

CHAPTER 8 GENERAL DISCUSSION

This thesis addresses the fundamental challenge of automated dental crown generation by developing integrated frameworks that simultaneously achieve morphological precision and functional constraint satisfaction. The research establishes four critical innovations that collectively transform dental crown generation from basic shape completion to comprehensive clinical deployment capability.

The morphological contributions represent fundamental advances in achieving precision required for dental applications. The transformer encoder-decoder architecture brings critical advantages through attention mechanisms particularly suited for complex geometric relationships in dental applications. Self-attention mechanisms allow networks to capture long-range dependencies across entire dental arches, permitting models to understand how crown design decisions in one region affect spatial relationships with distant anatomical structures. Cross-attention mechanisms facilitate effective integration between input dental context and target crown geometry, supporting simultaneous consideration of multiple anatomical constraints during generation. Unlike convolutional approaches that process local neighborhoods, the global receptive field of transformers provides comprehensive understanding of spatial relationships essential for crown design within complete oral environments.

Building upon this foundation, the development of end-to-end mesh generation through differentiable Poisson surface reconstruction represents a breakthrough in eliminating error accumulation inherent in traditional multi-stage pipelines. The central innovation lies in viewing 3D crown shapes as implicit indicator functions that are one inside the object and zero outside, rather than as explicit meshes or discrete voxels. By making classical Poisson surface reconstruction differentiable and embedding it as a layer within the transformer encoder-decoder architecture, the framework allows networks to learn vector fields whose divergence defines the indicator function. This integration of physics-based partial differential equations with deep learning provides mathematical rigor while facilitating end-to-end training that directly connects raw dental point clouds to high-quality mesh surfaces.

The integration of transformer architectures with explicit margin line constraints constitutes a paradigmatic shift in automated crown generation. Rather than treating boundary fitting as a post-processing concern, this research demonstrates how critical geometric constraints can be embedded directly into attention mechanisms through architectural integration of margin line data as direct network input rather than loss supervision. This approach permits immediate geometric constraint incorporation during feature extraction, ensuring that boundary

precision becomes integral to the generation process rather than an external constraint.

The introduction of contrastive Chamfer distance loss demonstrates how traditional geometric metrics can be enhanced through contrastive learning principles to achieve superior point distribution quality. This innovation addresses the critical challenge of maintaining geometric accuracy while ensuring uniform surface coverage, representing a significant advancement in loss function design for 3D generation applications.

However, morphological precision faces fundamental limitations from the inherent smoothness characteristics of transformer-based outputs, which compromise prediction of intricate surface features such as cusps, grooves, and surface textures essential for proper dental function. The DPSR integration encounters limitations in grid resolution trade-offs, where higher resolutions facilitate better detail capture but substantially increase computational requirements. The method's dependence on well-oriented point clouds with accurate normal vectors creates vulnerability when transformer-generated points lack sufficient detail, while DPSR's assumption of watertight manifold surfaces does not always align with complex dental geometries.

While these morphological advances establish the foundation for precise crown generation, the functional requirements present equally critical challenges. The functional contributions represent a comprehensive integration of spatial relationship modeling in automated crown generation, addressing the critical gap that existing approaches ignore spatial constraints essential for restoration viability. The intersection loss introduces a novel approach to regulate spatial relationships between crowns and adjacent teeth that goes beyond simple overlap penalization. Instead of merely preventing overlap, which could create gaps and compromise required contact, the method uses Poisson Surface Reconstruction fields to measure crown interpenetration with master arches. The innovation computes interpretable geometric quantities including interpenetration distance and contact area using oriented bounding boxes, training models to match these measurements to ground truth crowns and learn anatomically correct proximal contact.

The antagonist interaction loss addresses occlusal relationships through fundamentally different approaches than traditional geometric constraints. Rather than enforcing rigid contact requirements, the method measures similarity between spatial proximity patterns of predicted crowns and ground truth relative to antagonist arches using Chamfer distance-based discrepancy. This allows networks to learn crown morphology adjustments that match functional patterns observed in ground truth crowns, representing a conceptual shift from post-hoc spatial validation to integrated guidance during training.

These functional innovations establish automated crown generation as capable of addressing

comprehensive restoration workflows, moving beyond geometric similarity to ensure proper integration within dental arch contexts. The demonstrated improvements in spatial relationship modeling represent a systematic approach to incorporating functionality requirements in automated dental design systems.

Nevertheless, functional constraint modeling faces limitations in computational complexity, where PSR-based intersection detection requires careful threshold tuning to balance false positive detection against missing actual penetrations. Balancing multiple loss function weights across different spatial relationship types adds optimization complexity requiring careful tuning for different anatomical configurations without systematic approaches for parameter selection.

Beyond these architectural and computational challenges, the framework faces data-related limitations that impact its training and generalization capabilities. The supervised learning approach depends on training data consisting of prepared teeth paired with their corresponding expert-designed crowns. Acquiring such data requires significant time and expertise, as each crown must be professionally designed by dental technicians. This limits the size and diversity of available training datasets, potentially constraining the model's exposure to the full range of morphological variations—such as different cusp patterns and groove formations—present across patient populations.

