



Titre: Génération de maillage surfacique de dents par apprentissage profond
Title:

Auteur: Olivier Lessard
Author:

Date: 2022

Type: Mémoire ou thèse / Dissertation or Thesis

Référence: Lessard, O. (2022). Génération de maillage surfacique de dents par apprentissage profond [Mémoire de maîtrise, Polytechnique Montréal]. PolyPublie.
Citation: <https://publications.polymtl.ca/10424/>

 **Document en libre accès dans PolyPublie**
Open Access document in PolyPublie

URL de PolyPublie: <https://publications.polymtl.ca/10424/>
PolyPublie URL:

Directeurs de recherche: Farida Cheriet, & François Guibault
Advisors:

Programme: Génie informatique
Program:

POLYTECHNIQUE MONTRÉAL

affiliée à l'Université de Montréal

Génération de maillage surfacique de dents par apprentissage profond

OLIVIER LESSARD

Département de génie informatique et génie logiciel

Mémoire présenté en vue de l'obtention du diplôme de *Maîtrise ès sciences appliquées*

Génie informatique

Juillet 2022

© Olivier Lessard, 2022.

POLYTECHNIQUE MONTRÉAL

affiliée à l'Université de Montréal

Ce mémoire intitulé:

Génération de maillage surfacique de dents par apprentissage profond

présenté par **Olivier LESSARD**

en vue de l'obtention du diplôme de *Maîtrise ès sciences appliquées*

a été dûment accepté par le jury d'examen constitué de :

Michel DAGENAIS, président

Farida CHERIET, membre et directrice de recherche

François GUIBAULT, membre et codirecteur de recherche

Benjamin DE LEENER, membre

REMERCIEMENTS

J'aimerais remercier ma directrice de recherche Farida Cheriet ainsi que mon co-directeur François Guibault pour leur support moral et financier, pour l'organisation des rencontres d'équipes, pour avoir pris le temps de remplir des demandes de bourses et m'avoir aidé à réaliser mes premières publications scientifiques.

J'aimerais aussi remercier Julia Keren pour avoir aidé à l'avancement de ces travaux de ces recherches en partageant son expertise en dentisterie. De plus, j'aimerais remercier le laboratoire Kerenor Dental Studio pour le partage de leur expertise et de données, sans lesquels ce mémoire n'aurait pu être possible.

J'aimerais remercier Philippe Debanne pour sa direction des rencontres de laboratoire et son aide lors de réalisation de publications. Finalement, j'aimerais remercier Intellident Dentaire Inc., iMD Research Inc., the Natural Science, le conseil de recherches en sciences naturelles et génie du Canada (CRSNG) et MEDTEQ pour leur financement durant ces deux années de maîtrise. J'aimerais également remercier Calcul Canada pour leur ressources matérielles.

J'aimerais également remercier ma famille et mon entourage pour m'avoir supporté tout au long de cet apprentissage.

RÉSUMÉ

Lorsqu'une dent est abimée ou cariée, une couronne dentaire est habituellement conçue afin de restaurer la dent et de la protéger contre d'autres dommages. La génération de couronne dentaire implique la conception d'une forme 3D complexe. Cette forme comprend une surface externe visible une fois la couronne installée. En pratique, les techniciens dentaires utilisent des numériseurs intraoraux et des logiciels de conception assistée par ordinateur (CAD) pour réaliser cette tâche. Leur conception demande beaucoup de temps et d'expérience pour être bien réalisée, ce qui entraîne une grande variabilité dans la qualité des couronnes dentaires.

Jusqu'à maintenant les conceptions de couronnes sont toujours faites manuellement. Une des hypothèses qui motive l'équipe de recherche est que la génération de couronne dentaire peut être automatisée grâce à l'apprentissage profond. Nous croyons que les modèles 3D utilisés par les techniciens peuvent servir à développer des algorithmes d'intelligence artificielle permettant d'aider les techniciens dentaires dans leur conception.

L'objectif principal de ce mémoire est de compléter des maillages surfaciques de dents par apprentissage profond. Des modèles de dents partiels sont complétés grâce à un système. Le système a pour but de générer la partie manquante sous forme de maillage. L'étude de la complétion de dents guidera de futurs travaux sur la génération de couronnes dentaires.

Dans ce mémoire, un premier article présente des complétions de parties de mâchoires. L'algorithme utilisé permet de générer plusieurs types de dents complètes sous forme de nuages de points. Des maillages surfaciques peuvent être obtenus en effectuant une reconstruction de surface à partir de ces nuages de points. Toutefois, cette reconstruction est difficile en raison du bruit présent dans les dents générées.

Pour continuer, un second article présente des complétions de maillages de dent. Cet article comporte une nouvelle architecture de réseaux de neurones inspiré par les méthodes de déformations de maillages. Les complétions dans cet article sont effectuées pour un seul type de dent contrairement à l'article précédent. La solution présentée permet de déformer un gabarit afin d'obtenir directement des maillages de dents lisses avec détails.

Des expériences complémentaires ont également été réalisées afin de mieux comprendre les algorithmes utilisés dans ces deux articles. Dans certaines de ces expériences, une fonction de coût est utilisée afin de générer des dents broyant mieux les aliments.

Les travaux réalisés lors de cette maîtrise permettent de démontrer les limitations des complétions de nuages de points. Comme présenté dans les articles, ces méthodes ont de la difficulté à générer des formes lisses puisque la connectivité entre les points est négligée. De leur côté, les approches de déformations de maillages permettent de contrôler le lissage des formes générées en tirant avantage de la connaissance de la connectivité d'un gabarit. Elles permettent de générer de formes avec moins de flexibilité et plus de précision. La solution proposée dans de second article est donc préférable pour la complétion de dent puisqu'en pratique les formes à compléter sont connues et la complétion doit être faite avec une grande précision. Cette solution s'avère prometteuse pour la génération de couronne dentaire.

ABSTRACT

When a tooth is damaged or decayed, a dental crown is usually designed to restore the tooth and protect it from further damage. Dental crown generation involves the design of a complex 3D shape. This form includes an outer surface that is visible once the crown is installed. In practice, dental technicians use intraoral scanners and computer-aided design (CAD) software to perform this task. Their design requires a lot of time and experience to be done well, which leads to great variability in the quality of dental crowns.

Until now crown designs are still done manually. One of the hypotheses that motivates the research team is that the generation of dental crowns can be automated thanks to deep learning. We believe that the 3D models used by technicians can be used to develop artificial intelligence algorithms to help dental technicians in their design.

The main objective of this thesis is to complete surface meshes of teeth by deep learning. Partial tooth models are completed using a system. The system aims to generate the missing part in the form of a mesh. The study of tooth completion will guide future work on the generation of dental crowns.

In this master, a first article presents completions of jaw parts. The algorithm used makes it possible to generate several types of complete teeth in the form of point clouds. Surface meshes can be obtained by performing a surface reconstruction from these point clouds. However, this reconstruction is difficult due to the noise present in the generated teeth.

To continue, a second article presents tooth mesh completions. This paper features a new neural network architecture inspired by mesh deformation methods. The completions in this article are performed for a single type of tooth unlike the previous article. The presented solution makes it possible to deform a template to directly obtain meshes of smooth teeth with details.

Complementary experiments were also carried out to better understand the algorithms used in these two articles. In some of these experiments, a cost function is used in order to generate teeth that grind food better.

The work carried out during this master's degree makes it possible to demonstrate the limitations of point cloud completions. As presented in the articles, these methods have difficulty to generate smooth shapes since the connectivity between the points is neglected. For their part,

mesh deformation approaches make it possible to control the smoothing of the generated shapes by taking advantage of the knowledge of the connectivity of a template. They allow shapes to be generated with less flexibility and more precision. The solution proposed in the second article is therefore preferable for tooth completion since in practice the forms to be completed are known and the completion must be done with great precision. This solution is promising for the generation of dental crowns.

TABLE DES MATIÈRES

| | |
|---|------|
| REMERCIEMENTS | III |
| RÉSUMÉ..... | IV |
| ABSTRACT..... | VI |
| TABLE DES MATIÈRES | VIII |
| LISTE DES TABLEAUX..... | XI |
| LISTE DES FIGURES..... | XII |
| LISTE DES SIGLES ET ABRÉVIATIONS | XV |
| CHAPITRE 1 INTRODUCTION..... | 1 |
| 1.1 Anatomie d'une dent | 1 |
| 1.2 Couronne dentaire | 3 |
| 1.2.1 Conception en dentisterie..... | 4 |
| 1.3 Objectifs généraux..... | 5 |
| 1.4 Organisation du mémoire | 6 |
| CHAPITRE 2 REVUE CRITIQUE DE LA LITTÉRATURE | 7 |
| 2.1 Architectures de réseaux de neurones | 7 |
| 2.1.1 Auto-encodeur..... | 7 |
| 2.1.2 Réseaux antagonistes génératifs..... | 8 |
| 2.2 Apprentissage profond pour la génération de données 3D | 9 |
| 2.3 Apprentissage profond pour la complétion de nuages de points..... | 10 |
| 2.3.1 Réseaux de points..... | 10 |
| 2.3.2 Auto-encodeur de points | 11 |
| 2.3.3 Génération de nuages de points..... | 11 |
| 2.3.4 Complétion de nuage de points | 12 |

| | | |
|--|--|----|
| 2.4 | Reconstruction de surface à partir de nuages de points | 17 |
| 2.4.1 | Reconstruction assumant une surface lisse globale..... | 17 |
| 2.4.2 | Reconstruction avec gabarit | 18 |
| 2.5 | Conception de couronne dentaire avec apprentissage profond | 18 |
| CHAPITRE 3 DÉMARCHE DE L'ENSEMBLE DU TRAVAIL DE RECHERCHE ET ORGANISATION GÉNÉRALE DU DOCUMENT INDIQUANT LA COHÉRENCE DES ARTICLES PAR RAPPORT AUX OBJECTIFS DE RECHERCHE..... | | 20 |
| CHAPITRE 4 ARTICLE 1: DENTAL RESTORATIONS USING A MULTI-RESOLUTION DEEP LEARNING APPROACH..... | | 23 |
| 4.1 | Abstract | 23 |
| 4.2 | Introduction | 23 |
| 4.3 | Our approach | 25 |
| 4.4 | Experimental setting and results | 27 |
| 4.5 | Discussion | 31 |
| 4.6 | Conclusion..... | 31 |
| 4.7 | Acknowledgments | 32 |
| CHAPITRE 5 ARTICLE 2: MC-NET: MESH COMPLETION FOR DENTAL SCANS..... | | 33 |
| 5.1 | Abstract | 33 |
| 5.2 | Introduction | 33 |
| 5.3 | Related work | 38 |
| 5.3.1 | Point cloud completion..... | 38 |
| 5.3.2 | Mesh completion | 38 |
| 5.3.3 | Mesh deformation | 39 |
| 5.4 | Method | 40 |
| 5.4.1 | Initial surface..... | 41 |

| | | |
|--|---|----|
| 5.4.2 | Mesh deformation block..... | 42 |
| 5.4.3 | Graph unpooling layer..... | 42 |
| 5.4.4 | Discriminator's input..... | 43 |
| 5.4.5 | Losses..... | 44 |
| 5.5 | Experiments..... | 46 |
| 5.5.1 | Dataset and preprocessing..... | 46 |
| 5.5.2 | Implementation details..... | 47 |
| 5.5.3 | Comparison with state of the art..... | 48 |
| 5.5.4 | Ablation study..... | 50 |
| 5.6 | Discussion..... | 52 |
| 5.7 | Conclusion..... | 54 |
| 5.8 | Acknowledgments..... | 54 |
| CHAPITRE 6 EXPÉRIMENTATION COMPLÉMENTAIRE..... | | 56 |
| 6.1 | Analyse complémentaire des résultats l'article 1..... | 56 |
| 6.2 | Résultats complémentaires de l'ajout d'un gabarit de points dans le décodeur de PF-Net..... | 60 |
| 6.3 | Ajout d'une fonction de coût de fonctionnalité..... | 61 |
| 6.4 | Résultats complémentaires de l'ajout d'information conditionnelle au discriminateur..... | 65 |
| CHAPITRE 7 DISCUSSION GÉNÉRALE..... | | 68 |
| 7.1 | Synthèse des travaux..... | 68 |
| 7.2 | Limitations de la solution proposée..... | 69 |
| CHAPITRE 8 CONCLUSION..... | | 72 |
| 8.1 | Nouvelles avenues de recherche..... | 72 |
| RÉFÉRENCES..... | | 74 |

LISTE DES TABLEAUX

| | |
|---|----|
| Tableau 1.1 Numérotation dentaire | 2 |
| Table 4.1 Test Chamfer Distance. | 28 |
| Table 5.1 Tooth completion results after 6-fold cross-validation. The reported metrics are [Pred → GT / GT → Pred]. Resulting normalized CD values are scaled by 100. | 49 |
| Table 5.2 Ablation study to evaluate the effect of different components of our MC-Net model. Each row reports the average error metrics when removing the corresponding component. Resulting normalized CD values are scaled by 100. | 52 |
| Tableau 6.1 Résultats avec et sans recalage pour une génération de 768 points | 59 |
| Tableau 6.2 Comparaison des méthodes pour une génération de prémolaires de 768 points | 61 |
| Tableau 6.3 Effet de la fonction d'occlusion sur la base de données de l'article 1 | 63 |
| Tableau 6.4 Effet de la fonction d'occlusion sur la base de données de prémolaires | 64 |
| Tableau 6.5 Effet de la nouvelle formulation de l'entrée du discriminateur | 65 |

