



From Mesh Completion to AI Designed Crown

Golriz Hosseinimanesh¹(✉), Farnoosh Ghadiri², Francois Guibault¹,
Farida Cheriet¹, and Julia Keren³

¹ Polytechnique Montréal University, Montreal, Canada

{golriz.hosseinimanesh, francois.guibault, farida.cheriet}@polymtl.ca

² Centre d'intelligence artificielle appliquée (JACOBb), Montreal, Canada

farnoosh.ghadiri@jacobb.ca

³ Intellident Dentaire Inc., Montreal, Canada

info@kerenor.ca

Abstract. Designing a dental crown is a time-consuming and labor-intensive process. Our goal is to simplify crown design and minimize the tediousness of making manual adjustments while still ensuring the highest level of accuracy and consistency. To this end, we present a new end-to-end deep learning approach, coined Dental Mesh Completion (DMC), to generate a crown mesh conditioned on a point cloud context. The dental context includes the tooth prepared to receive a crown and its surroundings, namely the two adjacent teeth and the three closest teeth in the opposing jaw. We formulate crown generation in terms of completing this point cloud context. A feature extractor first converts the input point cloud into a set of feature vectors that represent local regions in the point cloud. The set of feature vectors is then fed into a transformer to predict a new set of feature vectors for the missing region (crown). Subsequently, a point reconstruction head, followed by a multi-layer perceptron, is used to predict a dense set of points with normals. Finally, a differentiable point-to-mesh layer serves to reconstruct the crown surface mesh. We compare our DMC method to a graph-based convolutional neural network which learns to deform a crown mesh from a generic crown shape to the target geometry. Extensive experiments on our dataset demonstrate the effectiveness of our method, which attains an average of 0.062 Chamfer Distance. The code is available at: <https://github.com/Golriz-code/DMC.git>

Keywords: Mesh completion · Transformer · 3D shape generation

1 Introduction

If a tooth is missing, decayed, or fractured, its treatment may require a dental crown. Each crown must be customized to the individual patient in a process, as depicted in Fig. 1. The manual design of these crowns is a time-consuming and labor-intensive task, even with the aid of computer-assisted design software.

Designing natural grooves and ensuring proper contact points with the opposing jaw present significant challenges, often requiring technicians to rely on trial and error. As such, an automated approach capable of accelerating this process and generating crowns with comparable morphology and quality to that of a human expert would be a groundbreaking advancement for the dental industry.

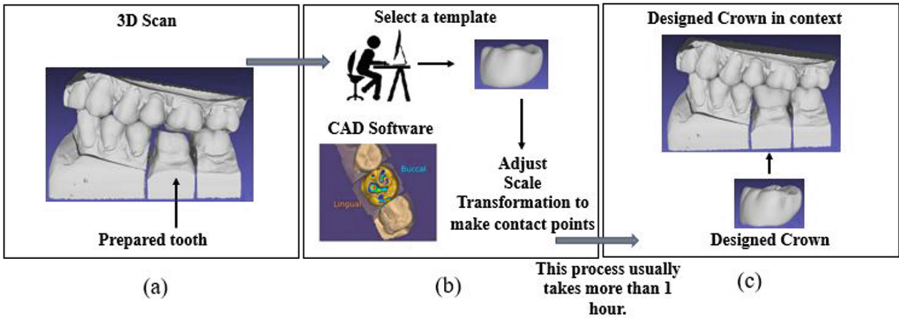


Fig. 1. Dental crown design process: a) Dentist prepares the tooth; b) Technician designs the crown; c) Dentist places the crown on the prepared tooth.

