Medical Image Analysis 101 (2025) 103439

Contents lists available at ScienceDirect

Medical Image Analysis

journal homepage: www.elsevier.com/locate/media

Personalized dental crown design: A point-to-mesh completion network

Golriz Hosseinimanesh [a,*], Ammar Alsheghri [b,c], Julia Keren [d], Farida Cheriet [a], Francois Guibault [a]

[a] Polytechnique Montréal University, Canada
[b] Mechanical Engineering Department, King Fahd University of Petroleum and Minerals (KFUPM), Dhahran, 31261, Kingdom of Saudi Arabia
[c] Interdisciplinary research center for Biosystems and Machines, King Fahd University of Petroleum and Minerals (KFUPM), Dhahran, Kingdom of Saudi Arabia
[d] Intellident Dentaire Inc., Canada

ARTICLE INFO

Keywords:
Dental crown generation
Transformer
Mesh completion
Margin line

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%.

1. 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. 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. 1) illustrates the detailed adjustments and customization involved in manually designing hundreds of crowns each year to meet individual patient needs.

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 (Tuan et al., 2018; Rajee and Mythili, 2021; Wu et al., 2017). 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 (Wu et al., 2020; Wang et al., 2023; Lei et al., 2021; Pandey et al., 2020; Ding et al., 2021; Tian et al., 2021b). In the field of 3D dental reconstruction, specific studies have advanced synthetic image reconstruction and

* Corresponding author.
E-mail address: golriz.hosseinimanesh@polymtl.ca (G. Hosseinimanesh).

https://doi.org/10.1016/j.media.2024.103439
Received 6 June 2024; Received in revised form 24 November 2024; Accepted 10 December 2024
Available online 17 December 2024
1361-8415/© 2024 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

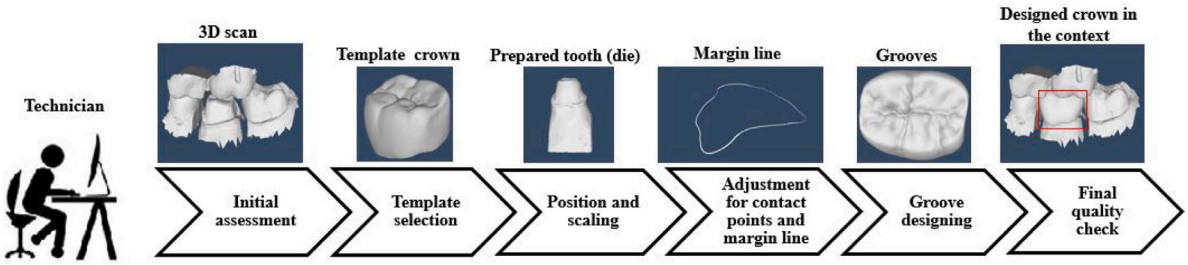


Fig. 1. Flowchart of manually designing a dental crown by a technician. This process usually takes more than 1 h.

prosthetic restoration, focusing on detailed crown morphology, gingival margin line generation, and occlusal surface reconstruction through innovative GAN and transformer-based models (Tian et al., 2022a; Ding et al., 2023; Tian et al., 2021a, 2022b; Alsheghri et al., 2024). 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 (Hwang et al., 2018; Yuan et al., 2020). The pix2pix technique (Isola et al., 2017) 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, Lessard et al. (2022) 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 (Zhu et al., 2022) 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 (Hosseinimanesh et al., 2023b), 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 (Hosseinimanesh et al., 2023a). 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) (Lin et al., 2023b), 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.

2. 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.

2.1. 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. Tian et al. (2022a) 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, Ding et al. (2023) 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 Tian et al. (2021b), 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 Tian et al. (2021a) 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, Tian et al. (2022b) 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 (Kazerouni et al., 2023) 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.

2.2. 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 (Qi et al., 2017a,b; Yang et al., 2018). According to Fei et al. (2022), 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 spatial 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 (Fei et al., 2022). Diffusion models, meanwhile, methodically convert noise into structured point clouds, enabling the creation of complex geometries in a controlled manner (Li and Liu, 2024).

Focusing on techniques that address missing regions, Huang et al. (2020) 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 (Yu et al., 2021), SnowflakeNet (Xiang et al., 2021), and AdapoinTr (Yu et al., 2023) 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 Wu et al. (2024) 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.

2.3. 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 (Sarkar et al., 2017; Sellán and Jacobson, 2023) 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 (Litany et al., 2018). 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 (Chen et al., 2021). 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 (Chu et al., 2024) 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 (Foti et al., 2020; Litany et al., 2018). However, the iterative optimization required at inference by these techniques, particularly by Point2Mesh (Hanocka et al., 2020), 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) (Huang et al., 2023a) 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.

2.4. Evaluation metrics

In the field of point cloud completion, the Chamfer Distance (CD) and Earth Mover's Distance (EMD) (Yuan et al., 2019) 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, cannot 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) (Wu et al., 2021) 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) (Lin et al., 2023a) 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 (Huang et al., 2023b) 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 (Lin et al., 2023b) 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.

3. 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. 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.

3.1. Point transformer encoder-decoder

We adapt the transformer encoder–decoder architecture (Yu et al., 2023) 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) (Qi et al., 2017b) and a lightweight dynamic graph convolution network (DGCNN) (Wang et al., 2019) 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.