LISTE DES FIGURES

| | |
|---|----|
| Figure 1.1 Numérotation dentaire selon la FDI, par Maduixa, 2007, Wikimedia Commons (https://commons.wikimedia.org/wiki/File:Raspored_ljudskih_zuba.jpg). CC BY-SA..... | 3 |
| Figure 1.2 Une couronne dentaire comprend une surface externe et une surface interne..... | 4 |
| Figure 2.1 Structure générale d'un auto-encodeur..... | 7 |
| Figure 2.2 Structure générale des réseaux antagonistes génératifs | 8 |
| Figure 2.3 Architecture de PF-Net. | 16 |
| Figure 4.1 Illustration of the problem formulation. A 3D model of the original teeth (a), the conditional information referred to as the context (b), and the extracted tooth (c) used as the ground truth. Their point clouds are shown at the bottom. | 26 |
| Figure 4.2 Histogram of test CDs. The horizontal axis corresponds to intervals of CD values scaled by 100..... | 29 |
| Figure 4.3 (a) a canine, (b) a premolar, (c) a molar seen from above. Green represents points with errors close to zero when red means errors of 1 millimeter or more. The left and right shapes are examples of GT2Pred and Pred2GT errors. | 30 |
| Figure 4.4 Predicted tooth and corresponding ground truth shown in green with their surroundings in red..... | 30 |
| Figure 5.1 Illustration of the tooth completion problem. Using a point cloud of the surrounding teeth, our model predicts the missing surface that the tooth should have, as shown by the mesh in cyan. | 35 |
| Figure 5.2 Architecture of MC-Net. The multi-resolution encoder (MRE) analyzes the three input resolutions to create global features. The mesh deformation decoder (MDD) uses the features to deform the template with three deformation blocks. (<i>CMLP means Combined Multi-Layer Perceptron; IFPS is Iterative Farthest Point Sampling; CD is Chamfer Distance.</i>)..... | 36 |
| Figure 5.3 The template mesh (in gray) is positioned and scaled to fit inside a bounding box defined by the segmented input..... | 41 |

| | |
|---|----|
| Figure 5.4 Details of the Mesh Deformation Decoder. MDD creates an output mesh along with 2 intermediate results. The graph unpooling layers preserve the connectivity during the coarse-to-fine generation. The plus circles represent add operations. Layers operating on points are shown in white; those operating on graphs are shown in green. | 43 |
| Figure 5.5 An example of the discriminator's input. The points in gray in the lower portion are sampled from the input using teeth segmentation and iterative farthest point sampling (IFPS). | 44 |
| Figure 5.6 Illustration of an input-output pair. Half of the tooth to be completed is removed using a planar slice. | 47 |
| Figure 5.7 Examples of mesh completions by the different architectures. a) partial input meshes used to create points clouds; b) generated teeth with PF-Net; c) generated teeth with PF-Net*; d) generated teeth with MC-Net; e) ground truth meshes. In all cases except d) , ball pivoting reconstruction was used to create meshes from point clouds. Holes in the ground truth and the prediction meshes are due to reconstruction artefacts. | 50 |
| Figure 5.8 Qualitative results of the ablation study. The images in each column reflect the contribution of each component of MC-Net by showing the resulting mesh without that component. The red circles overlaid on the images at bottom right show subtle difference . | 55 |
| Figure 6.1 Relation entre la quantité de données d'entraînement et les erreurs de l'ensemble de test par dent de l'article 1 | 56 |
| Figure 6.2 Relation entre la quantité de données dans l'ensemble d'entraînement et les erreurs de l'ensemble de test par catégorie de dents de l'article 1 | 57 |
| Figure 6.3 Erreurs des prédictions (en couleur) par rapport aux références (en noire). | 58 |
| Figure 6.4 Positionnement d'un gabarit pour la dent 14. | 60 |
| Figure 6.5 PPD avec gabarit. | 60 |
| Figure 6.6 Mesure de distance d'occlusion avec nuage de points | 62 |
| Figure 6.7 Histogramme illustrant le changement de distribution de CD de test | 65 |
| Figure 6.8 Analyse qualitative de l'amélioration avec une vue du dessous des dents. | 66 |

Figure 6.9 Autres résultats pour PF-Net-D+ 67

LISTE DES SIGLES ET ABRÉVIATIONS

| | |
|--------|-----------------------------------|
| AE | Auto-encodeur |
| CD | Chamfer Distance |
| CMLP | Combined Multi-Layer Perceptron |
| CNN | Convolutional Neural Network |
| DL | Deep Learning |
| FDI | World Dental Federation |
| GAN | Generative Adversarial Network |
| GCN | Graph Convolutional Network |
| GPA | Generalized Procrustes Analysis |
| IFPS | Iterative Farthest Point Sampling |
| MC-Net | Mesh completion Network |
| MDD | Mesh Deformation Decoder |
| MRE | Multi-Resolution Encoder |
| PCN | Point Completion Network |
| PPD | Point Pyramid Decoder |
| PRS | Point Rank Sampling |
| RL | Reinforcement Learning |

CHAPITRE 1 INTRODUCTION

Les bureaux de dentistes doivent faire des centaines de reconstructions dentaires par année (Elani et al., 2018). Pour produire chacune de ces prothèses dentaires, un professionnel dentaire doit spécifier la forme à concevoir en fonction de la denture du patient. Cette conception demande beaucoup de temps et d'expérience de la part du professionnel pour être réalisée adéquatement. Ces besoins entraînent une grande variabilité dans la qualité des prothèses dentaires. Présentement, aucun système intelligent n'aide les techniciens à faire cette conception. Ce manque d'outils est ce qui a motivé la création de ce projet de recherche. Le projet consiste à utiliser l'apprentissage profond pour générer des prothèses personnalisées. Pour ce faire, des données de scans dentaires fournies par un partenaire industriel serviront à entraîner des algorithmes pour générer des couronnes dentaires. Les prothèses générées par cet algorithme peuvent ensuite être ajustées par le technicien avec des logiciels CAD (Liao et al., 2015) au besoin. En produisant automatiquement des prothèses personnalisées, la rapidité de conception et la qualité des prothèses pourront être améliorées. Ce projet a donc potentiellement de grands impacts sur le domaine dentaire.

Dans les sections qui suivent, nous présenterons les notions essentielles de l'anatomie d'une dent permettant de comprendre la position des dents et des surfaces dentaires. Les couronnes dentaires seront également décrites ainsi que leurs différentes conceptions afin de comprendre l'objectif long terme de ce projet. Par la suite, les objectifs généraux pour cette maîtrise ainsi que la structure de ce mémoire seront présentés.

1.1 Anatomie d'une dent

Pour bien comprendre les couronnes dentaires pour les dents humaines, il faut tout d'abord comprendre l'anatomie d'une dent. La dent humaine est un organe dur dont la fonction est de capturer, mâcher et broyer les aliments. Étant donné que les humains sont omnivores, la forme de leurs dents et le nombre de celles-ci sont adaptées pour cette alimentation. Le terme dentition désigne la formation, la croissance et l'éruption des dents pendant la vie de l'individu. En revanche, le terme denture est utilisé pour désigner l'ensemble des dents.

Parmi les dents humaines, on distingue les incisives, les canines, les prémolaires et les molaires. Chez l'adulte, une arcade dentaire complète comporte 16 dents soit : 4 incisives, 2 canines, 4

prémolaires et 6 molaires. La numérotation de la « World Dental Federation » (FDI) et la position de ces catégories de dents sont représentées dans la Figure 1.1 et le Tableau 1.1.

Tableau 1.1 Numérotation dentaire

| Position relative | Nom de la dent |
|--------------------------|--------------------------------------|
| 1 | Incisive centrale |
| 2 | Incisive latérale |
| 3 | Canine |
| 4 | Première prémolaire |
| 5 | Deuxième prémolaire |
| 6 | Première molaire |
| 7 | Deuxième molaire |
| 8 | Troisième molaire ou dent de sagesse |

Dans la Figure 1.1, on peut voir la numérotation des quadrants lorsqu'on regarde une mâchoire de face. Par exemple, le premier quadrant se situe en haut à gauche. La FDI utilise une numération à deux chiffres où le chiffre des dizaines représente le quadrant et le chiffre des unités représente la position relative de la dent dans le quadrant. Ainsi, la position 18 représente la dent de sagesse en haut à gauche.

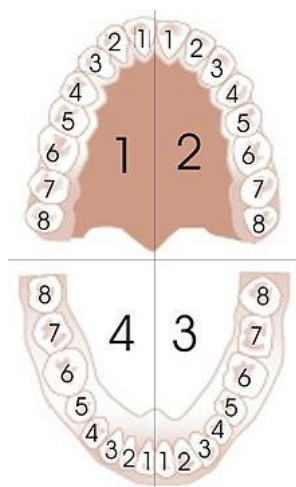


Figure 1.1 Numérotation dentaire selon la FDI, par Maduixa, 2007, Wikimedia Commons (https://commons.wikimedia.org/wiki/File:Raspored_ljudskih_zuba.jpg). CC BY-SA.

De plus, il est important de définir les faces d'une dent. La face occlusale est celle constituée de bosses sur laquelle on mord. La face vestibulaire est celle à l'extérieure contre la joue. La face cachée vers l'avant entre deux dents se nomme la mésiale et celle vers l'arrière se nomme la face distale. Finalement, pour les dents du bas, la face contre la langue est appelée face linguale tandis que pour celles du haut, c'est la face palatine.

1.2 Couronne dentaire

Les couronnes dentaires sont des dents artificielles qui ont pour but de restaurer les fonctions, la forme et l'apparence des dents abimées. Elles peuvent être fabriquées avec différents matériaux comme du métal ou de la porcelaine. Une dent abimée est typiquement une dent usée, avec des fissures ou des faiblesses. Si une telle dent ne se fait pas protéger, « elle peut fracturer et devenir non restaurable » (*La couronne dentaire : quand? pourquoi? comment? | Clinique dentaire Laetitia Trudel Baribault*, s. d.). Les couronnes dentaires servent à réparer ces défauts et elles peuvent être nécessaires à la suite de traitements de canal puisque la dent en question sera plus fragile à la suite de l'opération.

Une couronne dentaire comprend 2 parties soit la surface externe et la surface interne (voir la Figure 1.2). La surface externe est la surface visible lorsqu'on regarde dans une bouche. Cette surface doit avoir une apparence de dent naturelle. Elle doit posséder des points de contacts avec les dents voisines sur la même arche pour éviter que de la nourriture reste prise entre les dents. Elle doit également posséder des points de contacts avec les dents de l'arche opposée pour bien broyer les

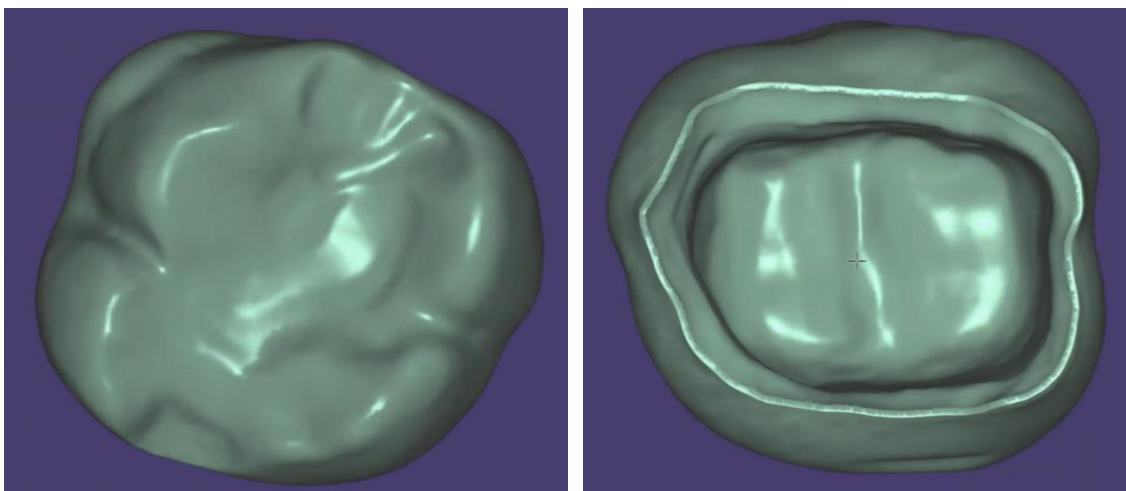


Figure 1.2 Une couronne dentaire comprend une surface externe et une surface interne.

aliments. Pour sa part, la surface interne inclut une bordure épaisse qui épouse la ligne de marge de la préparation. La surface interne doit laisser un mince espace entre la couronne et la préparation. Lors de l'installation de la couronne, cet espace est requis pour coller la couronne sur la préparation à l'aide d'une couche de ciment.