Additionally, the framework was designed and trained within a specific anatomical context: the prepared tooth with two adjacent teeth and nearest neighbor teeth from the antagonist arch. This contextual definition was deliberately chosen to enable simultaneous optimization of morphological and functional objectives. By including bilateral adjacent teeth, the model can control contact points and prevent crown interference from both sides, while antagonist teeth enable learning of proper occlusal relationships. This context confines the model to generate crowns that fit precisely within the defined spatial constraints without penetrating adjacent or antagonist structures. However, the framework's performance under alternative contextual configurations—such as scenarios with single adjacent teeth or different spatial arrangements—remains unexplored, potentially limiting its applicability across diverse clinical situations.

CHAPTER 9 CONCLUSION

The literature review revealed critical gaps in existing approaches: the absence of integrated frameworks addressing both geometric accuracy and spatial relationships, and the lack of methods incorporating functionality restrictions essential for restoration viability.

Through four progressive innovations detailed in the preceding chapters, this research systematically addressed these limitations. The development progressed from transformer-based point completion with margin line integration, through end-to-end mesh generation via differentiable Poisson surface reconstruction and advanced loss functions, to comprehensive spatial constraint modeling. This progression directly confronted the dual challenge identified in the literature, transforming crown generation from isolated shape completion to comprehensive spatial relationship modeling.

The research validates that integrated frameworks can successfully achieve both morphological precision and functional constraint satisfaction. Our work demonstrates substantial performance improvements over state-of-the-art methods, achieving 35.9-40.6% improvements in geometric accuracy with boundary errors reduced from 1.37 to 0.74 mm and 31.2% improvement in margin line precision. These results confirm that explicit spatial relationship modeling addresses the fundamental challenges preventing practical deployment of automated crown generation systems. The detailed analysis of contributions and limitations appears in the General Discussion chapter.

9.1 Future Work

The inherent smoothness of transformer architectures compromises fine-scale morphological detail capture, representing the most critical limitation requiring immediate attention. Developing hybrid generative approaches that combine transformer-based global shape generation with specialized detail refinement modules would address this challenge. Generative adversarial networks or diffusion models could add surface features such as cusps, grooves, and natural textures while preserving the learned spatial constraints that ensure functional viability. Alternatively, analytical approaches using spherical harmonics or procedural modeling techniques could provide precise control over surface features with reduced computational complexity while maintaining the mathematical smoothness properties essential for manufacturing.

The computational burden of DPSR grid resolution creates memory-performance trade-offs

that limit practical deployment. Implementing octree-based spatial data structures would enable adaptive resolution allocation, concentrating computational resources in geometrically complex regions while using coarser representations in uniform areas. This optimization would maintain reconstruction quality while dramatically reducing memory requirements and processing time.

Threshold sensitivity in PSR-based intersection detection requires careful manual tuning to balance false positive and false negative penetration detection. Replacing threshold-based methods with learning-based approaches that automatically recognize valid spatial relationships would eliminate parameter sensitivity while improving robustness. Alternative geometric representations for penetration detection, such as signed distance fields or collision detection algorithms, could provide more reliable functional constraint enforcement without extensive parameter tuning.

The dependency on supervised learning with expert-designed crown pairs limits dataset size and diversity, constraining exposure to the full spectrum of morphological variations across patient populations. Developing semi-supervised or unsupervised pattern learning approaches that extract surface feature characteristics from partially labeled or unlabeled datasets would reduce this dependency while capturing natural anatomical variations in cusp patterns and groove formations underrepresented in curated datasets.

While the framework was trained within a specific anatomical context—prepared tooth with two bilateral adjacent teeth and nearest antagonist neighbors—the modular architecture is inherently generalizable to alternative spatial configurations, including scenarios with single adjacent teeth or varying antagonist arrangements.

The spatial constraint modeling principles established in this work extend naturally to other dental prosthetics including implants, bridges, and complete denture design, where similar requirements for precise fitting and morphological accuracy exist. Extending to complete restoration planning involving multiple crowns would enable multi-crown optimization strategies ensuring proper occlusal relationships across comprehensive rehabilitations rather than isolated restorations. Finally, investigating applications to other medical devices requiring precise geometric fitting could establish broader impact, as the framework's modular design and validated constraint modeling methodologies provide foundations for advancing automated design systems across multiple medical domains.

REFERENCES

[1] J.-J. Hwang *et al.*, "Learning beyond human expertise with generative models for dental restorations," *arXiv preprint arXiv:1804.00064*, 2018.