A limited number of studies have focused on how to automate dental crown designs. In [2,3], a conditional generative adversarial based network (GAN) is applied to a 2D depth image obtained from a 3D scan to generate a crown for a prepared tooth. Depth images created from a 3D scan can be used directly with 2D convolutional neural networks (CNNs) such as pix2pix [4], which are well-established in computer vision. However, depth images are limited in their ability to capture fine-grained details and can suffer from noise and occlusion issues. By contrast, point clouds have the advantage of being able to represent arbitrary 3D shapes and can capture fine-grained details such as surface textures and curvatures. [5,8] use point cloud-based networks to generate crowns in the form of 3D point clouds. Input point clouds used by [5] are generated by randomly removing a tooth from a given jaw; then the network estimates the missing tooth by utilizing a feature points-based multi-scale generating network. [8] propose a more realistic setting by generating a crown for a prepared tooth instead of a missing tooth. They also incorporate margin line information extracted from the prepared tooth in their network to have a more accurate prediction in the margin line area. The crowns generated by both approaches are represented as point clouds, so another procedure must convert these point clouds into meshes. Creating a high-quality mesh that accurately represents the underlying point cloud data is a challenging task which is not addressed by these two works. [6] proposed a transformer-based network to generate a surface mesh of the crown for a missing tooth. They use two separate networks, one responsible for generating a point cloud and the other for reconstructing a mesh given the crown generated by the first network. Similar to [5,8], the point completion network

used by [6] only uses the Chamfer Distance (CD) loss to learn crown features. This metric’s ability to capture shape details in point clouds is limited by the complexity and density of the data.

Although all aforementioned methods are potentially applicable to the task of dental crown design, most of them fail to generate noise-free point clouds, which is critical for surface reconstruction. One way to alleviate this problem is to directly generate a crown mesh. In [28], the authors develop a deep learning-based network that directly generates personalized cardiac meshes from sparse contours by iteratively deforming a template mesh, mimicking the traditional 3D shape reconstruction method. To our knowledge, however, the approach in [28] has not been applied to 3D dental scans.

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

2 Related Work

In the field of 3D computer vision, completing missing regions of point clouds or meshes is a crucial task for many applications. Various methods have been proposed to tackle this problem. Since the introduction of PointNet [13,14], numerous methods have been developed for point cloud completion [12]. The recent works PoinTr [11] and SnowflakeNet [16] leverage a transformer-based architecture with geometry-aware blocks to generate point clouds. It is hypothesized that using transformers preserves detailed information for point cloud completion. Nonetheless, the predicted point clouds lack connections between points, which complicates the creation of a smooth surface for mesh reconstruction.

Mesh completion methods are usually useful when there are small missing regions or large occlusions in the original mesh data. Common approaches based on geometric priors, self-similarity, or patch encoding can be used to fill small missing regions, as demonstrated in previous studies [26,27], but are not suitable for large occlusions. [18] propose a model-based approach that can capture the variability of a particular shape category and enable the completion of large missing regions. However, the resulting meshes cannot achieve the necessary

precision required by applications such as dental crown generation. Having a mesh prior template can also be a solution to generate a complete mesh given a sparse point cloud or a mesh with missing parts. In [28], cardiac meshes are reconstructed from sparse point clouds using several mesh deformation blocks. Their network can directly generate 3D meshes by deforming a template mesh under the guidance of learned features.

We combine the advantages of point cloud completion techniques with a differentiable surface reconstruction method to generate a dental mesh. Moreover, we used the approach in [28] to directly produce meshes from 3D dental points and compared those results with our proposed method.

3 Method

3.1 Network Architecture

Our method is an end-to-end supervised framework to generate a crown mesh conditioned on a point cloud context. The overview of our network is illustrated in Fig. 2. The network is characterized by two main components: a transformer encoder-decoder architecture and a mesh completion layer. The following sections explain each part of the network.

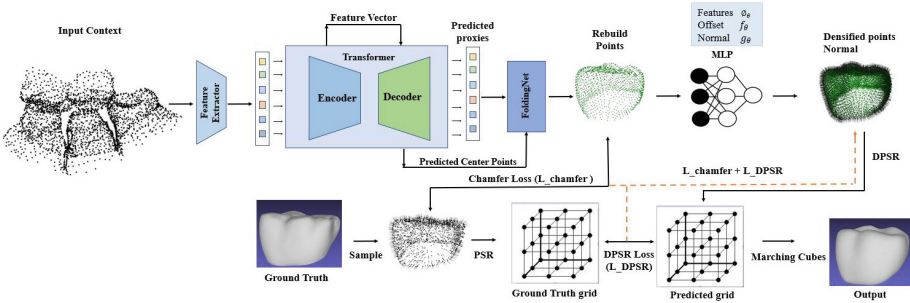


Fig. 2. Pipeline of our proposed network.