1.2.1 Conception en dentisterie

Pour construire une couronne dentaire, le spécialiste commence par obtenir une représentation de la denture avec un moule ou un modèle 3D. La représentation est ensuite envoyée à un laboratoire où la couronne est conçue. Cette conception peut être faite de manière manuelle ou digitale. Les conceptions manuelles utilisent un moule de la préparation dentaire et ses dents voisines. Un second moule est ensuite utilisé par-dessus le premier pour faire couler un liquide qui est ensuite chauffé à haute température. Avec le temps, ce liquide durcit en refroidissant. Une fois refroidi, le moule est ensuite brisé afin d'obtenir la couronne créée. Par opposition, certains laboratoires utilisent des logiciels de conception assistée par ordinateur (CAD). Ceux-ci permettent de

visualiser les modèles 3D obtenus préalablement et de concevoir la forme de la couronne avec précision.

Lors de la conception d'une couronne dentaire, le technicien commence par concevoir la surface interne de la couronne en fonction de la préparation. Il débute par détecter la ligne de marge de la préparation. Cette ligne détermine les limites de la couche de ciment. En effet, pour s'assurer que la couronne tienne sur la préparation, un espace doit être réservé pour introduire une couche de ciment entre la préparation et la couronne. Le technicien spécifie l'épaisseur de cette couche, il choisit habituellement une épaisseur de 0.08 millimètres. Il détermine également la bordure qui sera utilisée pour former le bas de la couronne. La bordure et la surface extérieure de la couche de ciment sont utilisées pour former la surface interne de la couronne.

Pour concevoir la surface extérieure, le technicien dentaire commence par sélectionner une forme de dent générale dans une base de données. Cette forme est utilisée comme point de départ dans la conception. Par la suite, le technicien ajuste la grosseur de la dent, sa position et son orientation selon les autres dents du patient pour obtenir une forme réaliste touchant potentiellement les dents voisines. Il coupe la forme au besoin pour qu'il n'y ait aucun espace. Avec ces manipulations, le technicien s'assure de ne pas laisser d'espace entre la couronne et les dents voisines pour éviter que de la nourriture puisse s'y infiltrer. Ces ajustements font parties des étapes les plus longues dans la conception. Pour continuer, le technicien ajuste les courbures de la surface d'occlusion pour lui donner de bons points de contact avec la dent opposée. Il s'assure de laisser environ 0.1 mm d'espace entre la couronne et la dent opposée pour permettre une fermeture normale de la bouche tout en permettant au patient de mastiquer avec la couronne.

1.3 Objectifs généraux

Ces aperçus des enjeux cliniques permettent de contextualiser les travaux de recherche liés à la conception de couronne dentaire et les défis associés. Le chapitre suivant présente une revue de littérature à propos de l'apprentissage profond géométrique se concentrant sur la complétion de nuages de points.

Dans le cadre de cette maîtrise, les travaux de recherches se sont concentrés sur la complétion de dent directement en 3D. L'objectif général est d'obtenir un système pouvant générer des surfaces de dents sous forme de maillage à partir de scans dentaires partiels.

1.4 Organisation du mémoire

Ce mémoire est structuré en 6 chapitres. Dans le premier chapitre, une revue de littérature sera présentée. Elle portera principalement sur la complétion de nuage de points.

Le second chapitre portera sur la démarche de l'ensemble du travail. Il permettra de préciser les objectifs et la pertinence des articles inclus dans ce travail.

Deux articles de recherches constituent le corps de ce mémoire. Ils sont présentés aux chapitres trois et quatre. Le premier a été publié en Avril 2022 pour la conférence *IEEE International Symposium on Biomedical Imaging*. Cet article porte sur la complétion de dent complète sous formes de nuages de points. Le second a été soumis en Juin 2022 à la revue *IEEE Transactions on Medical Imaging*. Celui-ci porte sur la complétion de dent partielle sous forme de maillage directement. La méthodologie utilisée repose sur une nouvelle architecture de réseaux de neurones tirant avantage d'une déformation de gabarit.

Dans le cinquième chapitre, des expériences complémentaires permettent de mieux comprendre les architectures utilisées. D'autres expériences utilisent une fonction de coût conçue pour la conception de couronne dentaire.

Le chapitre 6 comporte une discussion sur l'ensemble du travail, des explications des limitations et des recommandations pour des travaux futures.

Le chapitre 7 conclut ce mémoire en résumant l'ensemble du travail.

CHAPITRE 2 REVUE CRITIQUE DE LA LITTÉRATURE

Ce chapitre présente les travaux importants liés à la complétion de nuages de points. Avec le gain en popularité des équipements de scan 3D tels que scanneurs intraoraux, LiDAR, laser et RGB-D, les nuages de points sont de plus en plus utilisés dans différents domaines. Les sections suivantes permettront de justifier leur utilisation, de présenter l'état de l'art en complétion de nuage de points et leur pertinence pour la complétion de dents. De plus, des travaux liés à la génération de couronnes dentaires ou la complétion de dents seront présentées.

2.1 Architectures de réseaux de neurones

Dans cette section, deux architectures de réseaux de neurones importantes seront présentées. Ces architectures sont celles d'auto-encodeur et de réseaux antagoniste génératifs. Ces deux structures ont servi de base pour plusieurs travaux dans cette revue de littérature.

2.1.1 Auto-encodeur

Un auto-encodeur (AE) est un réseau de neurones artificiels entraîné pour effectuer une copie d'une entrée vers une sortie. À l'interne, le réseau possède une couche cachée avec une représentation h qui représente un encodage de l'entrée. Ce type de réseau comprend deux parties soit : une fonction d'encodage $h = f(x)$ et une fonction de décodage $r = g(h)$. Cette architecture est présentée à la Figure 2.1.

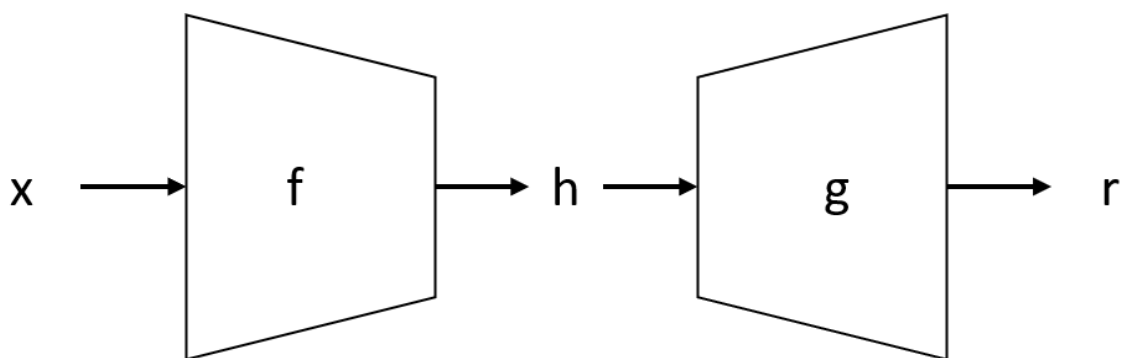


Figure 2.1 Structure générale d'un auto-encodeur

Un auto-encodeur parfait étant capable de reconstruire complètement une entrée ne serait pas très intéressant. Ces réseaux sont généralement conçus pour ne pas être capable de reconstruire

complètement les entrées. Les couches du réseau imposent des restrictions qui permettent uniquement de reconstruire approximativement l'entrée. Par exemple, une couche h ayant une dimension plus petite que celle de l'entrée force le réseau à compresser l'information et à conserver uniquement les caractéristiques importantes. En général, les réseaux avec un encodage à dimension inférieure à celle l'entrée sont entraîné afin de minimiser une fonction de coût

$$L(x, g(f(x))), \quad (2.1)$$

où L est une fonction de coût pénalisant $g(f(x))$ d'être différent de x . Par exemple, cette fonction peut mesurer l'erreur quadratique moyenne.

2.1.2 Réseaux antagonistes génératifs

Les réseaux antagonistes génératifs (GAN) (Goodfellow et al., 2014) ont eu un grand impact sur la recherche en génération de données. Ce type de réseaux a été utilisé à maintes reprises pour générer des images, des sons et des maillages (Cheng et al., 2019; Donahue et al., 2018; Isola et al., 2017). L'apprentissage avec GAN permet d'obtenir un réseau de neurones générant des données artificielles à partir d'un vecteur de bruit z . Un GAN comprend deux réseaux de neurones qui sont entraînés ensemble dans un scénario de théorie des jeux. Le premier réseau est un générateur G qui génère des échantillons, tandis que le second est un discriminateur D qui essaie de distinguer

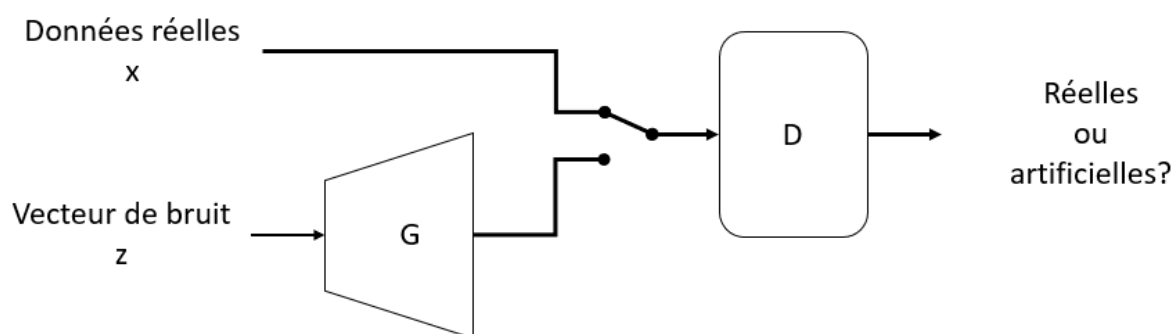


Figure 2.2 Structure générale des réseaux antagonistes génératifs

des échantillons réels des échantillons du générateurs. Le générateur a pour but de tromper le discriminateur. Les deux réseaux sont présentés à la Figure 2.2.

L'entraînement des deux réseaux peut être formulé comme un jeux minimax avec une fonction de valeur $V(G, D)$:

$$\min_G \max_G V(D, G) = \mathbb{E}_{x \sim p_{\text{données}}} \log D(x) + \mathbb{E}_{x \sim p_{\text{modèle}}} \log(1 - D(x)). \quad (2.2)$$

En pratique, les deux réseaux sont souvent mis-à-jour en alternance. Le générateur minimise la fonction $\log(1 - D(G(z)))$ afin de tromper le discriminateur et le discriminateur maximise la fonction $\log D(x) + \log(1 - D(G(z)))$ afin bien classifier les échantillons.

2.2 Apprentissage profond pour la génération de données 3D

Les réseaux de neurones à convolutions (CNN) ont été utilisés à maintes reprises pour générer des images et du son (Goodfellow et al., 2014; Van Den Oord et al., 2016). Une des raisons de leur succès est l'analyse à plusieurs niveaux de ces architectures permettant de reconnaître des motifs dans les données. Cependant, la généralisation de ces approches aux données 3D est loin d'être triviale. Contrairement aux images et aux sons, les données 3D n'ont pas de structure euclidienne. La nature non-euclidienne de ces données implique qu'il n'y a pas de système de coordonnées global, d'invariance à la translation et de décalage de position. Les opérations de bases des CNN comme la convolution ne sont pas bien définies pour ce type de données.

Les données 3D sont généralement représentées sous forme de graphes, de voxels, de nuages de points ou de maillages. Les surfaces sont souvent représentées par des maillages, des nuages de points ou des voxels. L'utilisation de voxels est intéressante puisqu'elle permet d'adapter facilement des méthodes 2D pour être utilisées en 3D. Cependant, en raison de leur grand besoin de mémoire, les méthodes basées sur les voxels sont moins populaires. Pour effectuer de la complétion, Wang et al. (Wang et al., 2020) ont proposé 3D-DCGAN. Leur approche utilise un premier GAN pour faire une prédiction locale. Par la suite, un second GAN est utilisé pour faire une prédiction globale avec une bonne jonction entre la forme existante et celle générée. Leur approche nécessite toutefois une reconstruction pour obtenir un maillage.

Les maillages polygonaux fournissent un mode de représentation des données intéressant pour analyser des surfaces. Ils permettent d'obtenir une représentation très détaillée des objets grâce aux points et arrêtes qu'ils contiennent. Plusieurs architectures modernes ont été développées pour générer des maillages triangulaires. Dans (Ranjan et al., 2018), les auteurs introduisent un AE de maillages à convolution utilisant une nouvelle opération d'échantillonnage qui permet de réduire

le nombre de paramètres requis. MeshGAN (Cheng et al., 2019) permet de générer des maillages directement en utilisant un réseau antagoniste génératif (GAN).