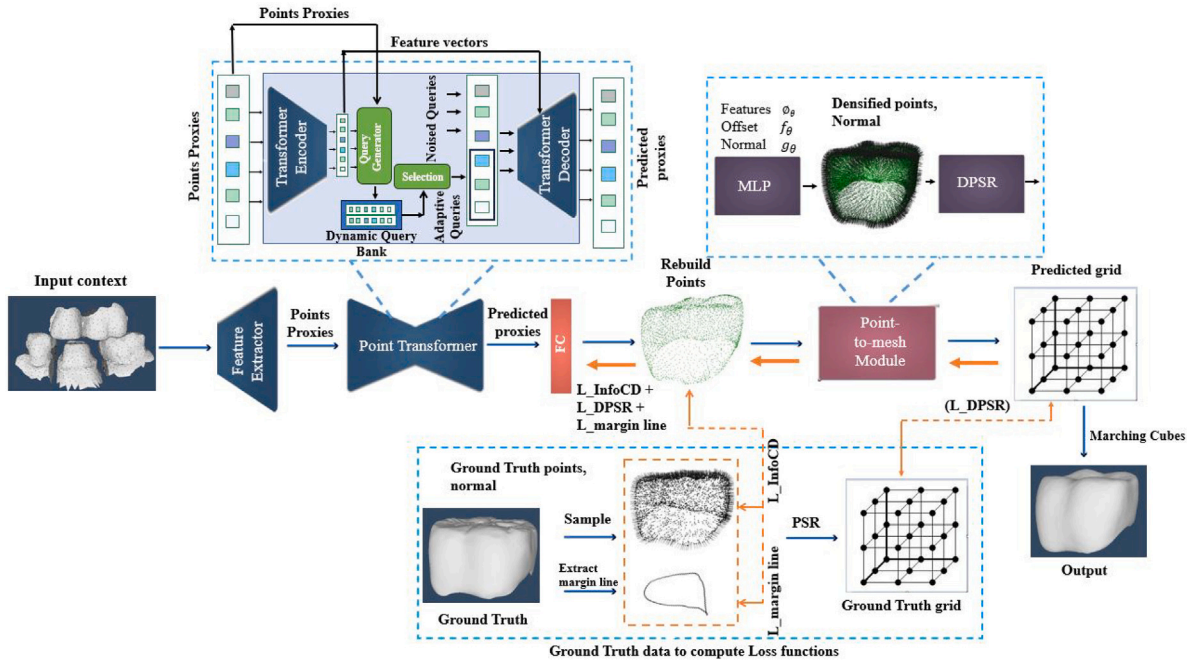


Fig. 2. Architecture flow of our proposed model, highlighting key components such as DPSR (Differentiable Poisson Surface Reconstruction), FC (Fully Connected Layer), and MLP (Multi-Layer Perceptron).

To generate queries, which serve as the starting state of predicted proxies, we employ a multi-layer perceptron (MLP) query generator by Yu et al. (2023). 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 (Yu et al., 2023) 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.

3.2. 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 Peng et al. (2021), 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) (Peng et al., 2021). 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 (Peng et al., 2021). The solution to the PDE is then derived by manipulating the unnormalized indicator function χ' with an inverse FFT as Eq. (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 Eq. (2):

$$\chi' = \mathrm{IFFT}\left(\hat{\chi}\right), \tag{1}$$

$$\chi = \frac{m}{\mathrm{abs}\left(\chi'_{|x=0}\right)}\left(\chi' - \frac{1}{|\{C\}|}\sum_{c\in\{C\}}\chi'_{|x=c}\right) \tag{2}$$

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) (Peng et al., 2021) 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 (Peng et al., 2021) to predict the indicator function χ. The extracted scalar field from this function serves as the basis for the

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 (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

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.

3.3. Loss function

We use three different losses to constrain the output shape and predict a high-quality mesh. We utilize the InfoCD loss (Lin et al., 2023b), 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 Eqs. (3), (4), (5), (6):

$$L_{\text{InfoCD}}(x_i, y_i) = \ell_{\text{InfoCD}}(x_i, y_i) + \ell_{\text{InfoCD}}(y_i, x_i), \tag{3}$$

where each term ℓ_{InfoCD} is calculated by:

$$\ell_{\text{InfoCD}}(x_i, y_i) = -\frac{1}{|y_i|}\sum_k \log\left(\frac{\exp\left\{-\frac{1}{\tau}\min_j d(x_{ij}, y_{ik})\right\}}{\left(\sum_k \exp\left\{-\frac{1}{\tau}\min_j d(x_{ij}, y_{ik})\right\}\right)^\lambda}\right), \tag{4}$$

$$L_{\text{InfoCD}}(x_i, y_i) \propto \frac{1}{\tau} L_{\text{CD}}(x_i, y_i) + R(x_i, y_i) \tag{5}$$

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:

$$R(x_i, y_i) = \log\left\{\sum_{m,n} \exp\left(-\frac{1}{\tau}\left[\min_j d(x_{ij}, y_{in}) + \min_k d(x_{im}, y_{ik})\right]\right)\right\}, \tag{6}$$

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 (Yu et al., 2021). 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 alignment, allowing for a flexible approach to point cloud completion. The InfoCD loss, as developed by Lin et al. (2023b), serves as a bridge between the traditional CD loss and a more structured, distribution-sensitive metric.

Algorithm 2 Margin Line Loss Function.

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

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) (Yuan et al., 2019) 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 Eq. (7) and Algorithm 2.