Transformer Encoder-Decoder. We adapt the transformer encoder-decoder architecture from [11] to extract global and local 3D features from our input (context) using the encoder and generate crown points via the decoder. A dynamic graph convolution network (DGCNN) [29] is used to group the input points into a smaller set of feature vectors that represent local regions in the context. The generated feature vectors are then fed into the encoder with a geometry-aware block. This block is used to model the local geometric relationships explicitly. The self-attention layer in the encoder updates the feature vectors using both long-range and short-range information. The feature vectors are further updated by a multi-layer perceptron. The decoder's role is to reason about the crown

based on the learnable pairwise interactions between features of the input context and the encoder output. The decoder incorporates a series of transformer layers with a self-attention and cross-attention mechanisms to learn structural knowledge. The output of the transformer decoder is fed into a folding-based decoder [15] to deform a canonical 2D grid onto the underlying 3D surface of the crown points.

Mesh Completion Layer. In this stage, to directly reconstruct the mesh from the crown points, we use a differentiable Poisson surface reconstruction (DPSR) method introduced by [17]. We reconstruct a 3D surface as the zero level set of an indicator function. The latter consists in a regular 3D point grid associated with values indicating whether a point is inside the underlying shape or not. To compute this function, We first densify the input unoriented crown points. This is done by predicting additional points and normals for each input point by means of an MLP network. After upsampling the point cloud and predicting normals, the network solves a Poisson partial differential equation (PDE) to recover the indicator function from the densified oriented point cloud. We represent the indicator function as a discrete Fourier basis on a dense grid (of resolution 128^3) and solve the Poisson equation (PDE) with the spectral solver method in [17].

During training, we obtain the estimated indicator grid from the predicted point cloud by using the differentiable Poisson solver. We similarly acquire the ground truth indicator grid on a dense point cloud sampled from the ground truth mesh, together with the corresponding normals. The entire pipeline is differentiable, which enables the updating of various elements such as point offsets, oriented normals, and network parameters during the training process. At inference time, we leverage our trained model to predict normals and offsets using Differentiable Poisson Surface Reconstruction (DPSR) [17], solve for the indicator grid, and finally apply the Marching Cubes algorithm [22] to extract the final mesh.

Loss Function. We use the mean Chamfer Distance (CD) [20] to constrain point locations. The CD measures the mean squared distance between two point clouds S_1 and S_2 . Individual distances are measured between each point and its closest point in the other point set, as described in Eq. (1). In addition, we minimize the L_2 distance between the predicted indicator function x and a ground truth indicator function x_{gt} , each obtained by solving a Poisson PDE [17] on a dense set of points and normals. We can express the Mean Square Error (MSE) loss as Eq. (2), where $f_\theta(X)$ represents a neural network (MLP) with parameters θ conditioned on the input point cloud X , D is the training data distribution, along with indicator functions x_i and point samples X_i on the surface of shapes. The sum of the CD and MSE losses is used to train the overall network.

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

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

4 Experimental Results

4.1 Dataset and Preprocessing

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

The first step in the preprocessing was to generate a context from a given 3D scan. To determine the context for a specific prepared tooth, we employed a pre-trained semantic segmentation model [21] to separate the 3D scan into 14 classes representing the tooth positions in each jaw. From the segmentations, we extracted the two adjacent and three opposing teeth of a given prepared tooth, as well as the surrounding gum tissue. To enhance the training data, we conducted data augmentation on the entire dental context, which included the master arch, opposing arch, and shell, treated as a single entity. Data augmentation involved applying 3D translation, scaling, and rotation, thereby increasing the training set by a factor of 10. To enrich our training set, we randomly sampled 10,240 cells from each context to form the input during training. We provide two types of ground truth: mesh and point cloud crowns. To supervise network training using the ground truth meshes, we calculate the gradient from a loss on an intermediate indicator grid. We use the spectral method from [17] to compute the indicator grid for our ground truth mesh.