D'autre part, les nuages de points permettent également de représenter des surfaces 3D. En raison du peu de ressources nécessaires avec ce type de données, les approches avec nuages de points ont été plus étudiées dans la communauté scientifique. Les défis principaux liés aux nuages de points sont l'irrégularité, la non-ordonnance et le manque de structure des points. Premièrement, l'échantillonnage utilisé pour obtenir les points peut avoir été imparfait. Certaines régions peuvent être sur-échantillonnées tandis que d'autres sont sous-échantillonnées. La densité variable dans les nuages de points peut être atténuée à l'aide de techniques de sous-échantillonnage mais elle ne peut être complètement éliminée. Deuxièmement, un nuage de points est un ensemble de points non-ordonné. Si deux points dans l'ensemble sont permutés, l'ensemble reste inchangé. Troisièmement, un nuage de points est non-structuré. Chaque point est considéré comme étant indépendant et la distance avec ses voisins n'est souvent pas fixée contrairement aux images où la distance entre les pixels est toujours fixée. Les prochaines sections comporteront plusieurs méthodes de génération et complétion de points.

2.3 Apprentissage profond pour la complétion de nuages de points

L'utilisation de l'apprentissage profond avec des nuages de points a été beaucoup exploré suite aux travaux de PointNet (Qi et al., 2017). Cette architecture a été la première à fournir une solution traitant les nuages de points directement malgré les défis qu'ils impliquent. Plusieurs avancées se sont basées sur PointNet pour ensuite mieux modéliser les régions locales ou pour accomplir d'autres tâches telles que la génération et la complétion de nuage de points.

2.3.1 Réseaux de points

PointNet utilise principalement 2 fonctions de base soit les perceptrons multicouches (MLP) et la fonction de mise en commun maximale (« max pooling »). Les perceptrons multicouches permettent d'apprendre les paramètres requis pour obtenir des fonctionnalités pertinentes pour effectuer une tâche. Dans PointNet, les MLP sont partagés entre tous les points à chaque couche. Pour obtenir des fonctionnalités globales, la fonction de mise en commun maximale est utilisée en raison de sa symétrie. Lorsque l'ordre de l'entrée d'une fonction symétrique est modifié, sa sortie reste inchangée. En plus, de ces deux fonctions de base, PointNet utilise un mini-réseau afin

d'apprendre la matrice de transformation à appliquer aux coordonnées des points pour que le réseau soit invariant aux transformations géométriques. L'architecture de PointNet a permis d'améliorer beaucoup la classification et la segmentation de nuages de points. Cependant, ses prédictions étaient peu détaillées parce qu'elle ne considérait pas la dépendance locale entre les points.

Qui et al. (Qi et al., 2017) ont proposé PointNet++, une approche qui consiste à appliquer PointNet de manière récursive sur différents groupes de points pour effectuer une meilleure analyse locale. L'approche utilise « iterative farthest point sampling » (IFPS) et « ball query » pour échantillonner et regrouper les points et obtenir différents niveaux d'abstraction des données. Leurs modifications ont permis d'améliorer l'état de l'art dans ces années sur la classification et segmentation de nuages de points.

2.3.2 Auto-encodeur de points

Dans Yang et al. (Yang et al., 2018), les auteurs proposent FoldingNet, un AE pour diverses tâches avec nuages de points. L'encodeur proposé est une amélioration de celui de PointNet. Il comprend des couches de graphes mentionnées dans (Atwood et Towsley, 2016) qui permettent d'obtenir un vecteur de caractéristiques. Le décodeur utilise deux opérations de *folding* afin d'apprendre à déformer une grille 2D en une surface 3D. Dans ces opérations, les caractéristiques produites par le décodeur sont utilisées pour chaque point de la grille à déformer. Leurs résultats démontrent que leur décodeur est plus efficace que celui de PointNet pour diverses tâches.

2.3.3 Génération de nuages de points

Certains chercheurs se sont concentrés sur la génération de nuages de points complets pour apprendre différentes représentations d'un même objet. Achlioptas et al. ont débuté en proposant une approche originale L-GAN (Achlioptas et al., 2018) qui consiste à générer des nuages de points en apprenant la distribution des caractéristiques latentes avec un GAN. Leur méthode utilise un AE pour apprendre les représentations latentes et ensuite entraîner un GAN avec celles-ci. Leur GAN arrive à générer des formes réalistes couvrant la plupart de la distribution des données. Valsesia et al. (Valsesia et al., 2018) proposent d'intégrer des opérations de convolutions de graphes dans un GAN pour générer des nuages de points. Contrairement aux convolutions de graphes traditionnelles, la connectivité des graphes n'est pas connue dans leur formulation. Leur méthode utilise plutôt des matrices d'adjacences à certaines couches qui permettent de générer des

formes locales mais elle possède une grande complexité. Wook Shu et al. proposent plutôt d'utiliser les structures d'arbres pour représenter les connectivités dans les graphes (Wook Shu et al., 2019). Leur modèle intitulé TreeGAN a surpassé les méthodes précédentes. Cependant, ces méthodes basées sur les GAN n'effectuent pas des générations conditionnelles. La prochaine section comporte des méthodes utilisant une information conditionnelle afin d'effectuer une complétion.

2.3.4 Complétion de nuage de points

La complétion de nuages de points est une tâche importante puisque souvent les données 3D sont incomplètes. La tâche consiste à prédire la forme géométrique complète d'un objet à partir d'une observation partielle. Les approches pour cette tâche sont soit basées sur la géométrie, sur l'alignement ou sur l'apprentissage. Les approches basées sur la géométrie utilisent uniquement des informations géométriques de l'objet partiel pour faire la reconstruction. Les techniques de reconstruction de surface (Berger et al., 2017) permettent d'obtenir des surfaces lisses sans trous. Cependant, ces techniques assument que l'objet est presque complet et que la géométrie manquante peut être prédite uniquement à partir de celle qui est observée. Cette supposition n'est pas souvent applicable dans les problèmes réels. Les méthodes basées sur l'alignement utilisent une forme générale obtenue en utilisant une base de données de formes. Certaines apprennent à déformer la forme générale pour qu'elle soit cohérente avec la forme partielle (Blanz et Vetter, 1999; Gupta et al., 2015). Cependant, elles requièrent une optimisation pour aligner la forme générale lors de l'inférence. Les méthodes basées sur l'apprentissage permettent de prédire directement la forme complète à partir de la représentation partielle. Ceci permet de réduire le temps d'inférence et d'obtenir une bonne généralisation.

2.3.4.1 Complétion basée sur la géométrie

Les approches basées sur la géométrie utilisent soit des continuités de surface ou des propriétés de symétrie. Celles utilisant la continuité permettent de réparer les trous dans les surfaces (Berger et al., 2014; Davis et al., 2002). D'autres exploitent plutôt la symétrie dans les objets. Ces méthodes détectent les objets avec symétries pour ensuite trouver la forme manquante avec une transformation d'une forme existante dans l'objet (Mitra et al., 2006, 2013). Cependant, toutes ces

approches sont très limitées. Elles ne permettent pas de faire une complétion adéquate lorsqu'une grande région est manquante ou qu'il n'y a pas suffisamment de symétrie dans l'objet.

2.3.4.2 Complétion basée sur un gabarit

Certaines approches développées utilisent des connaissances préalables de la forme des objets. En effet, les différents objets d'une classe ont souvent des caractéristiques générales qui nous permettent de reconnaître la classe à laquelle ils appartiennent. Inspirés par cette constatation, Cocias et al. ont développé GFPNet, un modèle utilisant un gabarit pour chaque classe d'objet à compléter (Cocias et al., 2020). Les formes générales sont obtenues en utilisant une analyse statistique avec « Generalized Procrustes Analysis » (GPA). GPA permet d'obtenir une forme moyenne lorsque la correspondance entre les points est connue, ce qui est le cas dans leur formulation. Leur méthode commence par aligner une forme générale avec les représentations partielles. Pour continuer, un modèle est entraîné afin d'apprendre à déformer la forme générale pour modéliser la forme partielle et apprendre les points manquants. Ce type d'approche nécessite de connaître la correspondance entre les points des différents objets, ce qui requiert beaucoup de travail manuel. Il demande aussi une optimisation pour aligner la forme générale lors de l'inférence, ce qui peut être problématique pour une application en ligne.

2.3.4.3 Complétion basée sur l'apprentissage

Inspiré par FoldingNet, Point Completion Network (PCN) a été développé pour compléter des formes sans aucune supposition concernant la symétrie ou l'annotation (Yuan et al., 2018). PCN comprend un encodeur et un décodeur entraîné de manière supervisée. L'encodeur utilise deux couches de PointNet pour produire un vecteur global de caractéristiques. Avec ce vecteur, le décodeur utilise tout d'abord un MLP pour produire une forme générale et ensuite un module de « folding » comme celui de FoldingNet (Yang et al., 2018) pour apprendre une surface locale à chaque point de la forme générale. Leur approche éprouve toutefois des difficultés pour compléter les parties fines et déconnectées. Afin d'agrandir l'espace de solution pendant l'apprentissage, Tchapmi et al. propose TopNet, un décodeur plus général qui utilise une structure d'arbre (Lyne Tchapmi et al., 2019). Cependant, la grosseur de leur arbre limite l'espace de solution. Contrairement à PCN et TopNet qui utilisent les opérations de « max pooling » entraînant une perte d'information au niveau des détails, Xie et al. propose d'utiliser des grilles 3D comme représentations intermédiaires afin d'utiliser des opérations de convolutions (Xie et al., 2020). Leur

méthode GR-Net utilise un CNN 3D pour générer une forme grossière et un MLP pour ajouter des détails. Elle comprend un encodage des coins de chaque voxel afin de réduire la perte d'information due à la conversion mais cela dépend de la résolution utilisée. Wang et al. ont également exploité le potentiel des voxels pour faire de la complétion de forme mais cette fois en intégrant un générateur de bords dans leur architecture VE-PCN (Wang et al., 2021). Ces mêmes auteurs ont également proposé d'améliorer le décodeur de GR-Net en utilisant un raffinement itératif qui permet d'augmenter la résolution (Wang et al., 2020). Leur décodeur utilise les caractéristiques globales en plus des détails de l'entrée partielle afin de minimiser la perte d'information causée par les opérations de « max pooling ».

D'autres chercheurs ont plutôt développé des approches basées sur les graphes. Shi et al. ont utilisé des déformations de graphes pour effectuer la complétion de nuage de points (Shi et al., 2021). Leur approche utilise un encodeur-décodeur pour obtenir une première complétion grossière. Par la suite, ils construisent un graphe avec les points en entrée comme points de contrôle pour raffiner leur complétion à l'aide de déformations. Leur formulation a l'avantage de réduire le bruit grâce aux opérations sur les graphes mais elle dépend fortement de la qualité de leur encodeur-décodeur.

Les avancées en apprentissage auto-supervisé ont inspiré plusieurs travaux en complétion de nuage de points. Récemment, (Wang et al., 2021) propose d'améliorer la dernière approche en ajoutant un AE. Celui-ci partage des poids avec le réseau de complétion et effectue une reconstruction de l'objet partiel. Cet ajout permet d'améliorer les caractéristiques globales de l'encodeur et par le fait même la prédiction finale. Xia et al. (Xia et al., 2021) ont proposé ASFM-Net, un nouveau réseau entraîné avec une stratégie de correspondance entre caractéristiques. Leur entraînement a pour but de faire correspondre les caractéristiques globales de complétion avec celles de reconstruction d'objets complets. Leur approche nécessite un AE pré-entraîné qui permet d'obtenir la représentation latente idéale que l'encodeur de nuages de points partiels pourrait fournir. Une fonction de coût pour évaluer la correspondance entre la représentation latente idéale et celle fournie par le décodeur est utilisée dans leur entraînement. Leur méthode a permis de surpasser plusieurs méthodes mentionnées précédemment telles que PCN, TopNet, GR-Net, CRN et SCRNet sur la tâche « Completion3D ».

Contrairement aux approches précédentes, pcl2pcl (Chen et al., 2019) propose d'effectuer la complétion de nuage de points de manière non supervisée pour être directement applicable à des

modèles réels. Cette approche entraîne deux AE séparément pour reconstruire des formes complètes et partielles et elle apprend la correspondance entre les espaces latents de chacun. Plus précisément, un GAN est entraîné pour générer des caractéristiques latentes complètes à partir de la représentation latente partielle d'un nuage de points incomplet. Le décodeur de caractéristiques complètes est ensuite utilisé pour obtenir une forme complète et réaliste. Les formes obtenues sont complètes et réalistes, cependant la forme peut être légèrement différente. Pour donner suite à ces travaux, Zhang et al. proposent de pré-entraîner un GAN pour générer des formes complètes et d'effectuer une recherche dans l'espace latent pour obtenir la forme qui reconstruit le mieux celle de l'entrée (Zhang et al., 2021). Leur approche permet de générer plusieurs complétions réalistes pour une entrée partielle. Inspiré par l'apprentissage par renforcement (RL), Sarmad et al. proposent un agent de RL pour contrôler un GAN qui complète des formes partielles (Sarmad et al., 2019). Leur approche requiert de préalablement entraîner un AE pour obtenir des représentations latentes des objets partiels. Par la suite, les caractéristiques encodées sont utilisées pour entraîner un GAN dans l'espace latent. Finalement, un agent RL est entraîné avec un environnement qui comprend l'AE et le GAN afin de trouver une entrée adéquate pour le GAN pour générer une forme qui complète bien le nuage de points partiel.