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

$$L_{\text{margin line}}(S_1, S_2) = \frac{1}{|S_1|}\sum_{x\in S_1}\min_{y\in S_2}|x-y|^2 + \frac{1}{|S_2|}\sum_{y\in S_2}\min_{x\in S_1}|y-x|^2 \tag{7}$$

Furthermore, we minimize the L_2 distance between the predicted and ground truth indicator functions, both derived from Poisson PDE solutions (Peng et al., 2021). This Mean Square Error (MSE) loss is computed as Eq. (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.

$$L_{DPSR}(\theta) = E_{X_i, x_i \sim D}\|Poisson(f_\theta(X_i)) - x_i\|_2^2 \tag{8}$$

The final loss function, represented as equal weighted sum of the individual loss functions, is described by Eq. (9):

$$L_{\text{total}} = L_{\text{InfoCD}} + L_{\text{margin line}} + L_{\text{DPSR}} \tag{9}$$

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 towards any particular loss component, supporting well-rounded development.

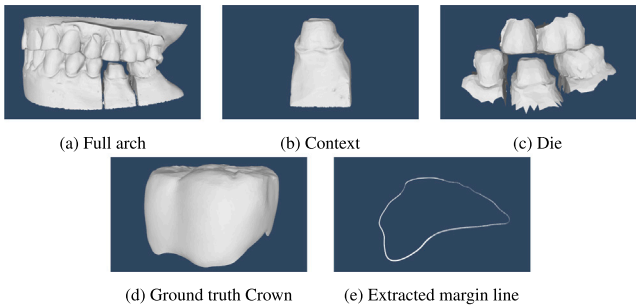


Fig. 3. (a) Dental arch with a prep, (b) generated context, (c) die (prepared tooth), (d) Ground truth crown, (e) extracted margin line.

4. 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.

4.1. 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. 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 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 Peng et al. (2021) to compute the indicator grid for our ground truth mesh.

4.2. Implementation details

Our model incorporates a transformer encoder–decoder module that aligns with the architecture detailed in Yu et al. (2023). For the network's mesh reconstruction phase, we employ the differentiable Poisson surface reconstruction (DPSR) method, as introduced in Peng et al. (2021). 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 h of computing time on an NVIDIA A100 GPU.

Table 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 (Hosseinimanesh et al., 2023a)	65.04	18.31	–
DMC (Hosseinimanesh et al., 2023b)	62.03	11.06	0.0028
Our model	**54.39**	**8.41**	**0.0015**

4.3. 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 1, comparing it with our previous methods. Pointr+ Margin Line (Hosseinimanesh et al., 2023a) builds on the point transformer approach (Yu et al., 2021) 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 (Hosseinimanesh et al., 2023b) 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 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. 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 1, the evaluation metrics for our new model — specifically, Chamfer distance and MSE — show improvements compared to the previous method (Hosseinimanesh et al., 2023b). This demonstrates

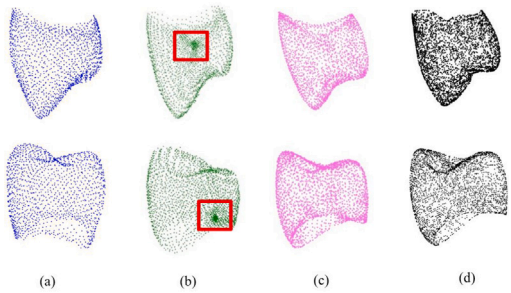


Fig. 4. Comparison of predicted point clouds using our proposed method with previous work. (*a*) Illustrates the predicted point cloud using the method by Hosseinimanesh et al. (2023b) with 1536 points. (*b*) Depicts the same method by Hosseinimanesh et al. (2023b), but with a higher resolution of 3072 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.

Table 2
Comparison of our model with the sate of the art methods in the domain. Metrics are the same as in Table 1.

Method	CD-L1 (↓)	CD-L2 (↓)	MSE (↓)
ToothCR (Zhu et al., 2022)	70.01	23.42	0.0024
AdapoinTr (Yu et al., 2023)	68.82	15.14	–
FSC (Wu et al., 2024)	72.86	17.62	–
Our model	**54.39**	**8.41**	**0.0015**

that the predicted points are more closely aligned with the ground truth, validating the enhanced accuracy of our approach.

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 Zhu et al. (2022), Yu et al. (2023) and Wu et al. (2024), 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 compare our model against state-of-the-art methods in the domain, as shown in Table 2. ToothCR (Zhu et al., 2022) employs a dual-network strategy, first predicting crown point clouds and then reconstructing surfaces from these points. AdaPoinTr (Yu et al., 2023), a point transformer-based model, completes the dental context by leveraging transformers. Few-point Shape Completion (FSC) (Wu et al., 2024) 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 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. 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. 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 provides visual results produced by our model, highlighting its ability to generate accurate, high-quality dental crowns compared to other methods.