4.2 Implementation Details

We adapted the architecture of [11] for our transformer encoder-decoder module. For mesh reconstruction, we used Differentiable Poisson Surface Reconstruction (DPSR) from [17]. All models were implemented in PyTorch with the AdamW optimizer [23], using a learning rate of $5e-4$ and a batch size of 16. Training the model took 400 epochs and 22 h on an NVIDIA A100 GPU.

4.3 Evaluation and Metrics

To evaluate the performance of our network and compare it with point cloud-based approaches, we used the Chamfer distance to measure the dissimilarity between the predicted and ground truth points. We employed two versions of CD: CD_{L_1} uses the L_1 -norm, while CD_{L_2} uses the L_2 -norm to calculate the distance between two sets of points. Additionally, we used the F-score [24] with a distance threshold of 0.3, chosen based on the distance between the predicted and ground truth point clouds. We also used the Mean Square Error (MSE) loss [17] to calculate the similarity between the predicted and ground truth indicator grids or meshes.

We conducted experiments to compare our approach with two distinct approaches from the literature, as shown in Table 1. The first such approach, PoinTr+margin line [8], uses the PoinTr [11] point completion method as a baseline and introduces margin line information to their network. To compare

our work to [8], we used its official implementation provided by the author. In the second experiment, PoinTr+graph, we modified the work of [28] to generate a dental crown mesh. To this end, we use deformation blocks in [28] to deform a generic template mesh to output a crown mesh under the guidance of the learned features from PoinTr. The deformation module included three Graph Convolutional Networks (GCNs) as in [9].

Table 1. Comparison of proposed method (DMC) with two alternate methods. Evaluation metrics: CD_{L1} , CD_{L2} ; MSE on output meshes; F-Score@0.3 .

Method	CD-L1 (\downarrow)	CD-L2 (\downarrow)	MSE (\downarrow)	$F1^{0.3}$ (\uparrow)
PoinTr + margin line [8]	0.065	0.018		0.54
PoinTr + graph	1.99	1.51		0.08
DMC	0.0623	0.011	0.0028	0.70

All experiments used the same dataset, which included all tooth positions, and were trained using the same methodology. To compare the results of the different experiments, we extracted points from the predicted meshes of our proposed network (DMC), as illustrated in Fig. 3. Table 1 shows that DMC outperforms the two other networks in terms of both CD and F-score. PoinTr+graph achieves poor CD and F-score results compared to the other methods. While the idea of using graph convolutions seems interesting, features extracted from the point cloud completion network don’t carry enough information to deform the template into an adequate final crown. Therefore, these methods are highly biased toward the template shape and need extensive pre-processing steps to scale and localize the template. In the initial two experiments, the MSE metric was not applicable as it was calculated on the output meshes. Figure 4 shows the visual results obtained from our proposed network (DMC). Furthermore, Fig. 5 presents a visual comparison of mesh surfaces generated by various methods for a sample molar tooth.

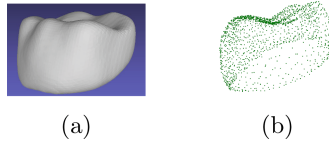


Fig. 3. Crown mesh predicted by DMC (a) and extracted point set (b).

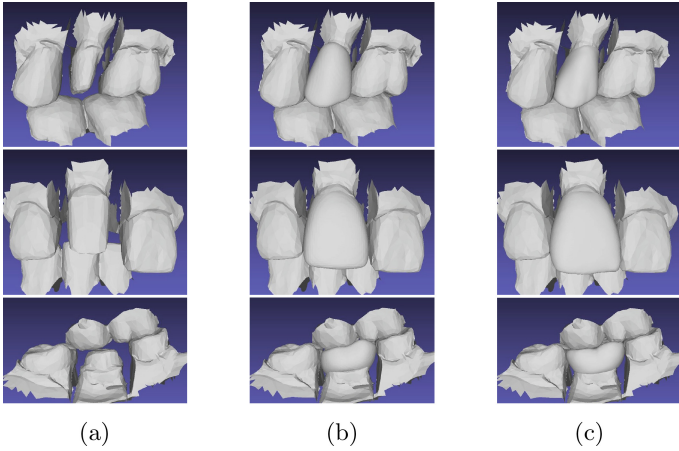


Fig. 4. Examples of mesh completions by the proposed architecture (DMC). a) Input context containing master arch, prepped tooth and opposing arch; b) Generated mesh in its context; c) Ground truth mesh in its context.

