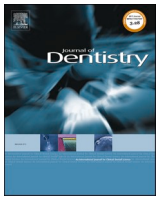
Journal of Dentistry 145 (2024) 104988

Contents lists available at ScienceDirect

Journal of Dentistry

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



Review article

Generative deep learning approaches for the design of dental restorations: A narrative review

Alexander Broll [a], Markus Goldhacker [b], Sebastian Hahnel [a], Martin Rosentritt [a,*]

[a] Department of Prosthetic Dentistry, University Hospital Regensburg, Regensburg, Germany
[b] Faculty of Mechanical Engineering, OTH Regensburg, Regensburg, Germany

ARTICLE INFO

Keywords:
Tooth reconstruction
Dental prosthesis design
Deep Learning
Digital dentistry

ABSTRACT

Objectives: This study aims to explore and discuss recent advancements in tooth reconstruction utilizing deep learning (DL) techniques. A review on new DL methodologies in partial and full tooth reconstruction is conducted.

Data/Sources: PubMed, Google Scholar, and IEEE Xplore databases were searched for articles from 2003 to 2023.

Study selection: The review includes 9 articles published from 2018 to 2023. The selected articles showcase novel DL approaches for tooth reconstruction, while those concentrating solely on the application or review of DL methods are excluded. The review shows that data is acquired via intraoral scans or laboratory scans of dental plaster models. Common data representations are depth maps, point clouds, and voxelized point clouds. Reconstructions focus on single teeth, using data from adjacent teeth or the entire jaw. Some articles include antagonist teeth data and features like occlusal grooves and gap distance. Primary network architectures include Generative Adversarial Networks (GANs) and Transformers. Compared to conventional digital methods, DL-based tooth reconstruction reports error rates approximately two times lower.

Conclusions: Generative DL models analyze dental datasets to reconstruct missing teeth by extracting insights into patterns and structures. Through specialized application, these models reconstruct morphologically and functionally sound dental structures, leveraging information from the existing teeth. The reported advancements facilitate the feasibility of DL-based dental crown reconstruction. Beyond GANs and Transformers with point clouds or voxels, recent studies indicate promising outcomes with diffusion-based architectures and innovative data representations like wavelets for 3D shape completion and inference problems.

Clinical significance: Generative network architectures employed in the analysis and reconstruction of dental structures demonstrate notable proficiency. The enhanced accuracy and efficiency of DL-based frameworks hold the potential to enhance clinical outcomes and increase patient satisfaction. The reduced reconstruction times and diminished requirement for manual intervention may lead to cost savings and improved accessibility of dental services.

1. Introduction

In recent years, computer aided design (CAD) and computer aided manufacturing (CAM) technologies have gained wide adoption in digital systems for dental restorations [1,2]. In contrast to conventional manual restoration methods, CAD/CAM systems introduce a significant reduction in labor intensity for both technicians and dentists as well as an improvement in reconstruction quality. Fully integrated chairside systems facilitate an end-to-end, in-house process for smaller restorations, involving intraoral scanning of the patient's dentition, the design of the fixed dental prosthesis (FDP), as well as the final fabrication.

The dental technician initiates the process by utilizing a proposal generated by the CAD software. Subsequently, they engage in a comprehensive re-design, re-adaptation, or re-modeling of the restoration. This intricate task involves considering prosthetic needs, addressing specific requests related to the case, and drawing upon years of accumulated experience. A crucial aspect of the restoration process involves preserving the proper function of the masticatory system. This includes dynamic contacts during excursive movements of the lower jaw with teeth in contact and static contacts in terminal occlusion. Motion-

* Corresponding author.

https://doi.org/10.1016/j.jdent.2024.104988
Received 23 January 2024; Received in revised form 13 March 2024; Accepted 3 April 2024
Available online 11 April 2024

tracking technologies allow for dynamic registration of mandible movements to capture the patient-specific function and hence simulate individual articulation [3,4]. Integration of the individual mandible movements in the used CAD software enables the assessment of occlusal contacts and further manual geometry adaptation to refine the restoration and ensure optimal results.

The manual adjustment of the restoration is susceptible to errors due to the complexity and individuality of the masticatory system. Consequently, restorations may induce occlusal discomfort, potentially causing temporomandibular disorders or the failure of the FDP [5,6]. Adaptation and reworking by the dentist are often necessary, resulting in increased treatment cost and duration. A reduction in overall stability due to the partial weakening of the material leads to increased roughness, surface damage, and the potential occurrence of fractures within the FDP.

The development of a completely automated reconstruction process using DL holds promise in addressing these issues and eliminating the necessity for dedicated motion-tracking systems. This premise relies on the assumption that the patient's individual function is implicitly evident in the intraoral scan data of the current clinical situation and can be captured by DL-based methods.

The aim of this study was to review and discuss recent advances in tooth reconstruction using DL techniques. An overview of DL methods in partial and full tooth reconstruction is provided.

2. Materials and methods

This section outlines the methodology employed for the review. The eligibility criteria, search strategy, data collection process, and outcomes are shown.

2.1. Eligibility criteria

The review process included articles published from 2003 to 2023. Eligible articles showcase novel approaches for tooth reconstruction. Those concentrating solely on the application or review of these methods are excluded. The review encompasses both conventional digital and DL-based approaches. The focus is on DL-based methods, as they represent the most recent advancements in the field.

2.2. Search strategy

The search was conducted in PubMed, Google Scholar, and IEEE Xplore databases. Ten search groups were used to build search queries. Each of the groups contained synonyms or related terms for the respective topic. These terms were concatenated with the boolean operator "OR". Different combinations of the search groups were concatenated with the boolean operator "AND".

The search groups were: Prosthesis, Implant, Dentistry, Artificial Intelligence, Machine Learning, Generative Architectures, Generative Architectures (short terms), Generative [Title], Reconstruction [Title], Method [Title/Abstract]. Square brackets in the search group title indicate that the terms are used as title or title/abstract search terms.

The articles were screened for eligibility based on title and abstract. Articles were then evaluated based on full text. The search was extended by screening the references as well as articles by the same first or last authors.

2.3. Data collection process

The following information was extracted from each DL study and collected in a database: year, title, dimension (2D/3D), network architecture type, training datatype, participant criteria, data acquisition method, raw datatype, preprocessing, network and loss function structure (reconstruction framework), amounts of data: (training data, validation data, testing data), reconstruction datatype, postprocessing, metrics, quantitative results.

Conventional digital methods were included with year, title, quantitative results, and a summary of their methodology.

2.4. Outcomes

The search queries resulted in 613 articles. The review finally includes 9 DL articles published from 2018 to 2023 and 15 conventional digital articles published from 2004 to 2017.

3. Results

3.1. Computer aided tooth reconstruction

This section provides an overview of existing methods for the computer-aided reconstruction of a crown, outlining conventional digital and the more recent DL-based approaches.

3.1.1. Conventional digital approaches

Over the last two decades, numerous approaches for conventional digital tooth reconstruction have been proposed. These methods either rely on the calculation of a statistical morphable tooth model, mirroring symmetrical teeth, or deforming a standard tooth from a template library.

Statistical morphable tooth model

Mehl et al. [7] laid the foundation for many subsequent studies [8–13] by introducing an average tooth model. This morphable model of a tooth is based on the vector space representation of the average deviations that a set of identically classified teeth exhibit w.r.t. a reference tooth. The algorithm is rooted in optical flow analysis [14], initially intended for image processing.

Building upon the representation of an average tooth, Mehl et al. [9] proposed the statistical biogeneric tooth model. This model provides an approximation of an infinite set of different teeth based on deviations from the average tooth. The level and direction of these deviations are defined by a principal component analysis (PCA) of the vector space representation of all deviations [7]. Specifically, the weighted principal components are used to deform the average tooth. The reconstruction of a missing tooth involves calculating the weights of the principal components of the PCA for the known features of the target tooth. The missing points are then reconstructed by a linear combination of the average tooth and the weighted principal components based on a reduced number of feature points.

Blanz et al. [8] employed a similar approach, emphasizing the mathematical solution to efficiently calculate the optimum for the proposed reconstruction problem. Subsequent studies, such as Mehl et al. [10] and Richter et al. [11], adapted the biogeneric tooth model. Variations of the biogeneric tooth model with additional modifications, such as a collision detection mechanism [12] or the inclusion of occlusal spatial constraints [13], were proposed.