In Fig. 7 , we present a color map analysis for four tooth types: Canine, Incisor, Molar, and Premolar, each visualized across instances

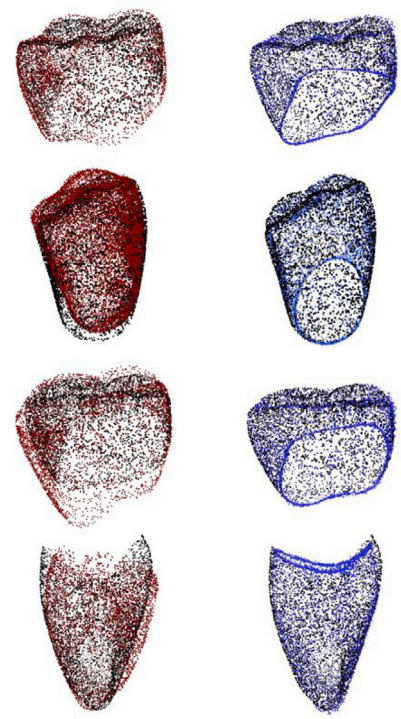
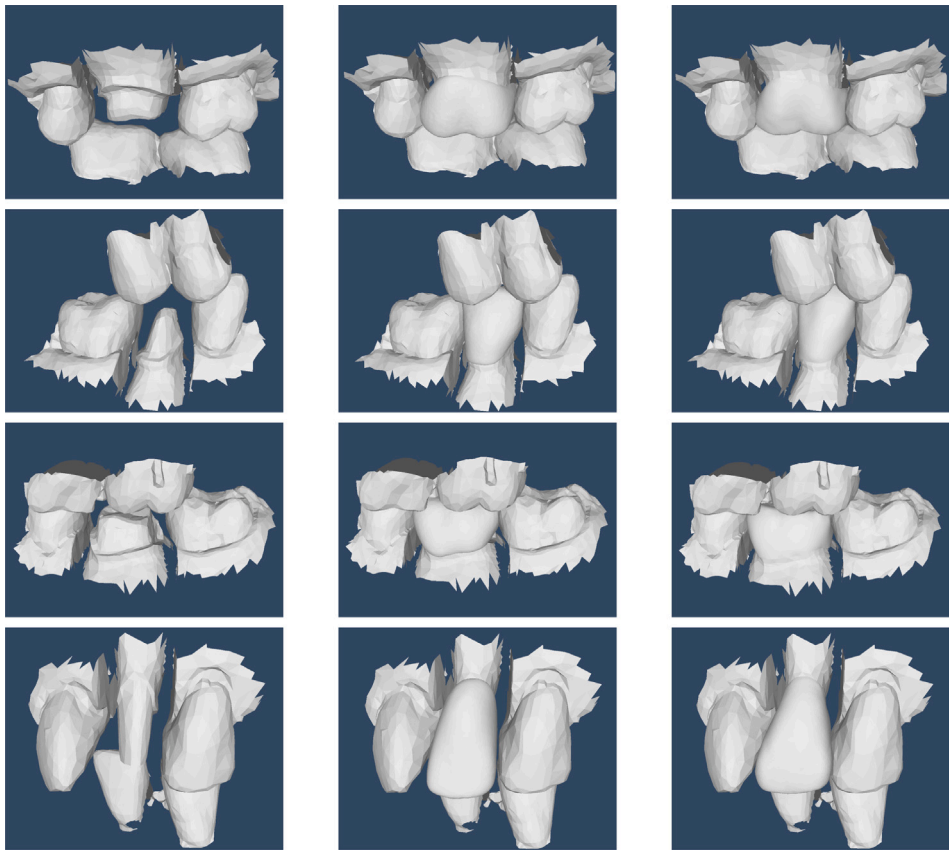


Fig. 5. Visual comparison of dental crown predictions: Predictions from the FSC model (left column) and AdaPoinTr model (right column) are shown for four cases. In each visualization, black points represent the ground truth, red points indicate predictions by the FSC model, and blue points correspond 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.

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 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 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



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

Fig. 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.

teeth. By generating a 3D grid around the crown and 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 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. 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.

Table 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

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.

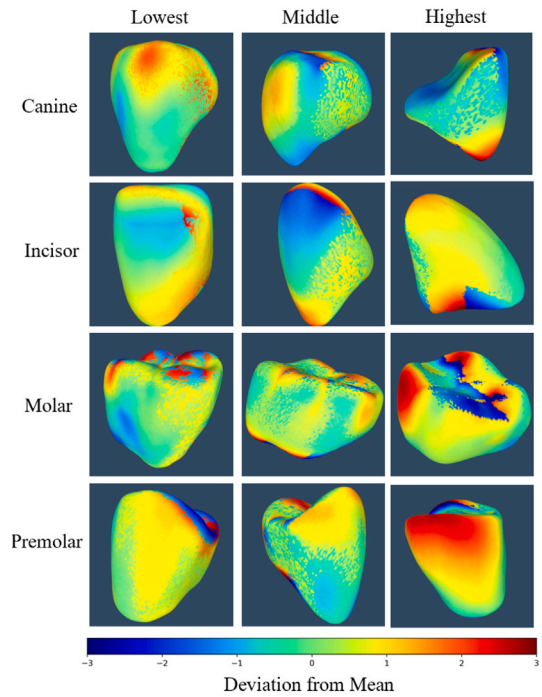


Fig. 7. Color maps showing the signed distance (in mm) between predicted and ground-truth 3D meshes for four tooth types — Canine, Incisor, Molar, and Premolar — each visualized at low, median, and high deviation levels. Positive values indicate mesh overestimation, and negative values indicate underestimation. 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.

4.4. Ablation analysis

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

4.4.1. 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 4. Subsequently, Fig. 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. 9 panel (c), predictions without the margin line loss (represented in blue) do not fit 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.

4.4.2. 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 4. The loss functions under comparison include Chamfer Distance (CD) (Fan et al., 2016), Density-aware Chamfer Distance (DCD) (Wu et al., 2021), and Hyperbolic Chamfer Distance (HyperCD) (Lin et al., 2023a). Our evaluation criteria encompass the average CD and Earth Mover's Distance (EMD) (Fan et al., 2016) across unseen datasets. The results demonstrate that our model, employing the InfoCD loss function (Lin et al., 2023b), 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.