[2] Airdrie Choice Dental, "The dental crown process: What to expect," 2024, accessed: [insert your access date]. [Online]. Available: https://www.airdriechoicedental.com/blog/the-dental-crown-process-what-to-expect/

[3] N. Rotroff, "What is dental occlusion and why is it important?" 2024, accessed: [insert your access date]. [Online]. Available: https://www.nancyrotroff.com/blog/what-is-dental-occlusion-and-why-is-it-important

[4] G. Hosseinimanesh, "Dental crown generation under geometric and functional constraints using deep learning," 2023, unpublished research proposal.

[5] Dentagama, "Tooth numbering systems in dentistry," 2024, accessed: September 27, 2025. [Online]. Available: https://dentagama.com/news/dental-numbering-systems

[6] B. Fei *et al.*, "Comprehensive review of deep learning-based 3d point cloud completion processing and analysis," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 12, pp. 22 862–22 883, 2022.

[7] Y. Li *et al.*, "Pccdiff: Point cloud completion with conditional denoising diffusion probabilistic models," *Symmetry*, vol. 16, no. 12, p. 1680, 2024.

[8] Y. Wu, H. Yan, and K. Ding, "Transformer based 3d tooth segmentation via point cloud region partition," *Scientific Reports*, vol. 14, no. 1, p. 28513, 2024.

[9] X. Yu *et al.*, "Pointr: Diverse point cloud completion with geometry-aware transformers," in *Proceedings of the IEEE/CVF international conference on computer vision*, 2021, pp. 12 498–12 507.

[10] S. Peng *et al.*, "Shape as points: A differentiable poisson solver," *Advances in Neural Information Processing Systems*, vol. 34, pp. 13 032–13 044, 2021.

[11] G. Hosseinimanesh *et al.*, "From mesh completion to ai designed crown," in *26th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI 2023)*, ser. Lecutre notes in computer science. Springer, 2023, pp. 555–565. [Online]. Available: https://publications.polymtl.ca/56092/

[12] X. Yu *et al.*, "Adapointr: Diverse point cloud completion with adaptive geometry-aware transformers," 2023. [Online]. Available: https://arxiv.org/abs/2301.04545

[13] R. Hanocka *et al.*, "Point2mesh: a self-prior for deformable meshes," *ACM Transactions on Graphics*, vol. 39, no. 4, Aug. 2020. [Online]. Available: http://dx.doi.org/10.1145/3386569.3392415

[14] J. Huang *et al.*, "Neural kernel surface reconstruction," 2023.

[15] G. Hosseinimanesh *et al.*, "Personalized dental crown design: A point-to-mesh completion network," *Medical Image Analysis*, vol. 101, p. 103439, 2025.

[16] F. Wang *et al.*, "Comparison of the morphological accuracy of automatic crowns designed by multiple computer-aided design software programs with different levels of dentition information acquisition," *The Journal of Prosthetic Dentistry*, vol. 132, no. 2, pp. 441–452, 2024.

[17] A. T.L.C., "3d data representations," 2023, accessed: 2025-04-21. [Online]. Available: https://www.antoinetlc.com/blog-summary/3d-data-representations

[18] Y. Feng *et al.*, "Meshnet: Mesh neural network for 3d shape representation," in *Proceedings of the AAAI conference on artificial intelligence*, vol. 33, no. 01, 2019, pp. 8279–8286.

[19] M. Kazhdan, T. Funkhouser, and S. Rusinkiewicz, "Rotation invariant spherical harmonic representation of 3 d shape descriptors," in *Symposium on geometry processing*, vol. 6, 2003, pp. 156–164.

[20] J. R. Driscoll and D. M. Healy, "Computing fourier transforms and convolutions on the 2-sphere," *Advances in applied mathematics*, vol. 15, no. 2, pp. 202–250, 1994.

[21] Y. Sahillioğlu, "Recent advances in shape correspondence," *The Visual Computer*, vol. 36, no. 8, pp. 1705–1721, 2020.

[22] N. J. Mitra *et al.*, "Structure-aware shape processing," in *ACM SIGGRAPH 2014 Courses*, 2014, pp. 1–21.

[23] A. Makadia, A. Patterson, and K. Daniilidis, "Fully automatic registration of 3d point clouds," in *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06)*, vol. 1. IEEE, 2006, pp. 1297–1304.

[24] D. T. Nguyen *et al.*, "A field model for repairing 3d shapes," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 5676–5684.

[25] A. Khatamian and H. R. Arabnia, "Survey on 3d surface reconstruction," *Journal of Information Processing Systems*, vol. 12, no. 3, pp. 338–357, 2016.

[26] H. Hoppe *et al.*, "Surface reconstruction from unorganized points," in *Proceedings of the 19th annual conference on computer graphics and interactive techniques*, 1992, pp. 71–78.

[27] J. Digne *et al.*, "Feature-preserving surface reconstruction and simplification from defect-laden point sets," *Journal of mathematical imaging and vision*, vol. 48, pp. 369–382, 2014.

[28] H. Q. Dinh, G. Turk, and G. Slabaugh, "Reconstructing surfaces using anisotropic basis functions," in *Proceedings eighth IEEE international conference on computer vision. ICCV 2001*, vol. 2. IEEE, 2001, pp. 606–613.