Contrairement aux approches précédentes, celles mentionnées dans le présent paragraphe se concentrent uniquement sur la génération de la partie manquante. Inspiré par les approches de peintures d'images et L-GAN, Yu et al. (Yu et al., 2020) propose Point Encoder GAN pour compléter des nuages de points. Leur approche effectue un apprentissage supervisé avec un GAN pour générer uniquement les points manquants. Le discriminateur a pour but de classifier les formes manquantes qu'il reçoit comme étant réelles ou artificielles. De son côté, le générateur a pour but de générer des formes de dents réalistes qui permettent de tromper le discriminateur et de générer des formes similaires à des formes de références. Cependant, le générateur n'arrive pas à bien modéliser les détails. Par une analyse multi-résolution, Huang et al. (Huang et al., 2020) développe l'architecture PF-Net. Cette architecture est présentée à la Figure 2.3.

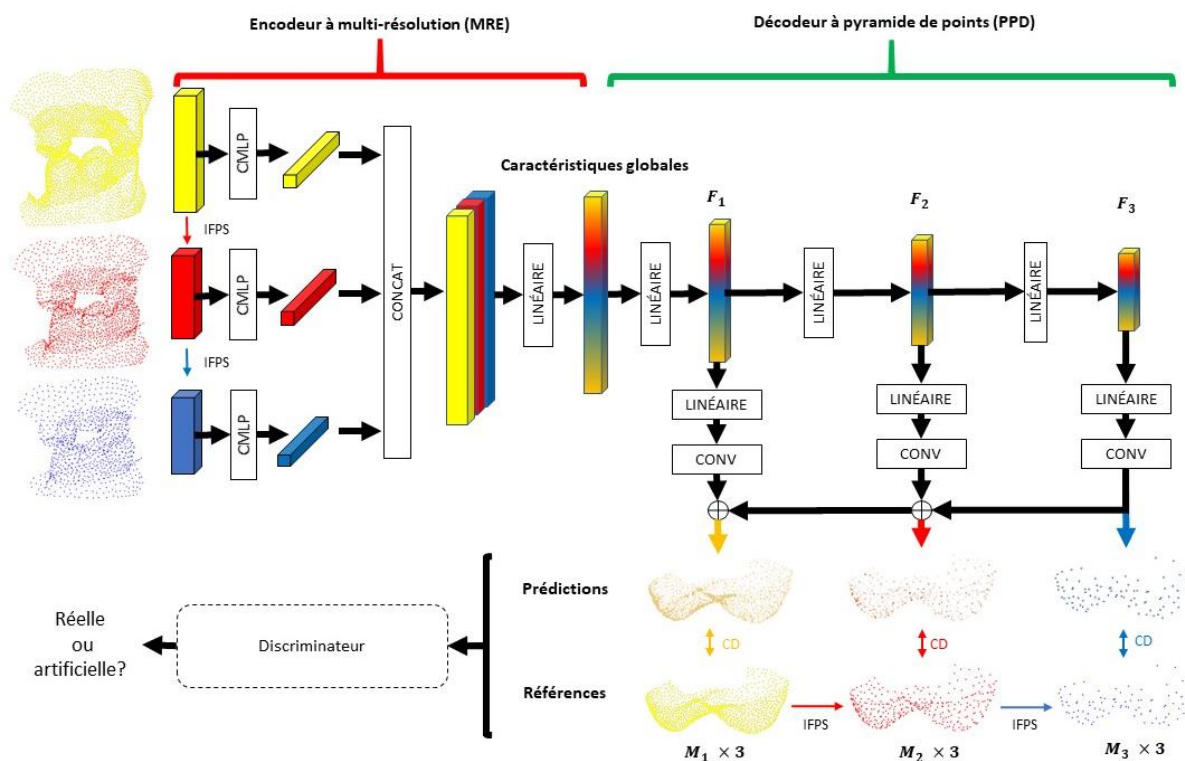


Figure 2.3 Architecture de PF-Net.

L'approche utilise 3 résolutions de l'entrée obtenues avec « Iterative Farthest Point Sampling » (IFPS). L'encodeur comprend des combinaisons de perceptrons multicouches appelées CMLP. PF-Net estime les points manquants hiérarchiquement et utilise une combinaison de fonction de coût adversarielle et de fonction de coût de complétion à multiples résolutions. Leurs résultats démontrent que PF-Net permet de prédire les points manquants avec précision contrairement à Point Encoder GAN. L'architecture de PF-Net a été récemment améliorée par Zhu et al. (Zhu et al., 2021) en utilisant des structures de graphes. Ces derniers proposent des opérations sur les graphes pour mieux combiner les caractéristiques des différentes résolutions et d'utiliser le « Point Rank Sampling » (PRS) plutôt que IFPS afin d'obtenir un meilleur échantillonnage. Wen et al. ont proposé de faire une connexion directe entre la partie manquante prédite et l'entrée partielle pour raffiner davantage la prédiction et mieux guider la complétion à l'aide des caractéristiques locales de l'entrée (Wen et al., 2020). Dans (Yu et al., 2021), les auteurs proposent plutôt d'utiliser les Transformers (Vaswani et al., 2017) pour apprendre les longues dépendances entre les parties locales des nuages de points. Pour ce faire, ils formulent la tâche de complétion de nuage de points

comme étant une traduction d'ensemble de points. Tout comme PF-Net, leur modèle nommé PoinTr utilise une fonction de coût à multiples résolutions afin de prédire uniquement la partie manquante.

2.4 Reconstruction de surface à partir de nuages de points

Dans plusieurs situations, une génération de maillages 3D à partir d'un nuage de points doit être faite. Pour obtenir un maillage triangulaire à partir d'un nuage de points non structuré, une reconstruction de surface doit être faite. Plusieurs méthodes permettent de réaliser cette tâche. Comme mentionné dans la revue de littérature de Berger et al. (Berger et al., 2014), les algorithmes de reconstruction peuvent porter soit sur la reconstruction urbaine, la complétion de surface ou l'interpolation de surface. L'interpolation de surface sera présentée plus en détails dans cette section.

La triangulation de Delaunay (Cazals et Giesen, 2004) permet d'obtenir un ensemble de triangles à partir de l'ensemble de points. La reconstruction assure qu'aucun point n'est compris dans un cercle formé avec les trois sommets de chaque triangle.

2.4.1 Reconstruction assumant une surface lisse globale

Lorsque les normales orientées d'un nuage de points sont connues, la reconstruction de poisson (Kazhdan et al., 2006) peut être utilisée pour produire des surfaces lisses. Cette méthode se base sur une fonction implicite qui permet de distinguer la partie intérieure de la partie extérieure de la forme. La reconstruction de poisson permet de trouver une fonction ayant un gradient autant près que possible du champ de normales N selon $\|\nabla X - N\|_2^2$ ou X représente le maillage reconstruit. Cependant, cette méthode assume que les points sont uniformément distribués et que les normales sont bien orientées.

Lorsque les normales ne sont pas connues ou difficiles à estimer, une reconstruction avec « ball pivoting » peut être utilisée (Bernardini et al., 1999). Cette technique utilise des balles avec des rayons prédéterminés afin de former des triangles à partir de trois points s'ils satisfont une condition. Pour former un triangle, les points doivent rentrer dans une balle sans contenir d'autres points. Les triangles sont formés en utilisant un triangle de départ et en explorant les contours de celui-ci. Le processus continue jusqu'à ce qu'il ne soit plus possible de former de triangle voisin,

pour ensuite choisir un nouveau point de départ jusqu'à ce que tous les points aient été considérés. Cette technique a l'avantage d'obtenir un maillage triangulaire sans modifier les coordonnées des points. Cependant, cette méthode ne permet pas de réduire le bruit dans les surfaces.

2.4.2 Reconstruction avec gabarit

Chen et al. (Chen et al., 2021) ont proposé d'apprendre à déformer un gabarit selon un nuage clairsemé de points. Au lieu de déformer un gabarit en plusieurs itérations, leur modèle prédit directement la déformation nécessaire à appliquer. Leur méthode génère un maillage 3D à partir d'un nuage de points. L'hypothèse utilisée dans leur approche est que le nuage de points est clairsemé, incomplet et peu bruité. L'utilisation du gabarit permet de déduire l'information manquante dans le nuage de points. Une conversion en volume 3D est utilisée afin d'obtenir la correspondance entre les points donnés en entrée et les arrêtes du gabarit avant d'effectuer la déformation. En revanche, leur conversion en volume 3D limite leur résolution en sortie et l'utilisation du gabarit peut entraîner une perte de détails initialement présents dans le nuage de points.

2.5 Conception de couronne dentaire avec apprentissage profond

Dans (Hwang et al., 2018), les auteurs convertissent les modèles dentaires 3D en image de profondeur pour ensuite appliquer une approche d'apprentissage profond 2D basée sur les GAN tel que pix2pix (Isola et al., 2017). Leur méthode permet de générer des images de profondeur de couronnes dentaires. Ces images sont ensuite utilisées pour reconstruire une forme 3D. Dans cet article, ils utilisent une fonction pour mesurer la fonctionnalité de la dent afin d'améliorer leurs résultats. Dans Yuan et al. (Yuan et al., 2020), les auteurs proposent d'utiliser une approche très semblable à la précédente en ajoutant une comparaison entre les creux des dents générées et les creux réels. Ces approches ont démontré l'importance d'inclure les dents opposées dans l'information conditionnelle. Elles utilisent toutefois une conversion des données pour appliquer des approches 2D. Cette conversion implique nécessairement une perte d'information qui n'est pas souhaitable.

Récemment, Ping et al. (Ping et al., 2021) ont proposé une approche utilisant les voxels intitulée SA-IFN qui combine les fonctions de représentations implicites avec le mécanisme d'auto-attention pour compléter des scans de dents 3D. Leur méthode permet de compléter des dents dans

les scans avec une distance de chanfrein de $0.069 * 10^{-2}$. Cependant, cette approche implique beaucoup de prédictions inutilisées qui ne correspondent pas à la surface de la dent. Ces prédictions limitent leur résolution.

Dans ce mémoire, les méthodes 3D ont été choisies pour éviter la perte d'information. Ce travail constitue une des premières applications d'approches 3D pour la conception de couronnes dentaires utilisant les nuages de points. Le chapitre suivant couvrira plusieurs méthodes 3D qui ont servi d'inspiration pour ce travail.

En raison des nombreuses publications sur les nuages de points et du faible besoin de ressources avec ceux-ci, les nuages de points ont été choisis comme type de données pour débiter cette recherche. Afin de compléter des maillages surfaciques de dents, nous avons choisi d'utiliser une complétion de nuage de points suivi d'une reconstruction de surface en raison de la popularité et de la simplicité de ces méthodes. Vu que le modèle développé a pour but d'être mis en production ultérieurement, nous avons choisi d'étudier les méthodes de complétions de points basées sur l'apprentissage pour avoir un court temps d'inférence et une bonne généralisation.

CHAPITRE 3 DÉMARCHE DE L'ENSEMBLE DU TRAVAIL DE RECHERCHE ET ORGANISATION GÉNÉRALE DU DOCUMENT INDIQUANT LA COHÉRENCE DES ARTICLES PAR RAPPORT AUX OBJECTIFS DE RECHERCHE

Lorsqu'une dent est abimée ou cariée, une couronne dentaire est habituellement conçue afin de restaurer la dent et de la protéger contre d'autres dommages. La génération de couronne dentaire implique la conception d'une forme 3D complexe. Cette forme comprend une surface externe visible une fois la couronne installée. En pratique, les techniciens dentaires utilisent des numériseurs intraoraux et des logiciels CAD pour réaliser cette tâche. Leur conception demande beaucoup de temps et d'expérience pour être bien réalisée, ce qui entraîne une grande variabilité dans la qualité des couronnes dentaires.

Jusqu'à maintenant les conceptions de couronnes sont toujours faites manuellement. Une des hypothèses qui motive l'équipe de recherche est que la génération de couronne dentaire peut être automatisée grâce à l'apprentissage profond. Nous croyons que les modèles 3D utilisés par les techniciens peuvent servir à développer des algorithmes d'intelligence artificielle permettant d'aider les techniciens dentaires dans leur conception.

Ce mémoire se concentre sur la complétion de dents afin d'éventuellement étudier la génération de couronne dans de futurs travaux. La complétion de dents est étroitement liée à la génération de couronne. Une distinction majeure est que la génération de couronne complète implique la conception de 2 surfaces tandis que la complétion de dents en implique une seule. Dans les deux situations, la génération doit fournir une surface externe lisse avec détails qui ne devrait pas chevaucher les dents voisines. Pour ce mémoire, le système développé doit donc arriver à compléter des dents automatiquement en 3D.

La recherche en apprentissage géométrique a connu dans les dernières années des avancées majeures permettant d'explorer de nouvelles applications. Avec l'engouement pour les données 3D, la complétion de données 3D gagne de plus en plus d'attention auprès des chercheurs en apprentissage profond.