Table 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 (Fan et al., 2016)	69.12	18.3	99.613
DCD (Wu et al., 2021)	67.8	15.4	97.61
HyperCD (Lin et al., 2023a)	66.27	12.83	86.38
InfoCD (Lin et al., 2023b)	**57.61**	**9.45**	**77.67**

Table 5
Comparative analysis of mesh Reconstruction methods. Metrics are the same as in Table 4.

Method	CD-L1 (↓)	CD-L2 (↓)	EMD (↓)
Point Transformer (PT) (Yu et al., 2023)	70.08	23.05	99.42
PT + SAP (Peng et al., 2021)	67.01	21.12	96.03
PT + NKSR (Huang et al., 2023a)	61.43	18.34	86.23
PT + Point2Mesh (Hanocka et al., 2020)	59.62	15.27	80.01
Our model	**54.39**	**8.41**	**75.31**

4.4.3. 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 (Yu et al., 2023) 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 (Peng et al., 2021) to investigate its synergistic effect on mesh reconstruction, leveraging the point cloud output generated by the transformer. Following this, we integrate Neural Kernel Surface Reconstruction (NKSR) (Huang et al., 2023a) into our evaluation, exploring its efficacy in refining meshes derived from the identical point cloud dataset. PoinT2Mesh (Hanocka et al., 2020) 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 5 and selectively visualized in Fig. 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.

5. 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 (Hosseinimanesh et al., 2023a,b; Zhu et al., 2022). Additionally, the integration with differentiable Poisson

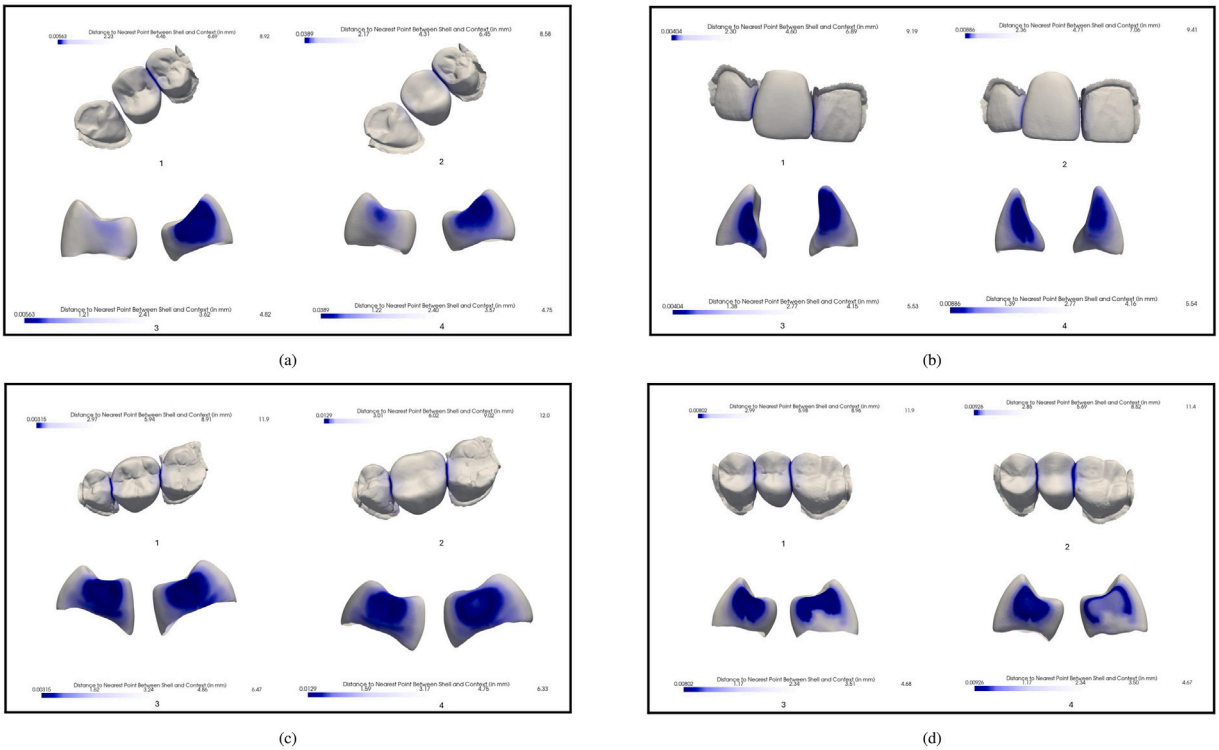


Fig. 8. Heatmap visualizations of the intersection function, displaying multiple cases for different tooth positions: (*a*) position 24, (*b*) position 11, (*c*) position 16, and (*d*) position 25. In each case, subfigures (1) and (2) represent the ground truth and predicted distances, respectively, 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.