[29] A. Nealen *et al.*, "Laplacian mesh optimization," in *Proceedings of the 4th international conference on Computer graphics and interactive techniques in Australasia and Southeast Asia*, 2006, pp. 381–389.

[30] M. Kazhdan and H. Hoppe, "Screened poisson surface reconstruction," *ACM Transactions on Graphics (ToG)*, vol. 32, no. 3, pp. 1–13, 2013.

[31] N. J. Mitra, L. J. Guibas, and M. Pauly, "Partial and approximate symmetry detection for 3d geometry," *ACM Transactions on Graphics (ToG)*, vol. 25, no. 3, pp. 560–568, 2006.

[32] N. J. Mitra *et al.*, "Symmetry in 3d geometry: Extraction and applications," in *Computer graphics forum*, vol. 32, no. 6. Wiley Online Library, 2013, pp. 1–23.

[33] M. Pauly *et al.*, "Example-based 3d scan completion," in *Symposium on geometry processing*, 2005, pp. 23–32.

[34] E. Kalogerakis *et al.*, "A probabilistic model for component-based shape synthesis," *Acm Transactions on Graphics (TOG)*, vol. 31, no. 4, pp. 1–11, 2012.

[35] J. Rock *et al.*, "Completing 3d object shape from one depth image," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 2484–2493.

[36] A.-L. Chauve, P. Labatut, and J.-P. Pons, "Robust piecewise-planar 3d reconstruction and completion from large-scale unstructured point data," in *2010 IEEE computer society conference on computer vision and pattern recognition*. IEEE, 2010, pp. 1261–1268.

[37] X. Zhang *et al.*, "View-guided point cloud completion," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2021, pp. 15 890–15 899.

[38] Y. Xia *et al.*, "Asfm-net: Asymmetrical siamese feature matching network for point completion," in *Proceedings of the 29th ACM international conference on multimedia*, 2021, pp. 1938–1947.

[39] L. Pan, "Ecg: Edge-aware point cloud completion with graph convolution," *IEEE Robotics and Automation Letters*, vol. 5, no. 3, pp. 4392–4398, 2020.

[40] X. Wen *et al.*, "Point cloud completion by skip-attention network with hierarchical folding," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 1939–1948.

[41] C. R. Qi *et al.*, "Pointnet: Deep learning on point sets for 3d classification and segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 652–660.

[42] ——, "Pointnet++: Deep hierarchical feature learning on point sets in a metric space," *Advances in neural information processing systems*, vol. 30, 2017.

[43] Y. Yang *et al.*, "Foldingnet: Point cloud auto-encoder via deep grid deformation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 206–215.

[44] W. Yuan *et al.*, "Pcn: Point completion network," in *2018 international conference on 3D vision (3DV)*. IEEE, 2018, pp. 728–737.

[45] Z. Wu *et al.*, "3d shapenets: A deep representation for volumetric shapes," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 1912–1920.

[46] D. Stutz and A. Geiger, "Learning 3d shape completion from laser scan data with weak supervision," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 1955–1964.

[47] H. Xie *et al.*, "Grnet: Gridding residual network for dense point cloud completion," in *European conference on computer vision*. Springer, 2020, pp. 365–381.

[48] H. Thomas *et al.*, "Kpconv: Flexible and deformable convolution for point clouds," in *Proceedings of the IEEE/CVF international conference on computer vision*, 2019, pp. 6411–6420.

[49] W. Wu, Z. Qi, and L. Fuxin, "Pointconv: Deep convolutional networks on 3d point clouds," in *Proceedings of the IEEE/CVF Conference on computer vision and pattern recognition*, 2019, pp. 9621–9630.

[50] A. V. Phan *et al.*, "Dgcnn: A convolutional neural network over large-scale labeled graphs," *Neural Networks*, vol. 108, pp. 533–543, 2018.

[51] K. Zhang *et al.*, "Linked dynamic graph cnn: Learning through point cloud by linking hierarchical features," in *2021 27th international conference on mechatronics and machine vision in practice (M2VIP)*. IEEE, 2021, pp. 7–12.

[52] H. Wu, Y. Miao, and R. Fu, "Point cloud completion using multiscale feature fusion and cross-regional attention," *Image and Vision Computing*, vol. 111, p. 104193, 2021.

[53] T. Groueix *et al.*, "A papier-mâché approach to learning 3d surface generation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 216–224.

[54] M. Liu *et al.*, "Morphing and sampling network for dense point cloud completion," in *Proceedings of the AAAI conference on artificial intelligence*, vol. 34, no. 07, 2020, pp. 11 596–11 603.

[55] L. P. Tchapmi *et al.*, "Topnet: Structural point cloud decoder," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2019, pp. 383–392.