Malgré cela, le déploiement de ces solutions algorithmiques est encore difficile pour quelques raisons :

- La quantité de données 3D est encore limitée et freine le déploiement d'algorithmes d'apprentissage profond. Il n'existe actuellement pas de grande base de données 3D de modèles dentaires publiques en raison de la protection des données dans le domaine de la santé.
- Jusqu'à récemment, les solutions en complétion de nuages de points manquaient de détails et l'analyse de données 3D était encore difficile.
- La conception d'une couronne dentaire est une tâche complexe que les gens ne savent pas programmer explicitement. Pour cette raison, la conception est encore majoritairement manuelle, ce qui entraîne une variabilité dans la qualité des couronnes et demande du temps.

À la vue de ces contraintes, le défi est de réussir à concevoir une base de données 3D pour entraîner un algorithme d'apprentissage profond à compléter des modèles de dents en 3D. Cette base de données doit être suffisamment riche pour représenter toutes les dentitions et elle doit contenir des formes de références approuvées par un technicien dentaire. Pour ce faire, nous proposons d'utiliser une approche de complétion de nuages de points.

Question de recherche : Peut-on utiliser la complétion de nuage de points pour générer des surfaces de dents manquantes?

L'objectif principal de ce mémoire est de compléter des maillages surfaciques de dents par apprentissage profond. Dans ce mémoire, des modèles de dents incomplets seront complétés grâce à un système générant la partie manquante sous forme de maillage.

Le premier objectif spécifique est de concevoir une base de données 3D. Pour réaliser cette base de données, des techniciens dentaires identifieront des dents considérées comme étant parfaites afin d'être utilisées comme référence et fourniront une segmentation de dent manuelle. La base de données comprendra des parties de mâchoires partielles et leur parties manquantes respectives. Elle servira à entraîner des modèles d'apprentissage profond à compléter des modèles dentaires.

Le second objectif est de concevoir un modèle pour compléter des dents sous forme de nuage de points. Ce modèle servira à prédire des dents complètes ou partielles à concevoir. Pour atteindre cet objectif, un système basé sur les modèles antagonistes génératifs a été conçu. Récemment, plusieurs travaux ont été réalisés en complétion de nuages de points. Ces travaux sont intéressants puisqu'ils permettent de prédire une forme 3D manquante en utilisant une entrée 3D.

Le troisième objectif est de concevoir un algorithme pour obtenir des maillages à partir des nuages de points générés. Cette reconstruction a pour but de fournir une surface lisse sans trous avec un contour. Grâce à cette reconstruction, le système peut compléter des dents sous forme de maillage. Finalement, le dernier objectif est d'évaluer la génération ainsi que sa rapidité. Pour ce faire, la distance de chanfrein sera utilisée afin d'évaluer la ressemblance entre les maillages de références et les maillages prédits.

Le premier article présenté dans le chapitre suivant présentera des complétions de dents utilisant l'architecture PF-Net (Huang et al., 2020). Cet article répondra aux deux premiers objectifs et l'évaluation du modèle y sera présentée.

Le second article présenté au chapitre 5 démontrera la pertinence des approches de déformation de maillage pour effectuer des complétions de maillages. Dans celui-ci, des maillages obtenus grâce à une complétion de nuage de points avec PF-Net et une reconstruction de surface seront comparés à ceux produits par une nouvelle architecture de réseaux de neurones générant directement des maillages. Cet article permettra de répondre à tous les objectifs spécifiques.

Au chapitre 6, des expériences permettent de compléter les résultats des articles et de présenter l'effet d'une nouvelle fonction de coût. Cette fonction a pour but de générer des dents broyant mieux les aliments. Elle encourage un réseau de neurones à générer de meilleurs nuages de points pour répondre au deuxième sous-objectifs.

CHAPITRE 4 ARTICLE 1: DENTAL RESTORATIONS USING A MULTI-RESOLUTION DEEP LEARNING APPROACH

Olivier Lessard¹, François Guibault¹, Julia Keren², Farida Cheriet¹

1. Department of Computer and Software Engineering, Polytechnique Montréal, Montréal, QC, Canada
2. Intellident Dentaire Inc., Montréal, QC, Canada

Published on the 26th April 2022 to the 2022 19th IEEE International Symposium on Biomedical Imaging

4.1 Abstract

Computer assisted design software is currently used by technicians to design dental crowns. However, this process involves manual adjustments that are time consuming and lead to great variability in quality of the design since they depend on the technician's experience. We developed a fully automatic approach that learns from natural teeth in dental scans using 3D conditional shape completion. Our work extends depth map-based approaches to generate crown shapes in 3D directly. Using a Generative Adversarial Network (GAN), our deep learning model is able to generate patient-specific point clouds of teeth starting from normalized incomplete point clouds. The model generates a crown's outer surface that looks realistic, with a mean Chamfer Distance (CD) of 0.55 millimeters when compared to real teeth.

4.2 Introduction

In dental centers, technicians need to manually design hundreds of dental crowns per year. Crown generation implies the design of a mesh surface that forms the external surface of the crown. The designed crown should not overlap with existing teeth and gingiva, and it must respect a space occlusion relationship with the opposing teeth to provide adequate chewing function. It should fit the patient's dentition and should have an aesthetically plausible shape. In the dental industry, the dentist first prepares the tooth where a crown is required. Secondly, using an intra-oral scanner, he scans the prepared tooth and its surroundings to have a 3D representation. A technician then uses that surrounding information to design the shape of the crown with respect to the aforementioned requirements.

The next steps of the design process are done with computer aided design (CAD) systems. These systems provide many tools to help the professional satisfy the dental requirements. However, human intervention is still much needed. Usually, the dental technician starts by picking a standard tooth model from a database, then manually adjusts it through positioning and scaling. The adjustments may involve complex transformations to get good occlusal relationships, grooves and cusps. According to dental technicians, the lengthiest task is the positioning and scaling to make the crown realistic. Thus, an automatic approach could reduce the time required for this step.

There have been a few studies to automate dental crown design. In two of these works (Hwang et al., 2018; Yuan et al., 2020), the authors convert the 3D scans to 2D depth maps to be able to use deep learning approaches based on pix2pix (Isola et al., 2017). In both approaches, it was shown that adding the opposing jaw helped the generation. By doing the 3D to 2D data conversion, though, some potentially useful information is lost. Recently, Ping et al. (Ping et al., 2021) propose a voxel-based approach (SA-IFN) by combining implicit function representations with a self-attention mechanism. They removed segmented teeth from dental scans and used their model to infer the missing teeth. However, this approach leads to unnecessary occupancy prediction of many points that are not part of the crown's outer surface. In addition, the marching cubes surface mesh generation step may limit the output resolution of the dental model.

3D deep learning has significantly advanced in recent years and many approaches are now used in real applications. Among various types of 3D data, point clouds have been studied extensively because of their small data size, especially since PointNet (Qi et al., 2017). Huang et al. (Huang et al., 2020) proposed PF-Net, a point cloud completion approach using a partial point cloud to produce the missing region. This approach preserves the spatial arrangements of the input data and estimates the missing points hierarchically by using a feature-point-based multi-scale generating network. PF-Net was evaluated on a large database of 13 object categories such as planes, cars and chairs. The incomplete point data was generated by randomly removing 25% of the original data for training and testing.

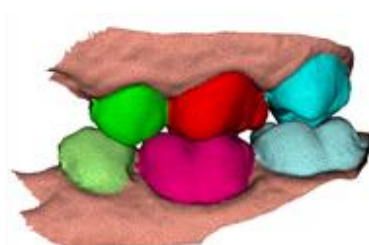
The main contribution of the present work is to adapt the architecture of PF-Net to generate a dental crown's outer surface from natural teeth in dental scans and the surrounding information typically used by the technician to design the shape of the crown. We changed the three output resolutions of the PF-Net model to be able to capture the semantics of complex shapes such as teeth.

Furthermore, Unlike PF-Net, where an object’s partial point cloud is taken as input, in our application only teeth-missing scans are used. These scans are refined to look more realistic, unlike the ones in SA-IFN. This will allow us to reuse the trained model on real teeth-missing scans. We evaluated our approach qualitatively and quantitatively.

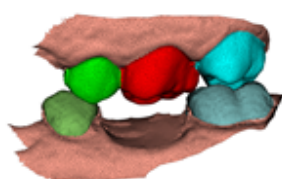
4.3 Our approach

We use a deep learning approach and formulate the dental restoration task as a point cloud completion problem. The meshes obtained from the input 3D scans are converted to point clouds to simplify the generation problem. We use point clouds of natural teeth with normal shapes as our ground truth and teeth-missing contexts as model input. By doing so, we learn from the shape of natural teeth directly. We generate realistic crown shapes by using an adapted PF-Net and compare our generated shapes with existing natural tooth shapes.

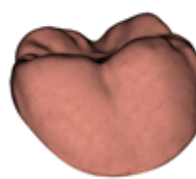
In this paper, we use the term “context” to refer to the surroundings of the tooth needing a crown. The context includes the two teeth adjacent to the preparation and the three closest teeth in the opposing jaw. Cases involving a crown design for the last molar were not considered in this study.



a) Original jaw (lower) and opposing jaw (upper)



b) Conditional information
(context)



c) Natural tooth shape

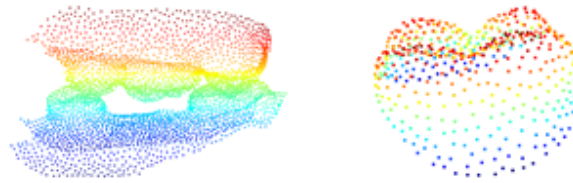


Figure 4.1 Illustration of the problem formulation. A 3D model of the original teeth (a), the conditional information referred to as the context (b), and the extracted tooth (c) used as the ground truth. Their point clouds are shown at the bottom.

We first create contexts from complete 3D scans. When dental technicians design crowns, they try to give them a natural shape. To use as much as possible of the available data, we use natural teeth as our ground truth, then replacing them with simulated gingiva as shown in Figure 4.1. The sides of the adjacent teeth are closed to make the surroundings realistic. These corrections are important because teeth always have closed sides. The model then uses the context thus obtained as input and predicts the natural tooth shape that the crown should have. The generated point cloud is then converted to a triangle mesh to be used in a CAD software.

The recent PF-Net model (Huang et al., 2020) has shown promising results on the Shapenet-Part dataset. Unlike previous point cloud completion methods (Yuan et al., 2018), PF-Net predicts the missing points only. The network uses an encoder-decoder as a generator G to predict the missing points Y and a discriminator D to classify them as real or fake. The encoder is composed of a Multi-Resolution Encoder (MRE) that uses different resolutions of the input X to create a final feature vector. The decoder is a Point Pyramid Decoder (PPD). It hierarchically creates the missing part from the features by first predicting primary centers denoted as $Y_{primary}$, then secondary centers denoted as $Y_{secondary}$, and the final prediction denoted as Y_{detail} . Iterative farthest point sampling (IFPS) is used to create different resolutions of the inputs and ground truths, this sampling strategy was used in PointNet++ to get a set of skeleton points (Qi et al., 2017). The discriminator is a classification network composed of a feature extractor block followed by a multilayer perceptron. The architecture uses a loss function inspired by GANs (Goodfellow et al., 2014). The adversarial loss is formulated as:

$$L_{adv} = \sum_{1 \leq i \leq S} \log(D(y_i)) + \sum_{1 \leq j \leq S} \log(1 - D(G(x_j))) \quad (4.1)$$

where $x_j \in X$, $y_i \in Y$ and S is the size of the dataset. x_j is a context point cloud and y_i is a ground truth point cloud. The model also uses a multistage completion loss to measure the similarity between the prediction and the ground truth. The metric used for that measurement is the Chamfer Distance (CD) proposed by Fan et al. (Fan et al., 2017) because it is differentiable and its computation is efficient:

$$d_{CD}(S_1, S_2) = \frac{1}{S_1} \sum_{x \in S_1} \min_{y \in S_2} \|x - y\|_2^2 + \frac{1}{S_2} \sum_{y \in S_2} \min_{x \in S_1} \|y - x\|_2^2 \quad (4.2)$$

CD measures the average nearest squared distance between the predicted point cloud S_1 and the ground truth S_2 in both directions. The completion loss uses the CD on the three output resolutions of the decoder's hierarchical prediction. It is formulated as:

$$L_{com} = d_{CD}(Y_{detail}, Y_{gt}) + \alpha d_{CD2}(Y_{primary}, Y'_{gt}) + 2\alpha d_{CD3}(Y_{secondary}, Y''_{gt}) \quad (4.3)$$

where d_{CD} is the Chamfer Distance and α is a weight. The final loss is:

$$L = \lambda_{com} L_{com} + \lambda_{adv} L_{adv} \quad (4.4)$$

λ_{com} and λ_{adv} are weights that satisfy the relation $\lambda_{com} + \lambda_{adv} = 1$.

4.4 Experimental setting and results

Our experimental dataset was composed of 160 healthy teeth. The training, validation and test sets included 93 teeth from 5 patients, 24 teeth from 2 patients and 43 teeth from 2 patients, respectively. The dataset included teeth in various positions in the jaw, namely molars, canines and incisors.