Mirroring

The second approach is based on the mirroring of symmetrical teeth [15,16]. This method relies on the assumption that the teeth of the upper and lower jaw are symmetrical. The reconstruction of a missing tooth is achieved by mirroring the corresponding tooth on the opposite side of the jaw. Subsequently, the mirrored tooth is aligned w.r.t. the new spatial occlusal constraints [15]. Zhang et al. [16] refined this approach by employing a semi-automatic method for the segmentation of the jaw and matching of symmetrical teeth. The pose (position and orientation) of the mirrored tooth is then optimized to match the target, and dental alignment is performed by a digital dental occlusion reestablishment algorithm [17] with the reconstructed complete dental model.

Standard tooth deformation

The following studies [18–23] center around the deformation of a standard tooth derived from a template library. In this process, the standard tooth is adjusted to match the configuration of the target tooth.

Song et al. [18] described the standard tooth utilizing a non-uniform rational B-spline (NURBS) surface fitting technique [24], a widely employed method in CAD applications for the creation of smooth surfaces. Subsequently, the outcome is locally fine-tuned to adhere to the new spatial constraints. Yoo et al. [19] employed 3D mesh deformation through direct manipulation of free-form deformation [25] and multiple wires deformation [26]. Their approach incorporated a functionally guided plane to account for the occlusion of the patient. Steinbrecher et al. [20] introduced Laplacian surface editing [27] for the construction of inlays and onlays. The extraction of the cavity line on the target tooth, essential for constructing the boundaries of the inlay, involved a hybrid approach combining a 2D mask image with ray collisions. Jiang et al. [22] extended the same methodology to full crowns. Preceding the deformation, salient features were extracted using a Morse function [28]. Zheng et al. [21] established pairs of feature points to establish a spatial mapping between the template tooth and the prepared target tooth. A radial basis function was then employed to achieve the deformation of the tooth surface. Zhang et al. [23] outlined a four-step process for their inlay reconstruction. Initially, a heuristic searching algorithm was utilized to extract the cavity contour. Subsequently, the tooth surface was segmented into the inside and outside of the cavity. Following this, the occlusal surface was reconstructed using a dual-factor constrained deformation. Finally, the inlay was modeled by merging the inner surface and the occlusal surface.

3.1.2. Deep learning based approaches

In contrast to the conventional digital methodologies, DL-based approaches refrain from the explicit definition of a tooth model or a template tooth. Instead, they leverage a deep neural network model to discern the inherent tooth morphology from a designated set of training data. Generative DL models, in particular, aim to capture the details within a given dataset by acquiring knowledge about the underlying patterns, structures, and statistical properties. These models can produce new samples that closely resemble the training data, thereby implicitly learning the distinctive features inherent in the dataset. The generative capabilities of the models extend beyond reconstructing the training data, enabling the generation and reconstruction of previously unseen samples. It is noteworthy that the initial training process from scratch demands a significantly larger volume of data compared to conventional digital methods. The application of transfer learning, which involves fine-tuning certain sections of a pre-trained network, can mitigate the requirement for extensive amounts of data.

Recent advances in general generative DL underscore the potential of this approach in both 2D applications [29–32] and, more recently, in the realm of 3D applications [33–39]. These advances have led to the emergence of generative DL-based approaches for the solution to the tooth reconstruction problem [40–49]. The following sections provide a detailed overview of these studies, including used data, data processing, and results. Table 1 provides a comprehensive comparison of all DL-based methods that will be discussed in this review article.

3.2. Data

This section gives a comprehensive examination of the data employed in the reviewed studies, offering insights into the differences in the data acquisition process and the final data types. The acquisition of data stands as a pivotal phase in the preparation for the training process of DL models.

3.2.1. Data acquisition

The selection criteria for participants exhibit variation across different studies. Raw analog data typically manifests in two forms: either as a plaster model representing the patient's dentition [43–45,47] or as the patient's dentition itself [40,46,48].

Plaster models undergo scanning using laboratory grade 3D scanners [43–45,47]. The resulting digital data is subsequently processed to generate a 3D mesh object. In contrast, the patient's dentition is directly digitalized with intraoral 3D scanners [48]. Although certain articles omit specifics about the analog data type and scanner employed [41,49], the provided information and figures suggest a scanning process occurred, leading to the presumption that the raw digital data format in all studies is a 3D mesh object.

Selection criteria for suitable patients are not consistently detailed across studies [40,44,46,49]. Some studies base participant selection on age [43,45], the presence of specific teeth, and the absence of diseases affecting the specified teeth, as well as their antagonists and adjacent teeth [41,48]. Notably, Chau et al. [47] define precise inclusion and exclusion criteria at both participant and tooth levels. The total number of samples utilized for training, validation, and testing for each study ranges from 169 [47] to 6109 [46]. It is noteworthy that the quantity of teeth each sample encompasses varies across the studies and ranges from three teeth to a whole jaw. The total count of healthy teeth included in each dataset, and hence the complexity of acquiring such a dataset, is not directly analogous.

3.2.2. Data type

Originating from a 3D mesh object as the primary input format, the typical transformation involves converting the data into a format that is suited as a direct input to the reconstruction network.

2D approach

The prevalence and accessibility of 2D generative network architectures designed for image inference [29–32] have led to the adoption of a 2D depth map representation for the data [40,41,43–45]. Nevertheless, the suitability of the 2D depth map representation is constrained by the inherent characteristics of the data, limiting its applicability to a specific maximum quantity of teeth prevalent in the input data. Specifically, studies employing 2D data have focused on reconstructing the occlusal surface of a single tooth with a maximum of three teeth per

Table 1
Comparison of the DL methodologies. The teeth are numbered according to FDI.

	Dimension	Network	Data aquisition	Data representation	n_{train}	n_{val}	n_{test}	Input Teeth	Target Tooth
Hwang et al. [40]	2D	GAN	intraoral scan	depth map $\in \mathbb{R}^{w \times h}$	1500	1570	243	35-37 ∧ 25-27	36
Yuan et al. [41]	2D	GAN	n/a	depth map $\in \mathbb{R}^{w \times h}$	500	n/a	100	35-37 ∧ 25-27	36
Tian et al. [43]	2D	GAN	plaster model scan	depth map $\in \mathbb{R}^{256 \times 256}$	750	n/a	80	35-37 ∧ 25-27, 45-47 ∧ 15-17	36, 46 (4 inlays, 1 onlay)
Tian et al. [45]	2D	GAN	plaster model scan	depth map $\in \mathbb{R}^{256 \times 256}$	850	90	60	35-37 ∧ 25-27, 45-47 ∧ 15-17	36, 46
Tian et al. [44]	2D	GAN	plaster model scan	depth map $\in \mathbb{R}^{256 \times 256}$	700	n/a	80	35-37 ∧ 25-27, 45-47 ∧ 15-17	36, 46
Zhu et al. [46]	3D	Transformer	intraoral scan	point cloud $\in \mathbb{R}^{3 \times 2048}$	5469	n/a	640	upper jaw, lower jaw	arbitrary
Chau et al. [47]	3D	GAN	plaster model scan	voxel $\in \{0,1\}^{x \times y \times z}$	159	n/a	10	upper jaw	16
Ding et al. [48]	3D	GAN	intraoral scan	voxel $\in \{0,1\}^{x \times y \times z}$	600	n/a	12	44-46	45
Feng et al. [49]	3D	Transformer	n/a	point cloud $\in \mathbb{R}^{3 \times 1024}$	480	n/a	120	11-13 ∧ 21-23	11-13, 21-23

input image. Examples include studies targeting the molar 36 [40,41] or 36, 46 [43–45]. Tian et al. [43] concentrated on partial reconstructions involving four types of inlays and one type of onlay. All employed methods are limited to the reconstruction of one tooth at a time. The target tooth preparations were complemented by their two adjacent teeth in the same depth map, along with their antagonists and the gap distance between the two jaws in terminal occlusion [40,41,43]. Each added feature is depicted in a separate image. Tian et al. [44,45] expanded the input data by incorporating extracted dental biological morphology, such as the occlusal groove and the occlusal fingerprint [50]. The ground truth (GT) in these studies comprised single crowns designed by a technician. In cases where the input data encompasses a substantial quantity of teeth, such as an entire upper or lower jaw, the 2D representation becomes impractical due to inherent data loss. This loss is attributed to undercuts in the data that cannot be adequately captured through projection with a plane.

3D approach

Consequently, 3D representation becomes essential. While Zhu et al. [46] and Feng et al. [49] explicitly specify the use of 3D point clouds as the data format, Chau et al. [47] and Ding et al. [48] only indicate the use of 3D data. The employed 3D-GAN architecture [51] necessitates voxelized point data as its input and is therefore assumed to be the used data format. The teeth selected for reconstruction vary from an arbitrary tooth [46] to specific molars [47,48] or incisors [49]. In contrast to the 2D methods, the input data is no longer confined to the occlusal surfaces of a limited number of teeth but allows for encompassment of the entire upper or lower jaw [46,47]. Ding et al. [48] did not increase the amount of data compared to the 2D methods. Feng et al. [49] expanded the dataset to include eight incisors. No study has utilized the entire upper and lower jaw as input data, resulting in information containment limited to the teeth that are included in the dataset. The reconstruction process is likely to derive advantages from incorporating additional teeth as the network can leverage the information contained in the remaining teeth to reconstruct the target tooth.