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 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 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 1 and 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 (Zhu et al., 2022), the Pointr+margin line approach (Hosseinimanesh et al., 2023a), DMC (Hosseinimanesh et al., 2023b), the AdapoinTr model (Yu et al., 2023), and FSC (Wu et al., 2024), 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. 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. Conclusion and future work

Our research proposes a novel model for creating detailed mesh models of dental crowns from incomplete point cloud data. Leveraging

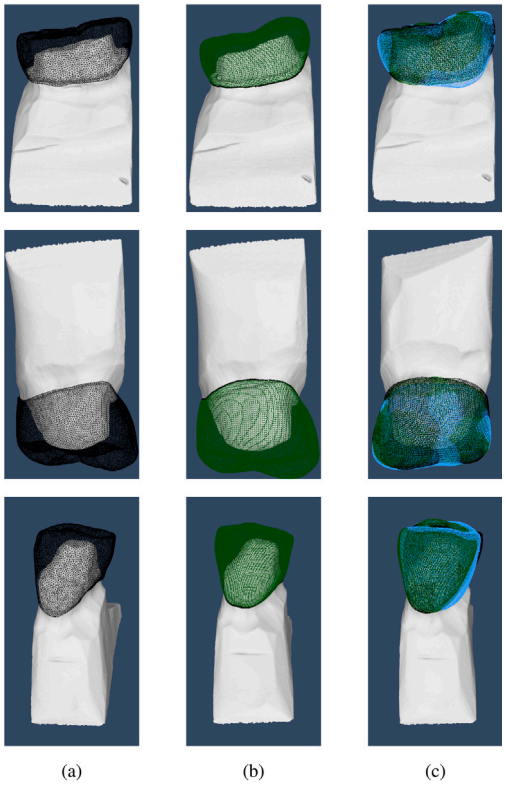


Fig. 9. Illustration of the impact of incorporating a margin line loss function on our model's ability to generate dental crown meshes. Panel (a) showcases the ground truth crown for the prepared tooth surface. Panel (b) displays the model's generated crown utilizing the margin line loss, illustrating 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.

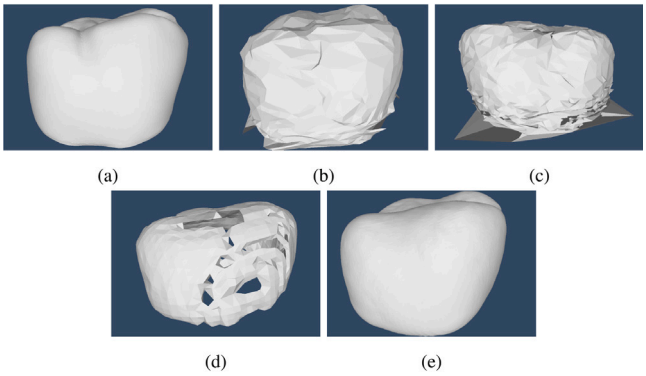


Fig. 10. Qualitative comparison of different approaches for crown mesh generation: (a) Ground truth shape; (b) point cloud processed using Point Transformer (Yu et al., 2023) and mesh generated by Shape as Points (Peng et al., 2021); (c) point cloud processed using Point Transformer (Yu et al., 2023) and mesh generated by Point2Mesh (Hanocka et al., 2020); (d) point cloud processed using Point Transformer (Yu et al., 2023) with mesh reconstruction by NKSR (Huang et al., 2023a); (e) proposed method.

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 (Wikipedia Contributors, 2024) 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.

CRediT authorship contribution statement

Golriz Hosseinimanesh: Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Ammar Alsheghri:** Validation, Writing – review & editing. **Julia Keren:** Funding acquisition. **Farida Cheriet:** Writing – review & editing, Supervision. **Francois Guibault:** Writing – review & editing, Supervision, Funding acquisition.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Golriz Hosseinimanesh reports financial support was provided by Intellident Dentaire Inc., iMD Research Inc., NSERC and MEDTEQ. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors would 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.

Funding

Golriz Hosseinimanesh, Julia Keren, Farida Cheriet, and Francois Guibault acknowledge the funding received from Intellident Dentaire Inc., iMD Research Inc., the Natural Science and Engineering Research Council of Canada (NSERC) [Ref: ALLRP 583415-23], and MEDTEQ [Ref: 19-D Volumétrie dentaire 2]. Author Ammar Alsheghri acknowledges the funding received under KFUPM Funded Grant [Ref: EC241009].

Data availability

The authors do not have permission to share data.

References

Alsheghri, A., Ladini, Y., Hosseinimanesh, G., Chafi, I., Keren, J., Cheriet, F., Guibault, F., 2024. Adaptive point learning with uncertainty quantification to generate margin lines on prepared teeth. Appl. Sci. 14 (20), 9486.

Chen, X., Ravikumar, N., Xia, Y., Attar, R., Diaz-Pinto, A., Piechnik, S.K., Neubauer, S., Petersen, S.E., Frangi, A.F., 2021. Shape registration with learned deformations for 3D shape reconstruction from sparse and incomplete point clouds. Med. Image Anal. 74, 102228. http://dx.doi.org/10.1016/j.media.2021.102228, URL: https://www.sciencedirect.com/science/article/pii/S1361841521002735.

Chu, R., Xie, E., Mo, S., Li, Z., Nießner, M., Fu, C.-W., Jia, J., 2024. Diffcomplete: Diffusion-based generative 3d shape completion. Adv. Neural Inf. Process. Syst. 36.

Ding, H., Cui, Z., Maghami, E., Chen, Y., Matinlinna, J.P., Pow, E.H.N., Fok, A.S.L., Burrow, M.F., Wang, W., Tsoi, J.K.H., 2023. Morphology and mechanical performance of dental crown designed by 3D-DCGAN. Dent. Mater. 39 (3), 320–332.