[56] P. Achlioptas *et al.*, "Learning representations and generative models for 3d point clouds," in *International conference on machine learning*. PMLR, 2018, pp. 40–49.

[57] Z. Huang *et al.*, "Pf-net: Point fractal network for 3d point cloud completion," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 7662–7670.

[58] M. Sarmad, H. J. Lee, and Y. M. Kim, "Rl-gan-net: A reinforcement learning agent controlled gan network for real-time point cloud shape completion," in *Proceedings*

of the IEEE/CVF conference on computer vision and pattern recognition, 2019, pp. 5898–5907.

[59] R. Li *et al.*, "Sp-gan: Sphere-guided 3d shape generation and manipulation," *ACM Transactions on Graphics (TOG)*, vol. 40, no. 4, pp. 1–12, 2021.

[60] P. Xiang *et al.*, "Snowflakenet: Point cloud completion by snowflake point deconvolution with skip-transformer," in *Proceedings of the IEEE/CVF international conference on computer vision*, 2021, pp. 5499–5509.

[61] X. Yan *et al.*, "Shapeformer: Transformer-based shape completion via sparse representation," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022, pp. 6239–6249.

[62] N. Sharp *et al.*, "Diffusionnet: Discretization agnostic learning on surfaces," *ACM Transactions on Graphics (TOG)*, vol. 41, no. 3, pp. 1–16, 2022.

[63] J. H. Lim *et al.*, "Score-based diffusion models in function space," *arXiv preprint arXiv:2302.07400*, 2023.

[64] S. Gu *et al.*, "Vector quantized diffusion model for text-to-image synthesis," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2022, pp. 10 696–10 706.

[65] Y. Shi *et al.*, "Mvdream: Multi-view diffusion for 3d generation," *arXiv preprint arXiv:2308.16512*, 2023.

[66] N. Wang *et al.*, "Pixel2mesh: Generating 3d mesh models from single rgb images," in *Proceedings of the European conference on computer vision (ECCV)*, 2018, pp. 52–67.

[67] M. Vakalopoulou *et al.*, "Atlasnet: Multi-atlas non-linear deep networks for medical image segmentation," in *Medical Image Computing and Computer Assisted Intervention–MICCAI 2018: 21st International Conference, Granada, Spain, September 16-20, 2018, Proceedings, Part IV 11*. Springer, 2018, pp. 658–666.

[68] X. Chen *et al.*, "Shape registration with learned deformations for 3d shape reconstruction from sparse and incomplete point clouds," *Medical image analysis*, vol. 74, p. 102228, 2021.

[69] A. X. Chang *et al.*, "Shapenet: An information-rich 3d model repository," *arXiv preprint arXiv:1512.03012*, 2015.

[70] A. Geiger *et al.*, "Vision meets robotics: The kitti dataset. the international journal of robotics research," *Int. J. Rob. Res*, pp. 1–6.

[71] L. Pan *et al.*, "Variational relational point completion network," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2021, pp. 8524–8533.

[72] E. Ahmed *et al.*, "A survey on deep learning advances on different 3d data representations," *arXiv preprint arXiv:1808.01462*, 2018.

[73] P. K. Vinodkumar *et al.*, "Deep learning for 3d reconstruction, augmentation, and registration: a review paper," *Entropy*, vol. 26, no. 3, p. 235, 2024.

[74] T.-H. Wu *et al.*, "Two-stage mesh deep learning for automated tooth segmentation and landmark localization on 3d intraoral scans," *IEEE transactions on medical imaging*, vol. 41, no. 11, pp. 3158–3166, 2022.

[75] A. Alsheghri *et al.*, "Semi-supervised segmentation of tooth from 3d scanned dental arches," 2022. [Online]. Available: https://arxiv.org/abs/2208.05539

[76] Z. Li *et al.*, "Multi-scale bidirectional enhancement network for 3d dental model segmentation," in *2022 IEEE 19th International Symposium on Biomedical Imaging (ISBI)*. IEEE, 2022, pp. 1–5.

[77] Z. Dong, J. Chen, and Y. Xu, "Transformer-based tooth alignment prediction with occlusion and collision constraints," *arXiv preprint arXiv:2410.20806*, 2024.

[78] P. Wang, H. Gu, and Y. Sun, "Tooth segmentation on multimodal images using adapted segment anything model," *Scientific Reports*, vol. 15, no. 1, p. 13874, 2025.

[79] M. S. H. Arian *et al.*, "Unsupervised tooth segmentation from three dimensional scans of the dental arch using domain adaptation of synthetic data," *International Journal of Medical Informatics*, vol. 195, p. 105769, 2025.

[80] A. P. Litzenburger *et al.*, "Fully automatic cad design of the occlusal morphology of partial crowns compared to dental technicians' design," *Clinical oral investigations*, vol. 17, no. 2, pp. 491–496, 2013.

[81] F. Yuan *et al.*, "Personalized design technique for the dental occlusal surface based on conditional generative adversarial networks," *International Journal for Numerical Methods in Biomedical Engineering*, vol. 36, no. 5, p. e3321, 2020.