3.3. Data processing

The data processing steps can be divided into preprocessing, reconstruction framework, and postprocessing (Fig. 1). The preprocessing steps (1) are defined as the steps required to transform the raw data into the input format for the learning-based reconstruction methods. The reconstruction framework (2) refers to the architecture of the employed network, including the utilized loss functions. The postprocessing steps (3) are defined as the steps required to transform the output of the network back to a properly positioned and aligned 3D mesh. The mesh can be used for further processing or fabrication.

3.3.1. Preprocessing

2D approach

For 2D methods, the outcome is consistently a 2D depth map of the input. The general approach in most studies involves orienting the data and projecting it onto a plane with a fixed grid (Fig. 2). However, the exact methodology behind the dimensionality reduction step is not always explicitly stated [40].

Yuan et al. [41] define the plane as parallel to the occlusal plane of the partial jaw, with a distance threshold d_{thresh} to exclude points beyond this threshold. The grid is fixed at 256×256, corresponding to the pixel count of the final image. The height resolution of the projection, represented by the grayscale pixel values, ranges from 0 to 255 (8-bit). This grid size and image bit depth remain constant across most studies utilizing 2D data [43–45]. Tian et al. [43,45] specify the orientation of the projection plane w.r.t. the data as parallel to its bounding box. The bounding box, a cuboid with minimal volume encompassing all object points, is orientation-invariant. Tian et al. [45] further refine the orientation of the plane through additional steps. They define a feature point $\boldsymbol{p}_{\text{feat},i}\ \forall\, i \in \{0,1,2\}$ for each tooth and calculate the normal vector $\boldsymbol{n}_1 = (\boldsymbol{p}_{\text{feat},1} - \boldsymbol{p}_{\text{feat},0}) \times (\boldsymbol{p}_{\text{feat},2} - \boldsymbol{p}_{\text{feat},0})$. Subsequently, the object is

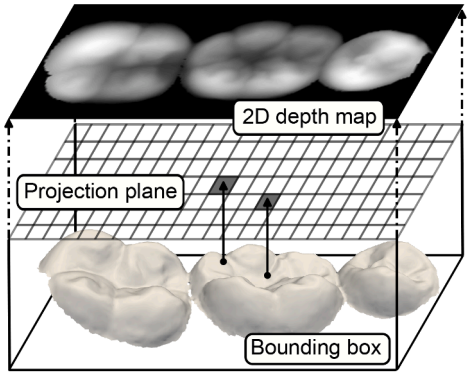


Fig. 2. 2D projection of 3D scan data. A projection plane with a fixed grid is defined to be parallel to the data's bounding box. Each point on the tooth is projected onto a 2D depth map with values ranging from 0 to 255 (8-bit). The height of each point determines its value on the depth map. A brighter value corresponds to higher areas on the 3D data.

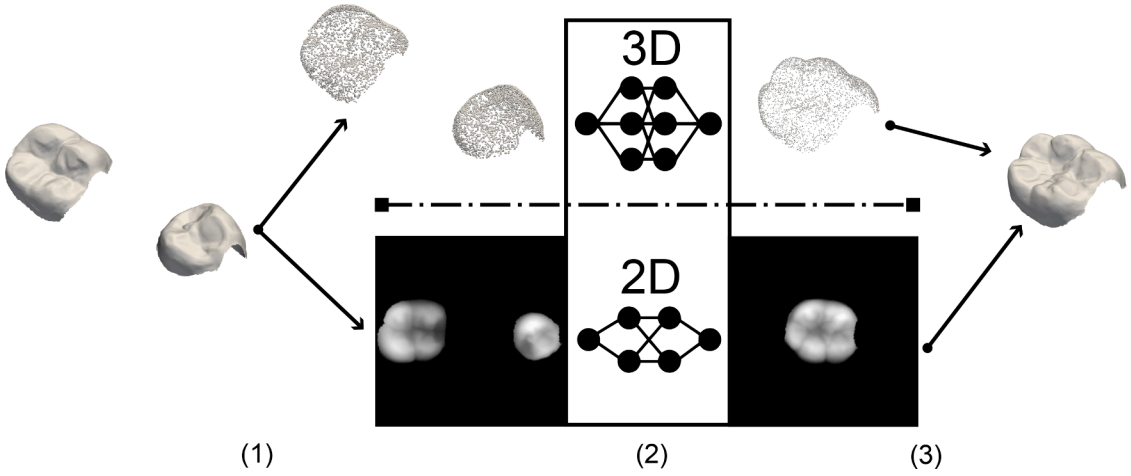


Fig. 1. DL data processing. (1) Preprocessing to transform the raw data (mesh) into the input format (point cloud, depth map) for the learning-based reconstruction methods. (2) Reconstruction framework. (3) Postprocessing to transform the output of the network back to a properly positioned and aligned 3D mesh. (The quantity of teeth displayed is not indicative of all methods and is provided merely for illustrative purposes.)

repositioned such that $\boldsymbol{p}_{\text{feat},1} = [0,0,0]^{\text{T}}$ and $\boldsymbol{n}_1$ is aligned with the global z_0-axis. After calculating the bounding box of the body, the object is rotated around the z_0-axis, aligning the global y_0-axis with the y_{b}-axis of the bounding box. The projection plane is then aligned parallel to the bounding box. The specific parameters for the projection are adjusted to maximize the entropy of the resulting image.

3D approach

In the case of 3D methods, data preprocessing is performed to streamline the training process. There is no necessity for dimensionality reduction. In the study by Chau et al. [47], where the entire upper or lower jaw serves as input data, alignment was achieved by orienting the respective occlusal planes parallel to one of the reference coordinate system planes. The specific definition of the occlusal plane is not further elucidated. Zhu et al. [46] and Ding et al. [48] do not provide details regarding the preprocessing of their data. Nevertheless, the 3D-GAN architecture [51] they employed necessitates voxelized point data as its input, implying that voxelization is likely part of the preprocessing. The voxelization process involves dividing 3D space into a grid of equally sized cubes (voxels) and assigning a binary value to each voxel. The resulting voxelized point cloud is a 3D matrix $\boldsymbol{P}_{\text{vox}} \in \{0,1\}^{x \times y \times z}$ with x, y, z dimensions corresponding to the grid.

In the work of Feng et al. [49], the requirement for decoupled features, specifically position and orientation, is essential for their reconstruction framework. The points $\boldsymbol{P}_{O,B}$ are defined as a transformation of the object's points represented in its own body O_B-coordinate system to the global O_O-coordinate system. This transformation is expressed as $\boldsymbol{P}_{O,B} = \boldsymbol{T}_{O,B}\boldsymbol{P}_{B,B}$. The transformation $\boldsymbol{T}_{O,B}$ is determined using the singular value decomposition (SVD) of the covariance matrix of the object points $\boldsymbol{P}_{O,B}$, i.e. the PCA. The eigenvectors span a new body coordinate system O_B. Subsequently, the transformation $\boldsymbol{T}_{O,B}$ is defined as the transformation matrix from the O_B-coordinate system to the global O_O-coordinate system. The transformation is defined with 3 Euler angles and a 3D translation vector. This reduces the number of variables from 16 to 6.

3.3.2. *Reconstruction framework*

The subsequent section offers a detailed overview of the network architectures utilized in the studies. Existing architectures for the reconstruction and generation of non-dental images and 3D shapes have shown promising results. Given the complex characteristics of the human dentition, the reconstruction of occlusal surfaces presents unique challenges. The morphology of the restoration is not only governed by explicit spatial requirements dictated by its antagonist, gap distance, and two adjacent teeth but also by implicit functional requirements derived from the data of the remaining teeth.

2D approach

While all 2D architectures [40,41,43–45] are rooted in the general concept of a GAN [52], some studies explicitly state the use of the pix2pix [31] Conditional Generative Adversarial Network (CGAN) architecture in their frameworks [40–43]. GANs are composed of two adversarial networks: a generator and a discriminator. The generator aims to produce samples that resemble the training data, while the discriminator's goal is to distinguish between the generator's output and the actual training data (Fig. 3).