Our training procedure was similar to the one used in PF-Net (Huang et al., 2020). The input contexts were downsampled with IFPS to get 4000, 2000 and 1000 points. The input point clouds were normalized using z-score normalization along each axis. We changed the output resolutions to $M_1 = 128$, $M_2 = 256$ and $M = 768$ points to adapt the existing approach to our dataset. The

low and medium resolutions needed to be increased because the tooth shapes are more complex than those in the Shapenet-Part dataset. Indeed, grooves and cusps can't be properly reconstructed with only 64 low resolution points. The training set was augmented using random translations, rotations and scaling. We alternated between one update of the discriminator and one update of the generator. The Adam optimizer was used with a learning rate of 0.001 and momentum parameters of 0.9 and 0.999. The training batch size was 6.

To evaluate our generation, we used the CD between the high-resolution output Y_{detail} and the ground truth, both containing 768 points. The context information is not used when we evaluate the prediction because PF-Net preserves the spatial arrangement of the incomplete point cloud. The CD uses a bidirectional measurement, from prediction to ground truth $Pred \rightarrow GT$ and ground truth to prediction $GT \rightarrow Pred$. The first one measures how different the prediction is from the ground truth while the second one indicates the extent to which the ground truth surface is covered by the predicted shape. Some post-processing was used to remove outliers. Points without neighbors within a sphere of radius 0.2 were removed. In addition, the mean distance from every point to its five closest neighbors was computed. Those with a distance three standard deviations over the mean value were removed. The mean CD of the test set is shown in Table 4.1.

Table 4.1 Test Chamfer Distance.

| | CD | Pred2GT | GT2Pred |
|--------|------------------|------------------|------------------|
| PF-Net | $2.27 * 10^{-2}$ | $2.51 * 10^{-2}$ | $2.03 * 10^{-2}$ |

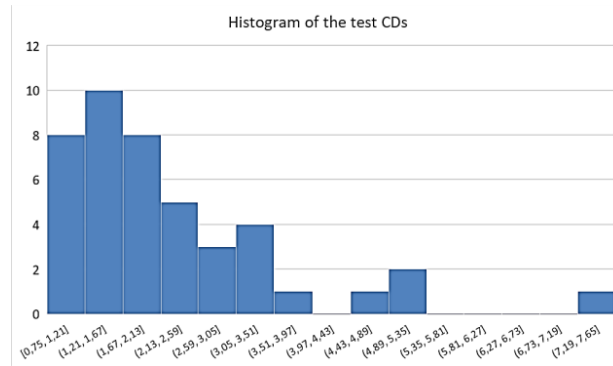


Figure 4.2 Histogram of test CDs. The horizontal axis corresponds to intervals of CD values scaled by 100.

The Chamfer Distance is $2.27 * 10^{-2}$ with the normalized data as shown in table 1, which correspond to an average of 0.55 millimeter error. The individual test CDs were used to compute the histogram shown in Figure 4.2. This is useful to identify difficult or under-represented cases in the training set. Examples of generated teeth are shown in Figure 4.3, while Figure 4.4 shows a sample prediction with its surroundings.

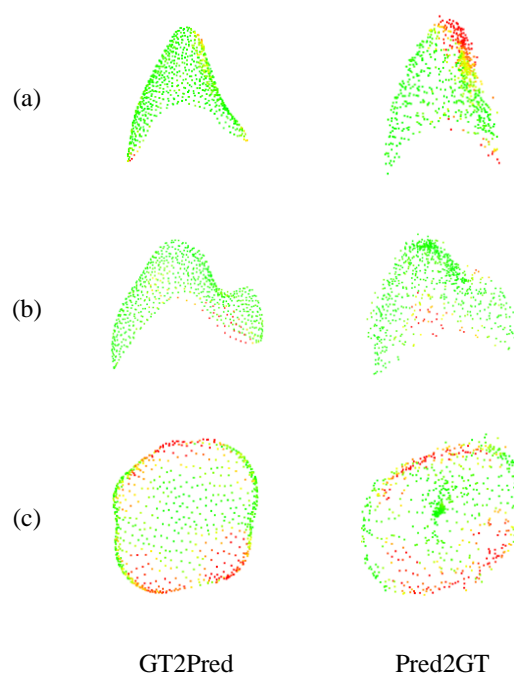


Figure 4.3 (a) a canine, (b) a premolar, (c) a molar seen from above. Green represents points with errors close to zero when red means errors of 1 millimeter or more. The left and right shapes are examples of GT2Pred and Pred2GT errors.

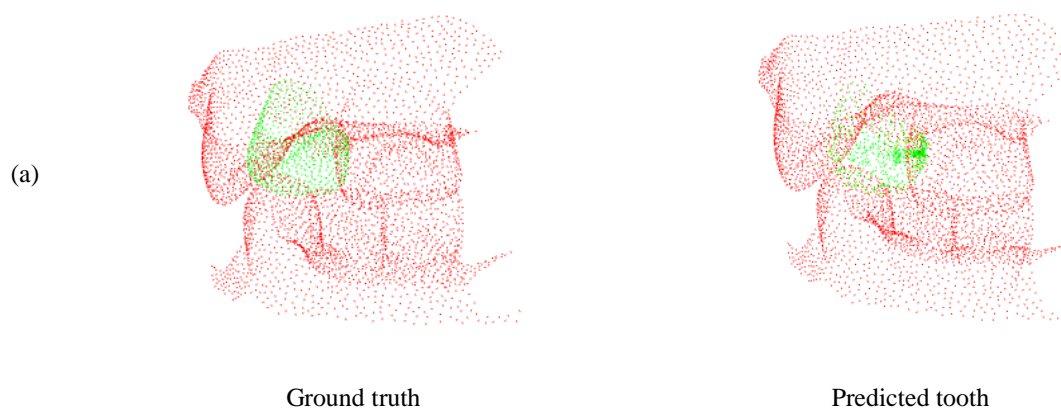


Figure 4.4 Predicted tooth and corresponding ground truth shown in green with their surroundings in red.

4.5 Discussion

In Figure 4.3, the model learns the various tooth shapes found in different positions in the jaw. Many red points are located at the bottom boundary. These errors are caused by the absence of preparation, which makes the generation challenging. A wrong positioning of the bottom can lead into incorrect positioning of the cusp. In Figure 4.3, most test cases have a mean Chamfer Distance in the interval $[0.75 * 10^{-2}, 3.97 * 10^{-2}]$. However, four cases have higher values. Three of those four challenging cases were designs for canines. They were more challenging because there were fewer canines than incisors and molars in the training set. A specialized data augmentation technique to create canine exemplars could alleviate this problem. In another instance, a slightly chipped tooth adjacent to the one to generate adversely affected the result. To obtain a more robust model, the training set should include more special cases such as this.

In some situations, the generated teeth have smaller grooves and cusps than the ground truths. This is likely because the training set comprises some teeth that have contact points with opposing teeth and some that do not. When the predicted point cloud is globally smaller, the post processing may remove outliers that are close to the ground truth points. This reduces the noise in the output, but also increases the CD. Nonetheless, since our goal is to generate a surface, we prefer to remove the outliers.

In Figure 4.4, the generated shape looks realistic. Indeed, it has a good position, size and orientation, and its overall morphology is correct. However, it may not have adequate occlusal contact points. The dental technician will thus need to verify this and make any necessary adjustments.

4.6 Conclusion

Our trained model can generate a realistic personalized crown shell for dental restoration. The method does not currently consider the shape of the tooth preparation. As future work, we will adjust the generated shell to fit onto the preparation and generate the inner surface to obtain a complete dental crown. A functionality-aware loss function could help to generate shapes with better occlusal relationships.

4.7 Acknowledgments

This work was funded by Kerenor Dental Studio, Intellident Dentaire Inc., iMD Research Inc., the Natural Science and Engineering Research Council of Canada (NSERC), the Institut de valorisation des données (IVADO), and MEDTEQ. We thank Compute Canada for providing the computational resources and acknowledge the help from JACOB Inc., Object Research Systems Inc and Philippe Debanne. This work has not been submitted elsewhere and it has been approved by the ethic committee for research with humans of Polytechnique Montreal.

CHAPITRE 5 ARTICLE 2: MC-NET: MESH COMPLETION FOR DENTAL SCANS

Olivier Lessard¹, François Guibault¹, Julia Keren², Farida Cheriet¹

1. Department of Computer and Software Engineering, Polytechnique Montréal, Montréal, QC, Canada
2. Intellident Dentaire Inc., Montréal, QC, Canada

Submitted on the June 24th 2022 the journal *Transactions on Medical Imaging*

5.1 Abstract

Dental centers need to design hundreds of dental crowns per year using computer assisted design (CAD) software. Typically, technicians manually modify a template shape to create crowns. That process requires a lot of time and experience to be done correctly, which leads to great variability in quality. In recent years, many deep learning methods have been proposed to do point cloud completion by predicting only the missing region. Although these methods are potentially applicable to the task of dental crown design, most of them fail to generate smooth point clouds, which is critical for surface reconstruction. In this paper, we propose an end-to-end approach called MC-Net for automatic mesh completion of dental scans. Using an input point cloud sampled at multiple resolutions and a template shape for the type of tooth to generate, MC-Net extracts features to guide a mesh deformation. The mesh generation follows a coarse-to-fine strategy and uses a mesh-related loss function to make the procedure stable. Our model can generate visually correct and accurate surfaces of the missing regions.

5.2 Introduction

In dental centers, technicians need to manually design hundreds of crowns per year. Crown generation implies designing an external surface that is visible once the crown is installed on the tooth preparation. That design is challenging even with today's computer assisted design (CAD) software. The technician must design a tooth shape with a complex morphology and patient-specific characteristics. Usually, the dentist starts by making a preparation with the damaged tooth. The preparation serves as a foundation on which the dental crown will be installed. Secondly, the

dentist scans the prepared tooth and the surrounding teeth with an intra-oral scanner to get a 3D representation. Finally, the technician uses that surrounding information to design a patient-specific crown.

Once the technician has the 3D scan of the patient's jaw, they will use a CAD system to create the crown shape. Modern systems have many dental application-specific tools to assist this process. However, considerable human intervention is still needed. Usually, the technician starts by picking a template tooth from a library and adjusting it through positioning and scaling to account for functional and esthetic constraints. Then, complex transformations may be applied to create contact points with the surrounding teeth. Thus, an automatic approach to generate patient-specific crowns with the desired positioning and scaling would be very useful. Also, an approach to create patient-specific teeth could help technicians to design more personalized crowns.

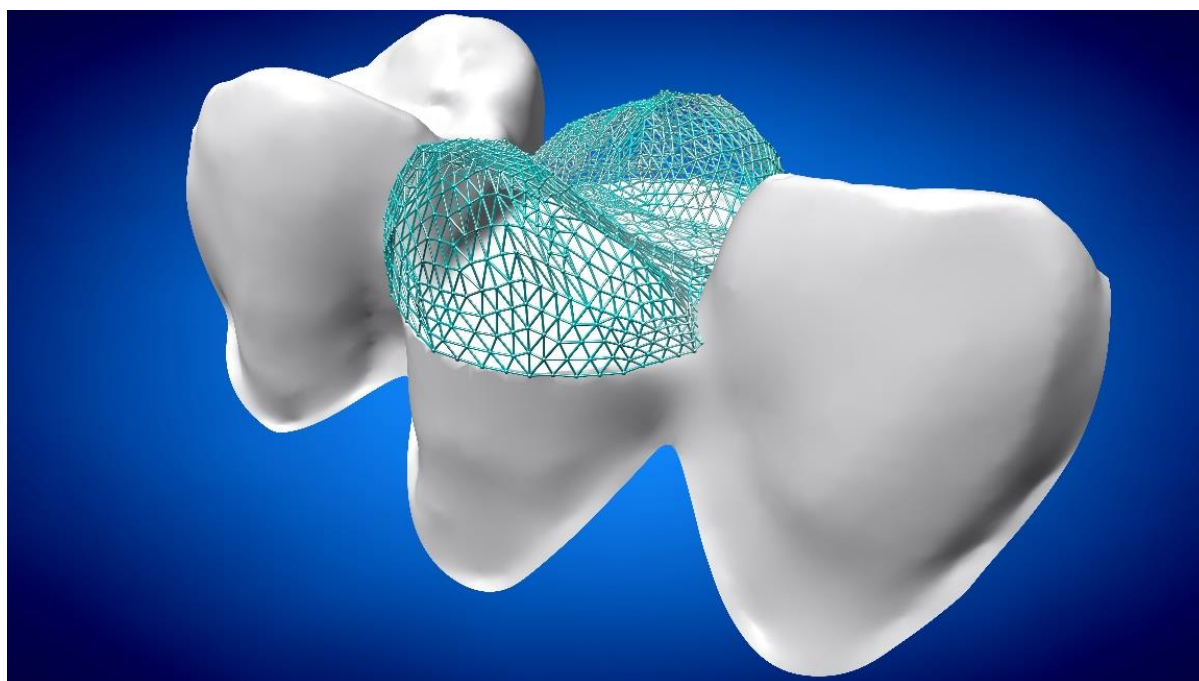


Figure 5.1 Illustration of the tooth completion problem. Using a point cloud of the surrounding teeth, our model predicts the missing surface that the tooth should have, as shown by the mesh in cyan.