[82] S. Tian *et al.*, "Dcpr-gan: dental crown prosthesis restoration using two-stage generative adversarial networks," *IEEE Journal of Biomedical and Health Informatics*, vol. 26, no. 1, pp. 151–160, 2021.

[83] Y. Ping *et al.*, "Self-attention implicit function networks for 3d dental data completion," *Computer Aided Geometric Design*, vol. 90, p. 102026, 2021.

[84] H. Zhu *et al.*, "Toothcr: a two-stage completion and reconstruction approach on 3d dental model," in *Pacific-Asia Conference on Knowledge Discovery and Data Mining*. Springer, 2022, pp. 161–172.

[85] T. Farook *et al.*, "Computer-aided design and 3-dimensional artificial/convolutional neural network for digital partial dental crown synthesis and validation. sci rep. 2023; 13: 1561," *DOI: https://doi. org/10.1038/s41598-023-28442-1*, p. 1561.

[86] G. Hosseinimanesh *et al.*, "Improving the quality of dental crown using a transformer-based method," in *Medical Imaging 2023: Physics of Medical Imaging*, vol. 12463. SPIE, 2023, pp. 802–809.

[87] P. Isola *et al.*, "Image-to-image translation with conditional adversarial networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 1125–1134.

[88] A. Vaswani *et al.*, "Attention is all you need," *Advances in neural information processing systems*, vol. 30, 2017.

[89] O. Lessard *et al.*, "Dental restoration using a multi-resolution deep learning approach," in *2022 IEEE 19th International Symposium on Biomedical Imaging (ISBI)*. IEEE, 2022, pp. 1–4.

[90] C. Lian *et al.*, "Deep multi-scale mesh feature learning for automated labeling of raw dental surfaces from 3d intraoral scanners," *IEEE transactions on medical imaging*, vol. 39, no. 7, pp. 2440–2450, 2020.

[91] F. Bernardini *et al.*, "The ball-pivoting algorithm for surface reconstruction," *IEEE transactions on visualization and computer graphics*, vol. 5, no. 4, pp. 349–359, 2002.

[92] G. Hosseinimanesh *et al.*, "From mesh completion to ai designed crown," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2023, pp. 555–565.

[93] K. Sarkar, K. Varanasi, and D. Stricker, "Learning quadrangulated patches for 3d shape parameterization and completion," in *2017 International Conference on 3D Vision (3DV)*. IEEE, 2017, pp. 383–392.

[94] O. Litany *et al.*, "Deformable shape completion with graph convolutional autoencoders," 2018.

[95] W. E. Lorensen and H. E. Cline, "Marching cubes: A high resolution 3d surface construction algorithm," in *Seminal graphics: pioneering efforts that shaped the field*, 1998, pp. 347–353.

[96] I. Loshchilov and F. Hutter, "Decoupled weight decay regularization," *arXiv preprint arXiv:1711.05101*, 2017.

[97] M. Tatarchenko *et al.*, "What do single-view 3d reconstruction networks learn?" in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2019, pp. 3405–3414.

[98] T. M. Tuan *et al.*, "Dental diagnosis from x-ray images: an expert system based on fuzzy computing," *Biomedical Signal Processing and Control*, vol. 39, pp. 64–73, 2018.

[99] M. Rajee and C. Mythili, "Gender classification on digital dental x-ray images using deep convolutional neural network," *Biomedical Signal Processing and Control*, vol. 69, p. 102939, 2021.

[100] C.-H. Wu *et al.*, "Model-based orthodontic assessments for dental panoramic radiographs," *IEEE Journal of biomedical and health informatics*, vol. 22, no. 2, pp. 545–551, 2017.

[101] H. Wu *et al.*, "Automated skin lesion segmentation via an adaptive dual attention module," *IEEE transactions on medical imaging*, vol. 40, no. 1, pp. 357–370, 2020.

[102] J. Wang *et al.*, "Xbound-former: Toward cross-scale boundary modeling in transformers," *IEEE Transactions on Medical Imaging*, vol. 42, no. 6, pp. 1735–1745, 2023.

[103] H. Lei *et al.*, "Unsupervised domain adaptation based image synthesis and feature alignment for joint optic disc and cup segmentation," *IEEE Journal of Biomedical and Health Informatics*, vol. 26, no. 1, pp. 90–102, 2021.

[104] S. Pandey, P. R. Singh, and J. Tian, "An image augmentation approach using two-stage generative adversarial network for nuclei image segmentation," *Biomedical Signal Processing and Control*, vol. 57, p. 101782, 2020.

[105] S. Ding *et al.*, "High-resolution dermoscopy image synthesis with conditional generative adversarial networks," *Biomedical Signal Processing and Control*, vol. 64, p. 102224, 2021.