Hwang et al. [40] extended the loss term of the pix2pix architecture by incorporating a spatially invariant functionality or histogram loss L_{hist} specifically designed for the reconstruction of the occlusal surface. The loss is calculated as the mean sum of all bin values in a histogram representing the error distances d_i between the GT and the generated image. The GT refers to a technician-designed crown. The loss is further enhanced by assigning variable weights to the bins. This ensures the generation of functional occlusal surfaces. Collisions, represented by negative distances $d_i \leq 0$, are therefore heavily penalized. Furthermore, the minimal 5% of distances are used to emphasize geometry deemed highly important for biting and chewing according to the authors. The final loss term is defined as the default CGAN-loss L_{cgan}, incorporating the additional histogram loss L_{hist} and an L1-loss L_{L1} between the GT and the generated image. The antagonist teeth and the gap distance between the two partial jaws serve as the conditional inputs to the network.

Although Yuan et al. [41] based their framework on the same architecture, they enhanced it by introducing new data and losses. To incorporate occlusal grooves, they implemented an occlusal groove filter network trained to extract these features from a full crown. The training data comprises both the original crowns and the occlusal grooves extracted by a technician. Similar to Hwang et al. [40], the pix2pix [31] loss is extended to include the L1-loss L_{L1} between the GT and the generated image. A perceptual loss L_{per} is introduced, utilizing the feature deviation of certain hidden layers of the discriminator as a quality measure. The inclusion of occlusal grooves into the loss term involves adding an L1-loss $L_{\text{L1,groove}}$ between the extracted occlusal grooves and the occlusal grooves calculated by the occlusal groove filter network. The final loss term is defined as the sum of the pix2pix loss L_{cgan}, perceptual loss L_{per}, L1-loss L_{L1}, and groove L1-loss $L_{\text{L1,groove}}$. The antagonist teeth serve as conditional inputs to the network, while the prepared partial jaw is used as the direct input to the generator.

Similar to Yuan et al. [41], Tian et al. [43] expanded their framework by introducing an occlusal groove filter network called GroNet to leverage the deviation of occlusal grooves as a quality metric $L_{\text{L1,groove}}$. Additionally, they adopted a dual discriminator approach. The local discriminator takes only the defective teeth as its input, while the global discriminator includes both the defective teeth and the adjacent teeth. Since the framework was designed for inlay generation, a certain

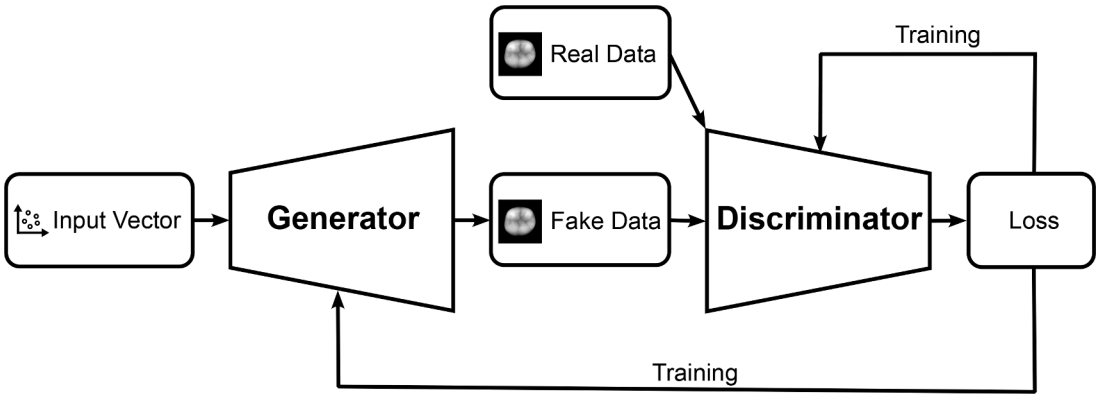


Fig. 3. Schematic training process of a GAN. This architecture consists of two competing networks: a generator and a discriminator. The generator network takes random noise as input and tries to generate data samples that resemble the real data as closely as possible. The discriminator network is trained to distinguish between real data samples from the dataset and fake data samples generated by the generator. It learns to differentiate between real and generated data. The discriminator's decision (data is fake/real) during the training process is used as the loss term for the optimizer to train both networks.

amount of the target tooth is retained. The rationale behind the dual discriminator approach is to ensure that the generated inlay surface is not only realistic but also consistent with the residual tooth surface and adjacent teeth. As mentioned earlier [41], the loss of the local discriminator was extended by a perceptual loss L_{per} to penalize the deviation of high-dimensional features. The objective function of the adversarial architecture was changed from the original Jensen-Shannon divergence [52] to the Wasserstein distance [53]. Finally, the histogram loss L_{hist}, introduced by Hwang et al. [40], is now included, resulting in a total of four new loss terms $L_{per}, L_{hist}, L_{L1}$, and $L_{L1,groove}$.

Tian et al. [45] maintained the dual discriminator approach [43] for the restoration of crowns. However, despite still utilizing occlusal grooves as inputs to the discriminators and the generator, the GroNet is no longer part of the framework, leading to the exclusion of the occlusal groove loss $L_{L1,groove}$. Additionally, the gap distance is no longer explicitly used in a dedicated loss term L_{hist} but is instead directly employed as an input to the framework. Similar to the previous approach, the local discriminator processes only the preparation tooth, the corresponding target crown, the antagonist, and the occlusal groove with an L1 loss L_{L1} and a newly introduced mean squared error term L_{mse}. The global discriminator includes the preparation tooth along with its adjacent teeth, the gap distance between the two jaws, the antagonist teeth, the occlusal groove, as well as the occlusal fingerprint [50]. This is done with the previously used perceptual loss L_{per} and the L1 loss L_{L1}. While the overall architectural approach remains similar to Tian et al. [43], the loss terms are reduced to L_{per}, L_{L1}, and L_{mse}.

The final 2D approach by Tian et al. [44] reintegrates the GroNet [43] into their newly proposed multi-stage framework. This framework is constructed upon a dual serialized CGAN approach. The first stage establishes the basic shape of an occlusal surface that satisfies spatial positional relationships, implementing the previously discussed L1-loss L_{L1} and perceptual loss L_{per} from its respective discriminator. The preparation tooth, the antagonist, the tooth type label, and the target crown excluding the occlusal fingerprint are used in conjunction with the gap distance as constraints to ensure a certain occlusal spatial relationship. The first stage results are then used to train the GroNet, which is fixed for the second stage. The second stage of the framework relies on the output of stage one as its input and aims to enhance the functional characteristics of the crown. Consequently, the second stage generator takes the stage one results and the occlusal fingerprint [50] as its input to generate the final crown. The loss terms of the second stage include the L1-loss L_{L1} and the GroNet loss L_{GroNet}.

3D approach

Similar to the 2D reconstruction case, GANs are also used as the foundation for 3D reconstruction frameworks [47,48]. The reported GAN architecture [51] is tailored for use with voxelized 3D data. However, no further details regarding the structure of the are provided. Other 3D methods [46,49] adopted a transformer network with a dynamic graph convolutional neural network (DGCNN) [54] for their multi-stage frameworks.

Zhu et al. [46] used furthest point sampling (FPS) [55] for the initial processing of the point cloud data of the target jaw with an arbitrary missing tooth, aiming to equalize the number of points for each jaw. The DGCNN [54] is utilized to group the point data and extract local features. Comprising graph convolutional layers that extract features from a point cloud, the DGCNN is a neural network architecture specifically designed for point cloud processing. These features are used to generate a latent representation of the point cloud. Two different multilayer perceptrons (MLPs) equalize the dimensions of the point sets and features, preparing them for input into the transformer network. The geometry-aware transformer network [56] extracts the latent vectors describing the geometry of the missing tooth. Another MLP generates a set of coarse tooth points $\boldsymbol{P}_{coarse}$ which is concatenated with the latent vectors, serving as input to the final multiscale generator network. This generator constructs a fine point cloud $\boldsymbol{P}_{fine}$ of the missing tooth in a coarse-to-fine manner. The final prediction is a combination of the coarse points $\boldsymbol{P}_{coarse}$ and the fine points $\boldsymbol{P}_{fine}$. Due to the unordered nature of a point cloud, a permutation invariant metric is needed for a final quality assessment of the generated crown. The reconstruction loss is calculated as the weighted sum of the chamfer distance (CD) [57] between $\boldsymbol{P}_{coarse}$, $\boldsymbol{P}_{fine}$, and the GT crown respectively. The weights are utilized to focus on the generation of tooth surfaces with logical feature points.