Ding, S., Zheng, J., Liu, Z., Zheng, Y., Chen, Y., Xu, X., Lu, J., Xie, J., 2021. High-resolution dermoscopy image synthesis with conditional generative adversarial networks. Biomed. Signal Process. Control 64, 102224.

Fan, H., Su, H., Guibas, L., 2016. A point set generation network for 3D object reconstruction from a single image. arXiv:1612.00603.

Fei, B., Yang, W., Chen, W.-M., Li, Z., Li, Y., Ma, T., Hu, X., Ma, L., 2022. Comprehensive review of deep learning-based 3D point cloud completion processing and analysis. IEEE Trans. Intell. Transp. Syst. 23 (12), 22862–22883. http://dx.doi.org/10.1109/tits.2022.3195555.

Foti, S., Koo, B., Dowrick, T., Ramalhinho, J., Allam, M., Davidson, B., Stoyanov, D., Clarkson, M.J., 2020. Intraoperative liver surface completion with graph convolutional VAE. In: Lecture Notes in Computer Science. Springer International Publishing, pp. 198–207. http://dx.doi.org/10.1007/978-3-030-60365-6_19.

Hanocka, R., Metzer, G., Giryes, R., Cohen-Or, D., 2020. Point2Mesh: a self-prior for deformable meshes. ACM Trans. Graph. 39 (4), http://dx.doi.org/10.1145/3386569.3392415.

Hosseinimanesh, G., Ghadiri, F., Alsheghri, A.A., Zhang, Y., Keren, J., Cheriet, F., Guibault, F., 2023a. Improving the quality of dental crown using a transformer-based method. In: Medical Imaging. p. n/a, URL: https://api.semanticscholar.org/CorpusID:257365370.

Hosseinimanesh, G., Ghadiri, F., Guibault, F., Cheriet, F., Keren, J., 2023b. From mesh completion to AI designed crown. In: 26th International Conference on Medical Image Computing and Computer Assisted Intervention. MICCAI 2023, In: Lecutre Notes in Computer Science, Springer, pp. 555–565. http://dx.doi.org/10.1007/978-3-031-43996-4_53, URL: https://publications.polymtl.ca/56092/.

Huang, T., Ding, Z., Zhang, J., Tai, Y., Zhang, Z., Chen, M., Wang, C., Liu, Y., 2023b. Contrastive Adversarial Loss for Point Cloud Reconstruction.

Huang, J., Gojcic, Z., Atzmon, M., Litany, O., Fidler, S., Williams, F., 2023a. Neural kernel surface reconstruction. arXiv:2305.19590.

Huang, Z., Yu, Y., Xu, J., Ni, F., Le, X., 2020. PF-net: Point fractal network for 3D point cloud completion. arXiv:2003.00410.

Hwang, J.-J., Azernikov, S., Efros, A.A., Yu, S.X., 2018. Learning beyond human expertise with generative models for dental restorations. ArXiv abs/1804.00064. URL: https://api.semanticscholar.org/CorpusID:4570708.

Isola, P., Zhu, J., Zhou, T., Efros, A.A., 2017. Image-to-image translation with conditional adversarial networks. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition. CVPR, IEEE Computer Society, Los Alamitos, CA, USA, pp. 5967–5976. http://dx.doi.org/10.1109/CVPR.2017.632, URL: https://doi.ieeecomputersociety.org/10.1109/CVPR.2017.632.

Kazerouni, A., Aghdam, E.K., Heidari, M., Azad, R., Fayyaz, M., Hacihaliloglu, I., Merhof, D., 2023. Diffusion models in medical imaging: A comprehensive survey. Med. Image Anal. 88, 102846.

Lei, H., Liu, W., Xie, H., Zhao, B., Yue, G., Lei, B., 2021. Unsupervised domain adaptation based image synthesis and feature alignment for joint optic disc and cup segmentation. IEEE J. Biomed. Health Inf. 26 (1), 90–102.

Lessard, O., Guibault, F., Keren, J., Cheriet, F., 2022. Dental restoration using a multi-resolution deep learning approach. In: 2022 IEEE 19th International Symposium on Biomedical Imaging. ISBI, IEEE, pp. 1–4.

Li, Y., Liu, L., 2024. Enhancing diffusion-based point cloud generation with smoothness constraint. arXiv:2404.02396.

Lin, F., Yue, Y., Hou, S., Yu, X., Xu, Y., Yamada, K.D., Zhang, Z., 2023a. Hyperbolic chamfer distance for point cloud completion. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. ICCV, pp. 14595–14606.

Lin, F., Yue, Y., Zhang, Z., Hou, S., Yamada, K., Kolachalama, V., Saligrama, V., 2023b. InfoCD: A contrastive chamfer distance loss for point cloud completion. In: Oh, A., Neumann, T., Globerson, A., Saenko, K., Hardt, M., Levine, S. (Eds.), Advances in Neural Information Processing Systems. Vol. 36, Curran Associates, Inc., pp. 76960–76973, URL: https://proceedings.neurips.cc/paper_files/paper/2023/file/f2ea1943896474b7cd9796b93e526f6f-Paper-Conference.pdf.

Litany, O., Bronstein, A., Bronstein, M., Makadia, A., 2018. Deformable shape completion with graph convolutional autoencoders. arXiv:1712.00268.

Pandey, S., Singh, P.R., Tian, J., 2020. An image augmentation approach using two-stage generative adversarial network for nuclei image segmentation. Biomed. Signal Process. Control 57, 101782.