[106] S. Tian *et al.*, "Efficient computer-aided design of dental inlay restoration: a deep adversarial framework," *IEEE Transactions on Medical Imaging*, vol. 40, no. 9, pp. 2415–2427, 2021.

[107] ——, "A dual discriminator adversarial learning approach for dental occlusal surface reconstruction," *Journal of Healthcare Engineering*, vol. 2022, no. 1, p. 1933617, 2022.

[108] H. Ding *et al.*, "Morphology and mechanical performance of dental crown designed by 3d-dcgan," *dental materials*, vol. 39, no. 3, pp. 320–332, 2023.

[109] S. Tian *et al.*, "Efficient tooth gingival margin line reconstruction via adversarial learning," *Biomedical Signal Processing and Control*, vol. 78, p. 103954, 2022.

[110] A. Alsheghri *et al.*, "Adaptive point learning with uncertainty quantification to generate margin lines on prepared teeth," *Applied Sciences*, vol. 14, no. 20, p. 9486, 2024.

[111] J.-J. Hwang *et al.*, "Learning beyond human expertise with generative models for dental restorations," *ArXiv*, vol. abs/1804.00064, 2018. [Online]. Available: https://api.semanticscholar.org/CorpusID:4570708

[112] F. Yuan *et al.*, "Personalized design technique for the dental occlusal surface based on conditional generative adversarial networks," *International Journal for Numerical Methods in Biomedical Engineering*, vol. 36, no. 5, p. e3321, 2020. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/cnm.3321

[113] P. Isola *et al.*, "Image-to-image translation with conditional adversarial networks," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Los Alamitos, CA, USA: IEEE Computer Society, jul 2017, pp. 5967–5976. [Online]. Available: https://doi.ieeecomputersociety.org/10.1109/CVPR.2017.632

[114] O. Lessard *et al.*, "Dental restoration using a multi-resolution deep learning approach," in *2022 IEEE 19th International Symposium on Biomedical Imaging (ISBI)*. IEEE, 2022, pp. 1–4.

[115] H. Zhu *et al.*, "Toothcr: A two-stage completion and reconstruction approach on 3D dental model," in *Advances in Knowledge Discovery and Data Mining: 26th Pacific-Asia Conference, PAKDD 2022, Chengdu, China, May 16–19, 2022,*

Proceedings, Part III. Berlin, Heidelberg: Springer-Verlag, 2022, p. 161–172. [Online]. Available: https://doi.org/10.1007/978-3-031-05981-0_13

[116] F. Lin *et al.*, "Infocd: A contrastive chamfer distance loss for point cloud completion," in *Advances in Neural Information Processing Systems*, A. Oh *et al.*, Eds., vol. 36. Curran Associates, Inc., 2023, pp. 76 960–76 973. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2023/file/f2ea1943896474b7cd9796b93e526f6f-Paper-Conference.pdf

[117] A. Kazerouni *et al.*, "Diffusion models in medical imaging: A comprehensive survey," *Medical Image Analysis*, vol. 88, p. 102846, 2023.

[118] C. R. Qi *et al.*, "Pointnet: Deep learning on point sets for 3d classification and segmentation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, July 2017, p. n/a.

[119] B. Fei *et al.*, "Comprehensive review of deep learning-based 3d point cloud completion processing and analysis," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 12, p. 22862–22883, Dec. 2022. [Online]. Available: http://dx.doi.org/10.1109/TITS.2022.3195555

[120] Y. Li and L. Liu, "Enhancing diffusion-based point cloud generation with smoothness constraint," 2024.

[121] X. Wu *et al.*, "Fsc: Few-point shape completion," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2024, pp. 26 077–26 087.

[122] S. Sellán and A. Jacobson, "Neural stochastic screened poisson reconstruction," 2023.

[123] X. Chen *et al.*, "Shape registration with learned deformations for 3d shape reconstruction from sparse and incomplete point clouds," *Medical Image Analysis*, vol. 74, p. 102228, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1361841521002735

[124] R. Chu *et al.*, "Diffcomplete: Diffusion-based generative 3d shape completion," *Advances in Neural Information Processing Systems*, vol. 36, 2024.

[125] S. Foti *et al.*, "Intraoperative liver surface completion with graph convolutional vae," in *Lecture Notes in Computer Science*. Springer International Publishing, 2020, p. 198–207. [Online]. Available: http://dx.doi.org/10.1007/978-3-030-60365-6_19

[126] T. Wu *et al.*, "Density-aware chamfer distance as a comprehensive metric for point cloud completion," 2021.

[127] F. Lin *et al.*, "Hyperbolic chamfer distance for point cloud completion," in *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, October 2023, pp. 14 595–14 606.

[128] T. Huang *et al.*, "Contrastive adversarial loss for point cloud reconstruction," *n/a*.

[129] Y. Wang *et al.*, "Dynamic graph cnn for learning on point clouds," *ACM Transactions on Graphics (tog)*, vol. 38, no. 5, pp. 1–12, 2019.