Feng et al. [49] introduced a novel approach to the reconstruction problem by decoupling the shape and pose of the input data. The framework comprises two networks: the pose estimation network (PEN) and the shape estimation network (SEN). The input data for the framework is always based on the seven existing incisors to reconstruct the missing tooth. The PEN functions as a regression network responsible for estimating the crown pose, utilizing a MLP. The six features of the data, representing their orientation with three Euler angles and position with a 3D translation vector, are extracted in advance. These features for the remaining teeth are then separately input into the same network with different weights to predict the position and orientation of the missing crown. The input dimension for the network is 21, with an output dimension of 3. During the training process, the loss L_{L1} of PEN is calculated based on the L1-norm. The SEN relies on a transformer encoder-decoder architecture and is employed to estimate the initial crown point cloud. It takes the pose-decoupled shapes $\boldsymbol{P}_{B,B}$ of the remaining teeth in their respective body coordinate systems as input, ordered according to their position in the jaw. Additionally, a unit sphere representing the missing tooth and the mirrored crown point cloud of the corresponding tooth on the opposite side are input into the network. This results in a total of nine sequenced point clouds. The loss of SEN is calculated based on the CD between the generated point cloud and the GT point cloud.

3.3.3. Postprocessing

2D approach

All 2D methods share a common final processing step. Although, this step is not always explicitly specified [40,41,44,45]. To create a 3D mesh object for further processing, an inverse of the employed 2D projection method must be applied to the generated 2D depth maps. Facing specific challenges in the fit of inlay restorations, Tian et al. [43] detailed a postprocessing routine to ensure a continuous transition between the generated inlay and the tooth preparation. This routine involves the back projection of the partial occlusal surface of the target tooth into 3D. The back projection is carried out using distance-gray mapping, which is inversely performed during preprocessing (Fig. 2), along with region growing. The outer segmented surface of the generated inlay is then merged with the internal segmented preparation area of the target tooth. The internal surface is offset by a given tolerance to simulate the adhesive/cement layer. To ensure a consistent distance between the internal surface boundary and its corresponding occlusal surface boundary, Laplacian deformation [58] is applied. The final inlay is then obtained by merging the occlusal surface and the internal surface using mesh stitching [20].

3D approach

The 3D methods [46–49] also require postprocessing to generate an adequately positioned and oriented 3D mesh object from the raw output point clouds. For this purpose, the studies employ different methods. Zhu et al. [46] outlined a four-step point cloud surface reconstruction process. The process commences with mean sampling of the points to prevent redundant meshes. The k-nearest-neighbors for each point are then utilized to create the initial mesh triangles, which are stored as a signed distance field representation of the surface in an octree format. The isosurface with a signed distance of zero is subsequently extracted from the octree yielding the final mesh surface object. Chau et al. [47] and Ding et al. [48] did not provide information about the mesh conversion process. They specified the use of superimposition to align the generated data with the target jaw. Chau et al. [47] detailed the process

by declaring four occlusal landmarks used for the alignment process. These landmarks were set as the mesiobuccal, mesiolingual, and distobuccal cusp tips, as well as the distolingual cusp ridge. Due to the structure of their network architecture, Feng et al. [49] do not require an dedicated alignment process. The final position of the reconstruction in its body coordinate system is obtained by transforming the generated point cloud into the global coordinate system using the pose estimated by PEN. An iterative method is then used to form a mesh object based on the oriented and positioned point cloud.

3.4. Performance metrics and quantitative results

All studies evaluated the performance of their respective methods. The following section provides an overview of the utilized metrics and evaluation methods.

3.4.1. Metrics

The metrics used to evaluate the performance of the methods can be categorized into general and 3D-only metrics. Unlike the ordered nature of 2D images, 3D point clouds are unordered, necessitating permutation-invariant metrics for their evaluation. While two images can be compared pixel by pixel, the same is not possible for two point clouds. Consequently, the metrics used to evaluate the 2D methods are not directly applicable to the unordered data representations. Due to the wide variety of metrics used in the reviewed studies, Table 2 and Table 3 provide a brief overview. The function behind each metric, the purpose of its use, and its direction are described.

3.4.2. Quantitative results

Table 4 presents the quantitative outcomes of the metrics presented in Section 3.4.1. The results for 2D and 3D methods are presented independently. They are not suitable for direct comparison due to the diverse nature of the respective metrics. Proprietary metrics are notably isolated, as they are not conducive to comparison with other metrics.

It is crucial to highlight that all displayed results are as reported directly by the authors. It is evident that RMSE, both variations of SSIM, PSNR, and FSIM are the most commonly utilized 2D metrics. Hwang et al. [40] stand out by incorporating metrics specifically tailored to the dental use case. The reported RMSE values across three of five studies are in the vicinity of approximately 0.07 mm. Tian et al. [45] introduce an RMSE that adversely deviates from the other methods by a factor of approximately 3, while also consistently reporting higher values for the other metrics. Tian et al. [43] also show a deviation from the other studies in the reported RMSE with a factor of approximately 2. These discrepancies are likely attributed to the utilization of a distinct data representation for the calculation of the metrics. A noteworthy factor contributing to this divergence can be the proportional size of the crown w.r.t. the entire image. The authors of the best-performing 2D studies report overall similar results.

A direct comparison of the 3D results is not feasible at this point due to a lack of overlap in the employed evaluation metrics. Nevertheless, the inclusion of all results serves the purpose of completeness, providing a comprehensive overview for future research endeavors.

4. Discussion

The review shows that the primary architectures employed to address the problem are GANs and Transformers. In the 2D domain, GANs are exclusively utilized. The data acquisition methods are roughly evenly split between intraoral scans and laboratory scans of dental plaster models. The total sample sizes range from 169 to 6109, encompassing training, validation, and test data. Regarding data representation, 2D methods exclusively use depth maps, while 3D methods employ point clouds and voxelized point clouds equally. For detailed representation and memory-efficient storage of the data, point clouds are favored due to locally specific resolution.

Table 2
Summary of general and 3D evaluation metrics. ↓↑ indicate the direction of the metric, signifying whether a higher or lower value is advantageous.

Metric	Function	Purpose
	General metrics	
Root Mean Square Error (RMSE) [62] ↓	Measures average difference between corresponding pixel values in predicted and true images.	Suitable for applications where accurate pixel-wise reconstruction is critical.
Structural Similarity Index (SSIM) [63] ↑	Evaluates structural information, luminance, and contrast similarity between images.	Reflects human perception of image quality, useful for compression or distortion assessment.
Multi-Scale SSIM (MS-SSIM) [64] ↑	Extends structural similarity index measure (SSIM) to multiple scales, capturing nuanced variations in image quality.	Useful for evaluating impact of image processing across different resolutions or scales.
Peak Signal-to-Noise Ratio (PSNR) [65] ↑	Measures image quality by comparing peak signal value to noise level.	Widely used for lossy image compression, provides intuitive measure of reconstruction fidelity.
Feature Similarity Index (FSIM) [66] ↑	Incorporates structural information for perceptually relevant image similarity.	Useful when structural content is crucial, e.g., in medical imaging or object recognition.
F-score [67] ↑	Combines precision and recall to quantify accuracy of object detection.	Commonly used in computer vision tasks for evaluating object detection algorithms.
Intersection over Union (IoU) [40] ↑	Measures the overlap of two objects with their respective bounding boxes.	Evaluates the localization and segmentation accuracy in computer vision tasks.
Mean Absolute Error (MAE) [49] ↓	Measures the average absolute deviation between the generated data and the GT.	Evaluates the overall accuracy of the generated data.
	3D metrics	
Chamfer Distance (CD) [57] ↓	Measures nearest average squared distance between points in predicted and true 3D point clouds.	Valuable for evaluating spatial alignment and accuracy of reconstructed 3D structures.
Earth Mover's Distance (EMD) [68] ↓	Quantifies dissimilarity between two point clouds by considering minimum cost of transformation.	Useful for comparing complex 3D shapes where direct point-to-point correspondence may not be straightforward.
Hausdorff Distance (HD) [69] ↓	Measures maximum distance between points in one point cloud to nearest point in the other.	Offers comprehensive measure of dissimilarity, relevant in medical imaging and shape matching.

In all cases, the size of the input data is larger than the output data. The reconstruction target is consistently limited to a single tooth, with six out of nine cases focusing on molar teeth. One case includes incisors, and only one method allows for the reconstruction of an arbitrary tooth. The reconstruction input data for all methods comprises a minimum of the target tooth's adjacent teeth, with a maximum of the whole respective upper or lower jaw. In three cases, antagonist teeth are included in both training and reconstruction data. Moreover, the data in four cases extends beyond the sole geometry of the teeth to include features such as occlusal grooves, occlusal fingerprints, and gap distance. There is no method that utilizes the entire dentition for the reconstruction process, thus benefiting from the full contextual information of the data.