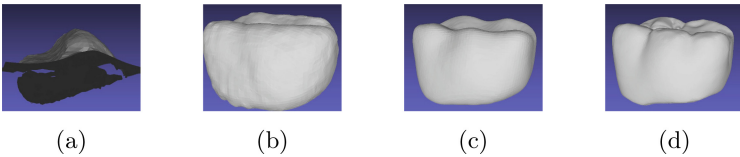


Fig. 5. Qualitative comparison of crown mesh generation approaches: a) Standard Poisson surface reconstruction; b) PoinTr [11] for point cloud and Shape as points [17] for mesh; c) Proposed method (DMC); d) Ground truth shape.

Table 2. Results of ablation study. Metrics are the same as in Table 1.

Method	CD-L1 (\downarrow)	CD-L2 (\downarrow)	MSE (\downarrow)	$F1^{0.3}$ (\uparrow)
PoinTr [11]	0.070	0.023		0.24
PoinTr + SAP	0.067	0.021	0.031	0.50
DMC without MSE	0.0641	0.015		0.65
DMC (Full Model)	0.0623	0.011	0.0028	0.70

4.4 Ablation Study

We conducted an ablation study to evaluate the components of our architecture. We started with the baseline PoinTr [8], a point completion method. To enhance it, we integrated Shape as Points (SAP) [17] as a separate network for mesh reconstruction from the PoinTr-generated point cloud. Next, we tested our proposed method (DMC) by excluding the Mean Square Error (MSE) loss function. Finally, we assessed DMC’s performance, including the MSE loss function. The results, shown in Table 2, demonstrate the consistent improvements achieved by our full model across all evaluation metrics.

5 Conclusion

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

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

References

1. <https://www.healthdirect.gov.au/dental-crown-procedure> (2021)
2. Hwang, J.-J., Azernikov, S., Efros, A.A., Yu, S.X.: Learning beyond human expertise with generative models for dental restorations (2018)
3. Yuan, F., et al.: Personalized design technique for the dental occlusal surface based on conditional generative adversarial networks. *Int. J. Numer. Methods Biomed. Eng.* **36**(5), e3321 (2020)
4. Isola, P., Zhu, J.-Y., Zhou, T., Efros, A.A.: Image-to-image translation with conditional adversarial networks. In: *CVPR* (2017)
5. Lessard, O., Guibault, F., Keren, J., Cheriet, F.: Dental restoration using a multi-resolution deep learning approach. In: *IEEE 19th International Symposium on Biomedical Imaging, (ISBI)* (2022)
6. Zhu, H., Jia, X., Zhang, C., Liu, T.: ToothCR: a two-stage completion and reconstruction approach on 3D dental model. In: Gama, J., Li, T., Yu, Y., Chen, E., Zheng, Y., Teng, F. (eds.) *Advances in Knowledge Discovery and Data Mining. PAKDD 2022. LNCS*, vol. 13282. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-05981-0_13
7. Ping, Y., Wei, G., Yang, L., Cui, Z., Wang, W.: Self-attention implicit function networks for 3D dental data completion. *Comput. Aided Geometr. Design* **90**, 102026 (2021)
8. Hosseinimanesh, G., et al.: Improving the quality of dental crown using a transformer-based method. *Medical Imaging 2023: Physics of Medical Imaging*, vol. 12463. SPIE (2023)
9. Wang, N., Zhang, Y., Li, Z., Fu, Y., Liu, W., Jiang, Y.-G.: Pixel2Mesh: generating 3D mesh models from single RGB images. In: Ferrari, V., Hebert, M., Sminchisescu, C., Weiss, Y. (eds.) *ECCV 2018. LNCS*, vol. 11215, pp. 55–71. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-01252-6_4
10. Huang, Z., Yu, Y., Xu, J., Ni, F., Le, X.: PF-Net: point fractal network for 3D point cloud completion. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 7662–7670 (2020)