[130] Wikipedia Contributors, "Spherical harmonics," https://en.wikipedia.org/wiki/Spherical_harmonics, 2024, accessed: May 8, 2024.

[131] K. He *et al.*, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

[132] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 3431–3440.

[133] A. Esteva *et al.*, "Deep learning-enabled medical computer vision," *NPJ digital medicine*, vol. 4, no. 1, p. 5, 2021.

[134] D. Shen, G. Wu, and H.-I. Suk, "Deep learning in medical image analysis," *Annual review of biomedical engineering*, vol. 19, no. 1, pp. 221–248, 2017.

[135] G. Litjens *et al.*, "A survey on deep learning in medical image analysis," *Medical image analysis*, vol. 42, pp. 60–88, 2017.

[136] R. Vashisht *et al.*, "Artificial intelligence in dentistry-a scoping review," *Journal of Oral and Maxillofacial Surgery, Medicine, and Pathology*, 2024.

[137] J. Abduo and K. Lyons, "Clinical considerations for increasing occlusal vertical dimension: a review," *Australian dental journal*, vol. 57, no. 1, pp. 2–10, 2012.

[138] P. C. Guess *et al.*, "All-ceramic systems: laboratory and clinical performance," *Dental clinics*, vol. 55, no. 2, pp. 333–352, 2011.

[139] F. Beuer, J. Schweiger, and D. Edelhoff, "Digital dentistry: an overview of recent developments for cad/cam generated restorations," *British dental journal*, vol. 204, no. 9, pp. 505–511, 2008.

[140] T. Miyazaki and Y. Hotta, "Cad/cam systems available for the fabrication of crown and bridge restorations," *Australian dental journal*, vol. 56, pp. 97–106, 2011.

[141] S. Yang *et al.*, "Dcrownformer: Morphology-aware point-to-mesh generation transformer for dental crown prosthesis from 3d scan data of antagonist and preparation teeth," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2024, pp. 109–119.

[142] S.-J. Byun *et al.*, "Analysis of proximal contact loss between implant-supported fixed dental prostheses and adjacent teeth in relation to influential factors and effects. a cross-sectional study," *Clinical Oral Implants Research*, vol. 26, no. 6, pp. 709–714, 2015. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/clr.12373

[143] Y. Huang *et al.*, "Collision-aware interactive simulation using graph neural networks," *Visual Computing for Industry, Biomedicine, and Art*, vol. 6, no. 1, pp. 1–10, 2023.

[144] J. Tan *et al.*, "Active learning of neural collision handler for complex 3d mesh deformations," in *Proceedings of the 39th International Conference on Machine Learning (ICML)*. PMLR, 2022.

[145] J. Wang, R. Lindenbergh, and M. Menenti, "Evaluating voxel enabled scalable intersection of large point clouds," *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. II-3/W5, pp. 25–31, 2015. [Online]. Available: https://isprs-annals.copernicus.org/articles/II-3-W5/25/2015/

[146] H. Chen *et al.*, "An approach to boundary detection for 3d point clouds based on dbscan clustering," *Pattern Recognition*, vol. 124, p. 108431, 2022.

[147] L. W. Graber *et al.*, *Orthodontics: current principles and techniques: first SA Edn.* Elsevier Health Sciences, 2016.

[148] A. Dawood *et al.*, "3d printing in dentistry," *British dental journal*, vol. 219, no. 11, pp. 521–529, 2015.

[149] K. M. Day, K. S. Gabrick, and L. A. Sargent, "Applications of computer technology in complex craniofacial reconstruction," *Plastic and Reconstructive Surgery–Global Open*, vol. 6, no. 3, p. e1655, 2018.

[150] T. Maruyama *et al.*, "Computer-aided determination of occlusal contact points for dental 3-d cad," *Medical and Biological Engineering and Computing*, vol. 44, no. 5, pp. 445–450, 2006.

[151] N. Z. Al-Rayes and M. Y. Hajeer, "Evaluation of occlusal contacts among different groups of malocclusion using 3d digital models," *The Journal of Contemporary Dental Practice*, vol. 15, pp. 46–55, 2014.

[152] H. Lee *et al.*, "Comparison of the occlusal contact area of virtual models and actual models: a comparative in vitro study on class i and class ii malocclusion models," *BMC Oral Health*, vol. 18, no. 1, p. 109, 2018.

[153] Z. Zhao *et al.*, "Construction of a novel digital method for quantitative analysis of occlusal contact and force," *BMC Oral Health*, vol. 23, no. 1, p. 190, 2023.

[154] F. Lin *et al.*, "Infocd: a contrastive chamfer distance loss for point cloud completion," *Advances in Neural Information Processing Systems*, vol. 36, pp. 76 960–76 973, 2023.

[155] J.-H. Cho *et al.*, "Tooth morphology, internal fit, occlusion and proximal contacts of dental crowns designed by deep learning-based dental software: a comparative study," *Journal of dentistry*, vol. 141, p. 104830, 2024.