The quantitative results demonstrate a high degree of similarity between the best performing 2D methods, with an RMSE in the near vicinity of 0.07 mm. In comparison to the MAE of the conventional digital methods for inlay restoration (0.125 mm to 0.15 mm [9,11]), the GAN-based 2D methods show an improvement in reconstruction quality. These conventional digital methods, akin to the 2D DL methods, utilized the occlusal surface of molar teeth. The reconstruction area for the inlay restorations was not explicitly reported, but even full crown

Table 3
Summary of propriatary evaluation metrics. ↓↑ indicate the direction of the metric, signifying whether a higher or lower value is advantageous. A heuristically determined ideal value for the metric is indicated by an $=^!$. The ideal values are based on the results of technician designed crowns.

Metric	Function	Purpose
	Propriatary metrics	
Penetration Rate (PR) [40] ↓	Measures relative amount of cases where the minimal gap distance was smaller than zero.	Comparison of the functional quality w.r.t. collisions.
Maximum Penetration Rate (MP) [40] ↓	Measures penetration depth for all cases where the gap distance was smaller than zero.	Comparison of the functional quality w.r.t. collisions.
Penetration Area (PA) [40] ↓	Measures the area where the gap distance is smaller than zero.	Comparison of the functional quality w.r.t. collisions.
Number of Clusters (NC) [40] $=^!$ 4.15	Measures the number of regions with neighboring pixels within a distance of two pixels.	Reflect the biting and chewing quality.
Spread (SPD) [40] $=^!$ 11.01	Measures the spatial standard deviation of critical pixels in a crown	Evaluates the contact points of the teeth. The contact points should not focus on one region.
Mean +/- Deviation (MPD/MND) [48] ↓	Measures the respective mean deviation between the generated and the GT point cloud.	Evaluates the overall accuracy of the generated point cloud.

reconstructions of the DL methods exhibit an improvement over the error for inlay restorations. The RMSE is sensitive to higher errors due to the squaring operation before taking the mean. Therefore, the RMSE would yield a higher value for the conventional digital methods, further enhancing the DL-based results. The MAE of the conventional digital methods was calculated based on 3D models of teeth, in contrast to the depth maps used by the DL-based methods. Feng et al. [49] reported an MAE of 0.13 mm for their full incisor 3D reconstruction method, surpassing the 2D and conventional digital methodologies that solely used occlusal surfaces in terms of reconstruction volume.

The similarity in errors among the 2D methods suggests that these methods may have reached a local equilibrium. Although the results show an improvement compared to existing conventional digital methods, there is still room for optimization. The limited amount of employable data underscores the need for further research, especially in the realm of 3D methodologies. Standardizing metrics for the comparison of different 3D methods is crucial for quantitative comparisons. A comparison is therefore currently not feasible.

While AI-based commercial software for dental crown restoration is emerging, further research is still needed. In particular, the incorporation of as much patient data as possible is essential. Clinical trials are needed to assess the performance of the methods in a real world setting. Restorations have to be generated based on a real patient's data and subsequently tested for fit and functionality. The results of these tests can then be used to further improve the methods.

With the current state of the art, the use of DL for dental crown reconstruction is feasible. The increasing adoption of digital techniques in dentistry will expand the available data, allowing for the training of more complex models and improving reconstruction quality. Furthermore, the application of DL for dental crown reconstruction is not confined to GANs and Transformers in combination with point clouds or voxels. Recent studies suggest promising results using diffusion-based architectures and data representations such as wavelets for 3D shape completion and inference problems [59–61].

5. Conclusions

The key insights derived from the analyzed articles are as follows:

- Acquisition of data is performed through intraoral scans or laboratory scans of dental plaster models.
- Sample sizes range from 169 to 6109.
- The reconstruction problem is approached from a 2D and 3D perspective.
- Depth maps, point clouds, and voxelized point clouds are the prevalent data representations.

Table 4
Quantitative results of the DL methods. The shown mean results are rounded to three decimal places. If available, the standard deviation is shown in parentheses. The pixel value RMSE by Tian et al. [43] is converted to mm by multiplication with 7/255 mm to conform with the other studies. The ordered voxel representation of Ding et al. [48] allows for the calculation of the RMSE in 3D. ↓↑ indicate the direction of the metric, signifying whether a higher or lower value is advantageous.

2D approach							
General metrics							
	RMSE (mm) ↓	SSIM ↑	MS-SSIM ↑	PSNR ↑	FSIM ↑	F-score ↑	IoU ↑
Hwang [40]	0.066	n/a	n/a	n/a	n/a	0.942	0.920
Yuan [41]	0.070	n/a	0.813	23.304	0.879	n/a	n/a
Tian [43]	0.196	n/a	0.820	29.250	0.890	n/a	n/a
Tian [45]	0.135 (0.117)	0.985 (0.005)	n/a	34.264 (1.228)	0.993 (0.008)	n/a	n/a
Tian [44]	0.068 (0.008)	0.845 (0.005)	n/a	23.304 (0.168)	0.887 (0.007)	n/a	n/a
Propriatary metrics							
	PR ↓	MP ↓	PA ↓	NC $=^!$ 4.15	SPD $=^!$ 11.01		
Hwang [40]	9.470	2.300	10.0	3.140	9.630		
3D approach							
General metrics							
	RMSE (mm) ↓	CD-L1 (mm) ↓	CD-L2 (mm) ↓	EMD (mm) ↓	HD (mm) ↓		
Zhu [46]	n/a	4.473×10^{-3}	0.055×10^{-3}	0.019	n/a		
Chau [47]	n/a	n/a	n/a	n/a	0.633 (0.961)		
Ding [48]	0.361 (0.116)	n/a	n/a	n/a	n/a		
Propriatary metrics							
	MAE (mm) ↓	MPD (mm) ↓	MND (mm) ↓				
Ding [48]	n/a	0.250 (0.050)	0.311 (0.122)				
Feng [49]	0.130	n/a	n/a				

- The focus of reconstructions is on individual teeth, leveraging data from neighboring teeth or the entire jaw.
- The number of teeth contained in the input data is larger than the number of reconstructed teeth.
- Only some studies incorporate data from antagonist teeth and features such as occlusal grooves and gap distance.
- The primary network architectures utilized are GANs and Transformers.
- When compared to conventional digital methods, tooth reconstruction using DL shows approximately half the error rate.

CRediT authorship contribution statement

Alexander Broll: Conceptualization, Investigation, Writing – original draft. **Markus Goldhacker:** Writing – review & editing. **Sebastian Hahnel:** Writing – review & editing. **Martin Rosentritt:** Writing – review & editing.

Declaration of competing interest

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

References

[1] F. Yuan, Y. Sun, Y. Wang, P. Lv, Computer-aided design of tooth preparations for automated development of fixed prosthodontics, Comput. Biol. Med. 44 (2014) 10–14, https://doi.org/10.1016/j.compbiomed.2013.10.019.

[2] H. Watanabe, C. Fellows, H. An, Digital Technologies for Restorative Dentistry, Dental Clin. North Am. 66 (4) (2022) 567–590, https://doi.org/10.1016/j.cden.2022.05.006.

[3] L. Lepidi, F. Grande, G. Baldassarre, C. Suriano, J. Li, S. Catapano, Preliminary clinical study of the accuracy of a digital axiographic recording system for the assessment of sagittal condylar inclination, J. Dentis. 135 (2023) 104583, https://doi.org/10.1016/j.jdent.2023.104583.

[4] Z. Nagy, A. Mikolicz, J. Vag, In-vitro accuracy of a novel jaw-tracking technology, J. Dentis. 138 (2023) 104730, https://doi.org/10.1016/j.jdent.2023.104730.

[5] V. Preis, T. Dowerk, M. Behr, C. Kolbeck, M. Rosentritt, Influence of cusp inclination and curvature on the in vitro failure and fracture resistance of veneered zirconia crowns, Clin. Oral Investigat. 18 (3) (2014) 891–900, https://doi.org/10.1007/s00784-013-1029-9.

[6] K. Schnitzhofer, A. Rauch, M. Schmidt, M. Rosentritt, Impact of the occlusal contact pattern and occlusal adjustment on the wear and stability of crowns, J. Dentis. 128 (2023) 104364, https://doi.org/10.1016/j.jdent.2022.104364.

[7] A.C. Mehl, V. Blanz, R. Hickel, A new mathematical process for the calculation of average forms of teeth, J. Prosthetic Dentis. 94 (6) (2005) 561–566, https://doi.org/10.1016/j.prosdent.2005.10.002.