11. Yu, X., Rao, Y., Wang, Z., Liu, Z., Lu, J., Zhou, J.: PoinTr: diverse point cloud completion with geometry-aware transformers. In: Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 12498–12507
12. Fei, B., Yang, W., Chen, W., Li, Z., et al.: Comprehensive review of deep learning-based 3D point clouds completion processing and analysis. *arXiv Prepr. arXiv2203.03311* (2022)
13. Qi, C.R., Su, H., Mo, K., Guibas, L.J.: PointNet: deep learning on point sets for 3D classification and segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 652–660 (2017)
14. Qi, C.R., Yi, L., Su, H., Guibas, L.J.: PointNet++: deep hierarchical feature learning on point sets in a metric space. *Deep Hierarchical Feature Learning on Point Sets in a Metric Space* (2017)
15. Yang, Y., Feng, C., Shen, Y., Tian, D.: FoldingNet: point cloud auto-encoder via deep grid deformation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 206–215 (2018)
16. Xiang, P., et al.: SnowflakeNet: point cloud completion by snowflake point deconvolution with skip-transformer. In: Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 5499–5509 (2021). <https://doi.org/10.48550/arxiv.2108.04444>
17. Peng, S., Jiang, C., Liao, Y., Niemeyer, M., Pollefeys, M., Geiger, A.: Shape as points: a differentiable poisson solver. *Journal, Advances in Neural Information Processing Systems* (2021)
18. Litany, O., Bronstein, A., Bronstein, M., Makadia, A.: Deformable shape completion with graph convolutional autoencoders. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 1886–1895 (2017)
19. Hui, K.-H., Li, R., Hu, J., Fu, C.-W.: Neural template: topology-aware reconstruction and disentangled generation of 3D meshes. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 18572–18582 (2022)
20. Yuan, W., Khot, T., Held, D., Mertz, C., Hebert, M.: PCN: point completion network. In: 3DV, pp. 728–737 (2018)
21. Alshegri, A., et al.: Semi-supervised segmentation of tooth from 3D scanned dental arches. In: *Medical Imaging, SPIE* (2022)
22. Lorensen, W.E., Cline, H.E.: Marching cubes: a high-resolution 3D surface construction algorithm. In: Proceedings of the 14th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH 87, pp. 163–169, New York, NY, USA, 1987. Association for Computing Machinery
23. Loshchilov, I., Hutter, F.: Decoupled weight decay regularization. *Journal, arXiv preprint arXiv:1711.05101* (2017)
24. Tatarchenko, M., Richter, S.R., Ranftl, R., Li, Z., Koltun, V., Brox, T.: What do single-view 3D reconstruction networks learn? *CoRR*, abs/1905.03678 (2019)
25. Foti, S., et al.: Clarkson: intraoperative liver surface completion with graph convolutional VAE. In: *Uncertainty for Safe Utilization of Machine Learning in Medical Imaging, and Graphs in Biomedical Image Analysis*, pp. 198–207 (2020)
26. Sarkar, K., Varanasi, K., Stricker, D.: Learning quadrangulated patches for 3D shape parameterization and completion. In: *International Conference on 3D Vision (3DV)*, pp. 383–392 (2017)
27. Kazhdan, M., Hoppe, H.: Johns Hopkins University: screened poisson surface reconstruction. *ACM Trans. Graph.* **32**(3), 1–13 (2013)

28. Chen, X., et al.: Shape registration with learned deformations for 3D shape reconstruction from sparse and incomplete point clouds. *J. Med. Image Anal.* **74**, 102228 (2021). Elsevier
29. Wang, Y., et al.: Dynamic Graph CNN for Learning on Point Clouds. *ACM Trans. Graph. (TOG)* **38**, 1–12 (2019)