Peng, S., Jiang, C.M., Liao, Y., Niemeyer, M., Pollefeys, M., Geiger, A., 2021. Shape as points: A differentiable Poisson solver. arXiv:2106.03452.

Qi, C.R., Su, H., Mo, K., Guibas, L.J., 2017a. PointNet: Deep learning on point sets for 3D classification and segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. CVPR, p. n/a.

Qi, C.R., Yi, L., Su, H., Guibas, L.J., 2017b. PointNet++: Deep hierarchical feature learning on point sets in a metric space. arXiv:1706.02413.

Rajee, M., Mythili, C., 2021. Gender classification on digital dental x-ray images using deep convolutional neural network. Biomed. Signal Process. Control 69, 102939.

Sarkar, K., Varanasi, K., Stricker, D., 2017. Learning quadrangulated patches for 3D shape parameterization and completion. arXiv:1709.06868.

Sellán, S., Jacobson, A., 2023. Neural stochastic screened Poisson reconstruction. arXiv:2309.11993.

Tian, S., Huang, R., Li, Z., Fiorenza, L., Dai, N., Sun, Y., Ma, H., 2022a. A dual discriminator adversarial learning approach for dental occlusal surface reconstruction. J. Healthc. Eng. 2022 (1), 1933617.

Tian, S., Wang, M., Dai, N., Ma, H., Li, L., Fiorenza, L., Sun, Y., Li, Y., 2021a. DCPR-GAN: dental crown prosthesis restoration using two-stage generative adversarial networks. IEEE J. Biomed. Health Inf. 26 (1), 151–160.

Tian, S., Wang, M., Ma, H., Huang, P., Dai, N., Sun, Y., Meng, J., 2022b. Efficient tooth gingival margin line reconstruction via adversarial learning. Biomed. Signal Process. Control 78, 103954.

Tian, S., Wang, M., Yuan, F., Dai, N., Sun, Y., Xie, W., Qin, J., 2021b. Efficient computer-aided design of dental inlay restoration: a deep adversarial framework. IEEE Trans. Med. Imaging 40 (9), 2415–2427.

Tuan, T.M., Fujita, H., Dey, N., Ashour, A.S., Ngoc, V.T.N., Chu, D.-T., et al., 2018. Dental diagnosis from X-ray images: an expert system based on fuzzy computing. Biomed. Signal Process. Control 39, 64–73.

Wang, J., Chen, F., Ma, Y., Wang, L., Fei, Z., Shuai, J., Tang, X., Zhou, Q., Qin, J., 2023. Xbound-former: Toward cross-scale boundary modeling in transformers. IEEE Trans. Med. Imaging 42 (6), 1735–1745.

Wang, Y., Sun, Y., Liu, Z., Sarma, S.E., Bronstein, M.M., Solomon, J.M., 2019. Dynamic graph cnn for learning on point clouds. ACM Trans. Graph. (tog) 38 (5), 1–12.

Wikipedia Contributors, 2024. Spherical harmonics. https://en.wikipedia.org/wiki/Spherical_harmonics. (Accessed 8 May 2024).

Wu, H., Pan, J., Li, Z., Wen, Z., Qin, J., 2020. Automated skin lesion segmentation via an adaptive dual attention module. IEEE Trans. Med. Imaging 40 (1), 357–370.

Wu, T., Pan, L., Zhang, J., Wang, T., Liu, Z., Lin, D., 2021. Density-aware chamfer distance as a comprehensive metric for point cloud completion. arXiv:2111.12702.

Wu, C.-H., Tsai, W.-H., Chen, Y.-H., Liu, J.-K., Sun, Y.-N., 2017. Model-based orthodontic assessments for dental panoramic radiographs. IEEE J. Biomed. Health Inform. 22 (2), 545–551.

Wu, X., Wu, X., Luan, T., Bai, Y., Lai, Z., Yuan, J., 2024. FSC: Few-point shape completion. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 26077–26087.

Xiang, P., Wen, X., Liu, Y.-S., Cao, Y.-P., Wan, P., Zheng, W., Han, Z., 2021. SnowflakeNet: Point cloud completion by snowflake point deconvolution with skip-transformer. arXiv:2108.04444.

Yang, Y., Feng, C., Shen, Y., Tian, D., 2018. FoldingNet: Point cloud auto-encoder via deep grid deformation. arXiv:1712.07262.

Yu, X., Rao, Y., Wang, Z., Liu, Z., Lu, J., Zhou, J., 2021. PoinTr: Diverse point cloud completion with geometry-aware transformers. arXiv:2108.08839.

Yu, X., Rao, Y., Wang, Z., Lu, J., Zhou, J., 2023. AdaPoinTr: Diverse point cloud completion with adaptive geometry-aware transformers. arXiv:2301.04545.

Yuan, F., Dai, N., Tian, S., Zhang, B., Sun, Y., Yu, Q., Liu, H., 2020. Personalized design technique for the dental occlusal surface based on conditional generative adversarial networks. Int. J. Numer. Methods Biomed. Eng. 36 (5), e3321. http://dx.doi.org/10.1002/cnm.3321, URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/cnm.3321. arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1002/cnm.3321.

Yuan, W., Khot, T., Held, D., Mertz, C., Hebert, M., 2019. PCN: Point completion network. arXiv:1808.00671.

Zhu, H., Jia, X., Zhang, C., Liu, T., 2022. 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. Springer-Verlag, Berlin, Heidelberg, pp. 161–172. http://dx.doi.org/10.1007/978-3-031-05981-0_13.