[8] V. Blanz, A.C. Mehl, T. Vetter, H.P. Seidel, A statistical method for robust 3D surface reconstruction from sparse data. IEEE 2nd International Symposium on 3D Data Processing, Visualization and Transmission (3DPVT), IEEE, 2004, pp. 293–300, https://doi.org/10.1109/TDPVT.2004.1335212.

[9] A.C. Mehl, V. Blanz, R. Hickel, Biogeneric tooth: a new mathematical representation for tooth morphology in lower first molars, Eur. J. Oral Sci. 113 (4) (2005) 333–340, https://doi.org/10.1111/j.1600-0722.2005.00224.x.

[10] A.C. Mehl, V. Blanz, New procedure for fully automatic occlusal surface reconstruction by means of a biogeneric tooth model, J. Prosthetic Dentis. 95 (5) (2006) 400, https://doi.org/10.1016/j.prosdent.2006.02.016.

[11] M.J. Richter, A.C. Mehl, Evaluation for the fully automatic inlay reconstruction by means of the biogeneric tooth model, Int. J. Comput. Dentis. 9 (2) (2006) 101–111.

[12] J. Sporring, K. Hommelhoff Jensen, Bayes Reconstruction of Missing Teeth, J. Math. Imag. Vis. 31 (2-3) (2008) 245–254, https://doi.org/10.1007/s10851-008-0081-6.

[13] C. Zhang, Statistical Reconstruction Algorithm for Restoring Broken Tooth Surface Based on Occlusion Spatial Constraint, J. Mech. Eng. 52 (1) (2016) 165, https://doi.org/10.3901/JME.2016.01.165.

[14] V. Blanz, T. Vetter, A morphable model for the synthesis of 3D faces, in: A. Rockwood (Ed.), Proceedings of the 26th annual conference on Computer graphics and interactive techniques, Computer graphics Annual conference series, ACM Press, New York, NY, 1999, pp. 187–194, https://doi.org/10.1145/311535.311556.

[15] F.A. Probst, A.C. Mehl, CAD reconstruction using contralateral mirrored anterior teeth: a 3-dimensional metric and visual evaluation, Int. J. Prosthodont. 21 (2008) 521.

[16] J. Zhang, J.J. Xia, J. Li, X. Zhou, Reconstruction-based digital dental occlusion of the partially edentulous dentition, IEEE J. Biomed. Health Inf. 21 (1) (2017) 201–210, https://doi.org/10.1109/JBHI.2015.2500191.

[17] Y.-B. Chang, J.J. Xia, J. Gateno, Z. Xiong, X. Zhou, S.T.C. Wong, An automatic and robust algorithm of reestablishment of digital dental occlusion, IEEE Trans. Med. Imag. 29 (9) (2010) 1652–1663, https://doi.org/10.1109/TMI.2010.2049526.

[18] Y.-L. Song, J. Li, L. Yin, T. Huang, P. Gao, The feature-based posterior crown design in a dental CAD/CAM system, Int. J. Adv. Manuf. Technol. 31 (11-12) (2007) 1058–1065, https://doi.org/10.1007/s00170-005-0289-1.

[19] K.-H. Yoo, J.-S. Ha, J.-S. Yoo, Modeling Inlay/Onlay prostheses with mesh deformation techniques. 7th International Conference on Computational Science (ICCS), Springer, 2007, pp. 154–157, https://doi.org/10.1007/978-3-540-72586-2_21.

[20] T. Steinbrecher, M. Gerth, Dental Inlay and onlay construction by iterative laplacian surface editing, Comput. Graph. Forum 27 (5) (2008) 1441–1447, https://doi.org/10.1111/j.1467-8659.2008.01284.x.

[21] S.-X. Zheng, J. Li, Q.-F. Sun, A novel 3D morphing approach for tooth occlusal surface reconstruction, Comput.-Aided Des. 43 (3) (2011) 293–302, https://doi.org/10.1016/j.cad.2010.11.003.

[22] X. Jiang, N. Dai, X. Cheng, J. Wang, Q. Peng, H. Liu, C. Cheng, Robust tooth surface reconstruction by iterative deformation, Comput. Biol. Med. 68 (2016) 90–100, https://doi.org/10.1016/j.compbiomed.2015.11.001.

[23] C. Zhang, T. Liu, W. Liao, T. Yang, L. Jiang, Computer-aided design of dental inlay restoration based on dual-factor constrained deformation, Adv. Eng. Softw. 114 (2017) 71–84, https://doi.org/10.1016/j.advengsoft.2017.06.005.

[24] P.J. Hartley, C.J. Judd, Parametrization and shape of B-spline curves for CAD, Comput.-Aided Des. 12 (5) (1980) 235–238, https://doi.org/10.1016/0010-4485(80)90028-7.

[25] W.M. Hsu, J.F. Hughes, H. Kaufman, Direct manipulation of free-form deformations, ACM SIGGRAPH Compu. Graphic. 26 (2) (1992) 177–184, https://doi.org/10.1145/142920.134036.

[26] S. Karan, E. Fiume, Wires: a geometric deformation technique. 25th annual conference on Computer graphics and interactive techniques, 1998, pp. 405–414.

[27] Y. Lipman, O. Sorkine, D. Cohen-Or, D. Levin, C. Rossl, H.P. Seidel, Differential coordinates for interactive mesh editing. Shape Modeling Applications, IEEE, 2004, https://doi.org/10.1109/smi.2004.1314505.

[28] P.-T. Bremer, H. Edelsbrunner, B. Hamann, V. Pascucci, A topological hierarchy for functions on triangulated surfaces, IEEE Trans. Visual. Comput. Graph. 10 (4) (2004) 385–396, https://doi.org/10.1109/tvcg.2004.3.

[29] T. Karras, S. Laine, M. Aittala, J. Hellsten, J. Lehtinen, T. Aila, Analyzing and Improving the Image Quality of StyleGAN. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, 2020, pp. 8110–8119, https://doi.org/10.1109/cvpr42600.2020.00813.

[30] T. Karras, M. Aittala, J. Hellsten, S. Laine, J. Lehtinen, T. Aila, Training generative adversarial networks with limited data. 34th International Conference on Neural Information Processing Systems (NIPS), 2020, pp. 12104–12114.

[31] P. Isola, J.-Y. Zhu, T. Zhou, A.A. Efros, Image-to-image translation with conditional adversarial networks. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, 2017, https://doi.org/10.1109/cvpr.2017.632.

[32] R. Rombach, A. Blattmann, D. Lorenz, P. Esser, B. Ommer, High-Resolution Image Synthesis with Latent Diffusion Models, 2021, 2112.10752.

[33] J. Zhang, K. Zhou, Y. Luximon, T.-Y. Lee, P. Li, MeshWGAN: mesh-to-mesh wasserstein GAN with multi-task gradient penalty for 3d facial geometric age transformation, IEEE Trans. Visual. Comput. Graph. PP (2023), https://doi.org/10.1109/TVCG.2023.3284500.

[34] X. Yu, Y. Rao, Z. Wang, J. Lu, J. Zhou, AdaPoinTr: diverse point cloud completion with adaptive geometry-aware transformers, 2023, http://arxiv.org/pdf/2301.04545.pdf. 2301.04545v1.

[35] S. Bipasha, A. Agarwal, G. Singh, B. Brojeshwar, S. Sridhar, M. Krishna, SCARP: 3D shape completion in arbitrary poses for improved grasping. IEEE International Conference on Robotics and Automation (ICRA), IEEE, 2023, pp. 3838–3845, https://doi.org/10.1109/ICRA48891.2023.10160365.

[36] Y. Li, Y. Dou, X. Chen, B. Ni, Y. Sun, Y. Liu, F. Wang, Generalized deep 3D shape prior via part-discretized diffusion process. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, 2023, https://doi.org/10.1109/cvpr52729.2023.01610.

[37] Y. Li, J. Wang, Y. Wang, R. Xiang, Y. Wang, Y. Pan, TGNet: learning 3D shape from sparse and incomplete point cloud. 9th International Conference on Virtual Reality (ICVR), IEEE, 2023, pp. 64–70, https://doi.org/10.1109/ICVR57957.2023.10169375.

[38] S. Hong, M. Yavartanoo, R. Neshatavar, K.M. Lee, ACL-SPC: adaptive closed-loop system for self-supervised point cloud completion. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, 2023, https://doi.org/10.1109/cvpr52729.2023.00910.

[39] Y.-C. Cheng, H.-Y. Lee, S. Tulyakov, A. Schwing, L. Gui, SDFusion: multimodal 3D shape completion, reconstruction, and generation. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, 2023, https://doi.org/10.1109/cvpr52729.2023.00433.

[40] J.-J. Hwang, S. Azernikov, A.A. Efros, S.X. Yu, Learning beyond human expertise with generative models for dental restorations, 2018, 1804.00064v1.

[41] F. Yuan, N. Dai, S. Tian, B. Zhang, Y. Sun, Q. Yu, H. Liu, Personalized design technique for the dental occlusal surface based on conditional generative adversarial networks, Int. J. Numer. Method. Biomed. Eng. 36 (5) (2020) e3321, https://doi.org/10.1002/cnm.3321.

[42] H. Lee, N. Thoummala, H. Park, S. Ham, J. Yu, J. Hwang, S.-M. Heo, S.H. Kim, AI-based dental prostheses fabrication using generative adversarial networks, Quant. Bio-Sci. 40 (1) (2021) 39–44.

[43] S. Tian, M. Wang, F. Yuan, N. Dai, Y. Sun, W. Xie, J. Qin, Efficient computer-aided design of dental inlay restoration: a deep adversarial framework, IEEE Trans.

Medical Imag. 40 (9) (2021) 2415–2427, https://doi.org/10.1109/tmi.2021.3077334.
[44] S. Tian, M. Wang, N. Dai, H. Ma, L. Li, L. Fiorenza, Y. Sun, Y. Li, DCPR-GAN: dental crown prosthesis restoration using two-stage generative adversarial networks, IEEE J. Biomed. Health Inf. 26 (1) (2022) 151–160, https://doi.org/10.1109/jbhi.2021.3119394.
[45] S. Tian, R. Huang, Z. Li, L. Fiorenza, N. Dai, Y. Sun, H. Ma, A dual discriminator adversarial learning approach for dental occlusal surface reconstruction, J. Healthcare Eng. 2022 (2022) 1–14, https://doi.org/10.1155/2022/1933617.
[46] H. Zhu, X. Jia, C. Zhang, T. Liu, ToothCR: a two-stage completion and reconstruction approach on 3D dental model. 26th Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD), Springer, [S.l.], 2022, pp. 161–172, https://doi.org/10.1007/978-3-031-05981-0_13.
[47] R.C.W. Chau, R.T.-C. Hsung, C. McGrath, E.H.N. Pow, W.Y.H. Lam, Accuracy of artificial intelligence-designed single-molar dental prostheses: a feasibility study, J. Prosthetic Dentis. (2023), https://doi.org/10.1016/j.prosdent.2022.12.004.
[48] H. Ding, Z. Cui, E. Maghami, Y. Chen, J.P. Matinlinna, E.H.N. Pow, A.S.L. Fok, M. F. Burrow, W. Wang, J.K.H. Tsoi, Morphology and mechanical performance of dental crown designed by 3D-DCGAN, Dental Mater. 39 (3) (2023) 320–332, https://doi.org/10.1016/j.dental.2023.02.001.
[49] Y. Feng, B. Tao, J. Fan, S. Wang, J. Mo, Y. Wu, Q. Liang, 3D reconstruction for maxillary anterior tooth crown based on shape and pose estimation networks, Int. J. Comput. Assist. Radiol. Surg. 18 (8) (2023) 1405–1416, https://doi.org/10.1007/s11548-023-02841-1.
[50] O. Kullmer, S. Benazzi, L. Fiorenza, D. Schulz, S. Bacso, O. Winzen, Technical note: Occlusal fingerprint analysis: quantification of tooth wear pattern, Am. J. Phys. Anthropol. 139 (4) (2009) 600–605, https://doi.org/10.1002/ajpa.21086.
[51] J. Wu, C. Zhang, T. Xue, B. Freeman, J. Tenenbaum, Learning a probabilistic latent space of object shapes via 3D generative-adversarial modeling. 30th International Conference on Neural Information Processing Systems (NIPS), 2016, pp. 82–90.
[52] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio, Generative adversarial nets, Adv. Neural Inf. Process. Syst. 27 (2014).
[53] Q. Yang, P. Yan, Y. Zhang, H. Yu, Y. Shi, X. Mou, M.K. Kalra, Y. Zhang, L. Sun, G. Wang, Low-Dose CT image denoising using a generative adversarial network with wasserstein distance and perceptual loss, IEEE Trans. Med. Imag. 37 (6) (2018) 1348–1357, https://doi.org/10.1109/TMI.2018.2827462.
[54] Y. Wang, Y. Sun, Z. Liu, S.E. Sarma, M.M. Bronstein, J.M. Solomon, Dynamic graph CNN for learning on point clouds, 2018, 1801.07829v2.
[55] R.Q. Charles, H. Su, M. Kaichun, L.J. Guibas, PointNet: deep learning on point sets for 3D classification and segmentation. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, 2017, https://doi.org/10.1109/cvpr.2017.16.
[56] X. Yu, Y. Rao, Z. Wang, Z. Liu, J. Lu, J. Zhou, PoinTr: diverse point cloud completion with geometry-aware transformers. IEEE/CVF International Conference on Computer Vision (ICCV), IEEE, 2021, https://doi.org/10.1109/iccv48922.2021.01227.
[57] H. Fan, H. Su, L. Guibas, A point set generation network for 3D object reconstruction from a single image. 30th IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, Piscataway, NJ, 2017, pp. 2463–2471, https://doi.org/10.1109/CVPR.2017.264.
[58] M. Botsch, O. Sorkine, On linear variational surface deformation methods, IEEE Trans. Visual. Comput. Graph. 14 (1) (2008) 213–230, https://doi.org/10.1109/TVCG.2007.1054.
[59] K.-H. Hui, R. Li, J. Hu, C.-W. Fu, Neural wavelet-domain diffusion for 3D shape generation, in: S.K. Jung (Ed.), SIGGRAPH Asia, ACM Digital Library, Association for Computing Machinery, New York,NY,United States, 2022, pp. 1–9, https://doi.org/10.1145/3550469.3555394.
[60] X. Zeng, A. Vahdat, F. Williams, Z. Gojcic, O. Litany, S. Fidler, K. Kreis, LION: latent point diffusion models for 3D shape generation, Adv. Neural Inf. Process. Syst. 35 (2022) 10021–10039.
[61] L. Wu, D. Wang, C. Gong, X. Liu, Y. Xiong, R. Ranjan, R. Krishnamoorthi, V. Chandra, Q. Liu, Fast point cloud generation with straight flows. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, 2023, https://doi.org/10.1109/cvpr52729.2023.00911.
[62] H. Zhang, T. Xu, H. Li, S. Zhang, X. Wang, X. Huang, D.N. Metaxas, StackGAN++: realistic image synthesis with stacked generative adversarial networks, IEEE Trans. Pattern Anal. Mach. Intell. 41 (8) (2019) 1947–1962, https://doi.org/10.1109/TPAMI.2018.2856256.
[63] S.U. Dar, M. Yurt, L. Karacan, A. Erdem, E. Erdem, T. Cukur, Image synthesis in multi-contrast MRI with conditional generative adversarial networks, IEEE Trans. Med. Imag. 38 (10) (2019) 2375–2388, https://doi.org/10.1109/TMI.2019.2901750.
[64] Z. Wang, A.C. Bovik, H.R. Sheikh, E.P. Simoncelli, Image quality assessment: from error visibility to structural similarity, IEEE Trans. Image Process. 13 (4) (2004) 600–612, https://doi.org/10.1109/tip.2003.819861.
[65] Y. Luo, K. Chen, L. Liu, J. Liu, J. Mao, G. Ke, M. Sun, Dehaze of cataractous retinal images using an unpaired generative adversarial network, IEEE J. Biomed. Health Inf. 24 (12) (2020) 3374–3383, https://doi.org/10.1109/JBHI.2020.2999077.
[66] L. Zhang, L. Zhang, X. Mou, D. Zhang, FSIM: a feature similarity index for image quality assessment, IEEE Trans. Image Process. : Publ. Ieee Signal Process. Soc. 20 (8) (2011) 2378–2386, https://doi.org/10.1109/tip.2011.2109730.
[67] A. Knapitsch, J. Park, Q.-Y. Zhou, V. Koltun, Tanks and temples: benchmarking large-scale scene reconstruction, ACM Trans. Graph. 36 (4) (2017) 1–13, https://doi.org/10.1145/3072959.3073599.
[68] Y. Rubner, C. Tomasi, L. Guibas, A metric for distributions with applications to image databases. Sixth International Conference on Computer Vision (ICCV), Narosa Publishing House, [S.l.]n[s.n.], 1998, pp. 59–66, https://doi.org/10.1109/ICCV.1998.710701.
[69] N. Aspert, D. Santa-Cruz, T. Ebrahimi, MESH: measuring errors between surfaces using the Hausdorff distance. IEEE International Conference on Multimedia and Expo (ICME), IEEE, 2002, https://doi.org/10.1109/icme.2002.1035879.