In this paper, we focus on the task of tooth completion. An example of tooth completion is shown in Figure 5.1. From this figure, we can see that tooth completion implies generating a missing shape to get a complete arch. The missing shape corresponds to a mesh surface similar to a crown's external surface. In both cases, the surface needs to be smooth, detailed, and patient-specific.

The completion problem has been addressed in the computer vision community since real world 3D data is often captured by scanners. Due to occlusions, light reflections and limitations of the scanners, 3D data such as point clouds are often incomplete. The task of point cloud completion consists in generating a complete point cloud from a partial observation. Deep Learning (DL) methods became popular because of their ability to complete shapes even with large missing regions. Since the advent of PointNet (Qi et al., 2017), many methods were proposed to achieve point cloud completion. Among them, Huang et al. proposed PF-Net (Huang et al., 2020). Their architecture enables a multi-resolution representation of the completions and generates realistic missing regions. Contrary to most shape completion methods in the literature, missing shapes in tooth completion and dental crown design can be seen as complex shapes in and of themselves. Furthermore, in dental crown design, the technicians always know where they must be located. For

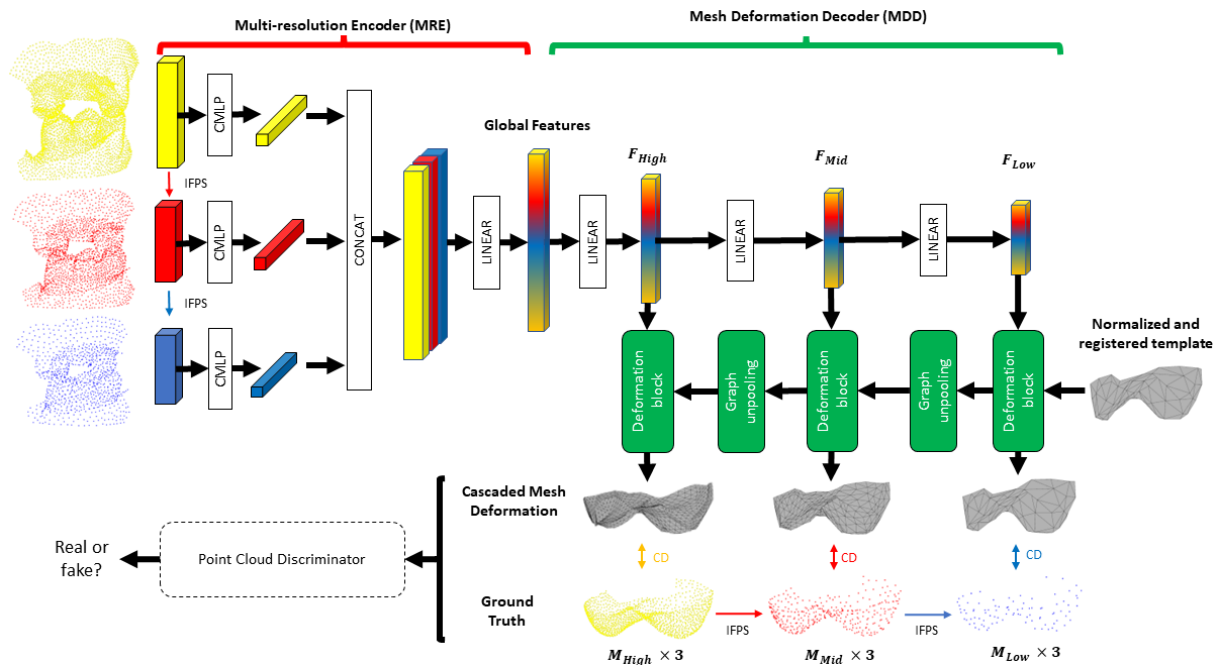


Figure 5.2 Architecture of MC-Net. The multi-resolution encoder (MRE) analyzes the three input resolutions to create global features. The mesh deformation decoder (MDD) uses the features to deform the template with three deformation blocks. (*CMLP* means *Combined Multi-Layer Perceptron*; *IFPS* is *Iterative Farthest Point Sampling*; *CD* is *Chamfer Distance*.)

this reason, we focus on tooth-specific completions.

There have been a few studies on how to automate dental crown design and how to complete teeth. In two of these works (Hwang et al., 2018; Yuan et al., 2020), the authors converted 3D dental scans to 2D depth images. Then, they used pix2pix-based approaches (Isola et al., 2017) to generate depth images of dental crowns. In this manner, they could benefit from the advances in image generation to achieve good performance. But their conversion from 3D to 2D necessarily involves information loss, especially on the sides of the teeth. By generating depth maps, these approaches focus on the crown's occlusal surface but neglect the lateral portions, which are equally important to produce an appropriate dental crown. The 2D approaches do not generate complete tooth surfaces directly and obtaining the lateral surfaces from the occlusion surface is not a trivial task. To complete dental scans, Ping et al. proposed a voxel-based approach called SA-IFN (Ping et al., 2021). Their approach uses implicit functions and a self-attention mechanism to capture nonlocal features to do shape completion. Their model can predict 3D voxels that represent the complete shapes, in contrast to previous methods. However, voxel-based approaches need a lot of

unnecessary occupancy predictions, and this limits the output resolution. Dental technicians always work with high resolution scans. Thus, a scalable approach that can work at higher resolution would be preferred. In (Lessard et al., 2022), the recently proposed PF-Net was used to complete teeth. The model can predict a variety of tooth shapes, but the predictions are sometimes wrongly positioned and noisy. The approach produces point clouds that cannot be directly used to reconstruct meshes. Zhu et al. (Zhu et al., 2022) used a transformer-based model to complete teeth point clouds and apply a new voxel-based reconstruction. However, their method reconstructs a watertight mesh instead of a thin surface with a boundary. In addition, by generating point clouds rather than meshes including connectivity, their method has little control over the smoothness of the resulting shapes.

The present work is an extension of a preliminary conference publication (Lessard et al., 2022) to achieve teeth completion. In practice, technicians use a preparation which is very important to guide their crown design. Similarly, in our proposed approach, we use the base of the partial tooth (referred to as a *sliced tooth*) to guide the tooth completion. The dataset to train our DL model now contains sliced teeth, which improves the positioning of the predictions. In addition, this paper presents 1) a new architecture using a formulation of mesh deformation that generates meshes instead of point clouds; 2) a comparison with the original PF-Net for mesh generation; and 3) an ablation study to evaluate the effectiveness of the key components of our model.

The contributions of this work are twofold: I) We propose to directly learn high-level geometric features of the dental scans for automatic tooth completion; II) We propose an end-to-end DL method called Mesh Completion Network (MC-Net).

MC-Net can be seen as an extension of PF-Net. First, instead of learning directly a low resolution representation of the missing region, the proposed network learns a deformation from a fixed template. The template introduces both mesh connectivity and dental shape characteristics. Second, we propose to replace the concatenation and addition mechanism of PF-Net by the edge-based graph unpooling proposed in Pixel2Mesh (Wang et al., 2018). This allows us to add new points at better positions and maintain the uniformity of the meshes. Third, we add a Laplacian regularization (Wang et al., 2018) to control the coarse-to-fine mesh deformation. This ensures a smooth deformation and avoids creating intersecting triangles and spikes. Finally, graph convolutions and a new formulation of the discriminator input are used to enhance the shape generation.

The rest of the paper is organized as follows. Related works concerning point cloud completion, mesh completion and mesh deformation are briefly reviewed in Section 5.3. The proposed method is explained in Section 5.4. Section 5.5 presents the experimental settings and results, including an ablation study. The limitations and future work are presented in Section 5.6. Finally, Section 5.7 concludes this paper.

5.3 Related work

5.3.1 Point cloud completion

Point cloud completion is crucial in 3D computer vision because 3D data are often incomplete. The task consists in predicting the complete shape of an object from a partial observation. This problem has received increasing attention in the computer vision community since the pioneering works of PointNet and FoldingNet (Qi et al., 2017a, 2017b; Yang et al., 2018). Many methods have been proposed to do point cloud completion. Some of them are based on generative adversarial networks (GAN) (Goodfellow et al., 2014). As mentioned in Fei et al. (Fei et al., 2022), these can be categorized into point-based, convolution-based, graph-based, folding-based, GAN-based, and transformer-based methods. Among these different architectures, some were proposed to predict only the missing region (Huang et al., 2020; Yu et al., 2021). For instance, the GAN-based architecture PF-Net uses a multi-resolution encoder and a point pyramid decoder to generate the missing points hierarchically. However, their resulting shapes lack geometric details. PF-Net does not use the connectivity between points, and thus the new points added in the upsampling mechanism have a large degree of freedom. Hence, the resulting point cloud does not represent a smooth surface, which makes the mesh reconstruction challenging. Recently, transformer-based approaches like PoinTr and SnowflakeNet (Xiang et al., 2021; Yu et al., 2021) have been proposed. These approaches require long training times and are memory-consuming, which is not ideal for practical applications in dentistry.

5.3.2 Mesh completion

Similarly, mesh completion involves predicting a complete mesh from a partial observation. Mesh completion can be used for either small missing regions or large occlusions. For small missing regions like holes, geometric priors, self-similarity or patch encoding can be used as in (Kazhdan et Hoppe, 2013; Sarkar et al., 2017). However, these methods work only for small missing regions.

To complete larger occlusions, model-based approaches are used to capture the variability of a shape category. In (Litany et al., 2017), the authors propose an autoencoder to learn a latent space of complete shapes to achieve possible completions. Their method can complete meshes directly without any conversion or reconstruction, but the resulting meshes do not have the precision needed for dental applications. In (Foti et al., 2020), the authors propose a similar method to complete surfaces of the liver. They train a variational auto-encoder (VAE) on complete shapes and perform an optimization at inference time. The method takes as input a partial point cloud and outputs a complete mesh. However, the two latter methods require a time-consuming optimization at inference. This is not suitable in our case as we would like to reduce the overall time required for a dental crown design. Importantly, only the missing tooth region is interesting in our application. Existing mesh completion methods generate the complete shape, but extracting the missing region afterwards is not trivial. By generating the missing surface only, a model can focus on the details of that region and maximize its resolution.

5.3.3 Mesh deformation

Deep learning has become a popular approach for mesh deformation because it can learn complex transformations that would be difficult to determine explicitly. Mesh deformation has been employed to reconstruct 3D meshes from images. Early works on mesh reconstruction used voxels to generate 3D shapes (Choy et al., 2016; Girdhar et al., 2016). However, these methods are memory consuming, which limits the resolution and constitutes an important limitation of this shape representation in practice. Instead, Fan et al. (Fan et al., 2017) propose to generate a point set from a single image. But due to the large number of degrees of freedom of point clouds, their predictions are noisy, which makes subsequent mesh reconstruction difficult. To overcome that problem, Wang et al. propose Pixel2Mesh to generate a 3D mesh directly (Wang et al., 2018). They formulated the reconstruction problem as a mesh deformation problem and use mesh losses to control a deformation of a template mesh. Their template is a uniform general shape that provides an initial connectivity. The deformation is done with graph convolution networks (GCN) and a perceptual pooling that provides meaningful local features. They later extended their architecture to reconstruct a mesh from multiple views (Wen et al., 2019). More recently, new methods have achieved better performance than Pixel2Mesh, especially on shapes with positive genus (Guillard et al., 2021; Shi et al., 2021; Smith et al., 2019). In this paper, we focus on the generation of shapes

with genus 0 (i.e., closed surfaces with no holes) as in Pixel2Mesh. Chen et al. proposed MR-Net (Chen et al., 2021) to reconstruct cardiac meshes from sparse point clouds using mesh deformation. They propose to use voxels to map the point clouds to template vertices to guide the deformation which is done using GCN blocks. However, that voxel correspondence limits the quality of the deformations.

The method we present here takes inspiration from all three approaches seen above. Indeed, point cloud completion methods analyze a 3D input to generate a missing 3D output. They are interesting mainly because dental scans contain a lot of information and point cloud encoders can analyze such input efficiently with light algorithms. However, point cloud outputs cannot be directly used by technicians. For this reason, mesh completion also influenced our methodology because those approaches provide an output in the form of a mesh, as required in our problem. However, such methods are designed for hole filling or completion without high precision requirements. Because of the ability of mesh deformation methods to generate shapes that capture important details, our insight was to propose an architecture that analyzes an input point cloud to deform a template mesh into an output mesh. The details of the architecture will be discussed in the next section.

5.4 Method

Our model is an end-to-end deep learning framework that takes a partial point cloud as input and generates the missing region in a mesh format. The overview of our framework is illustrated in Figure 5.2. The network is a GAN-based model (Goodfellow et al., 2014) like PF-Net but with a new decoder. It includes a generator composed of a Multi-Resolution Encoder (MRE) and a Mesh Deformation Decoder (MDD), and a discriminator network. The encoder network is a point-based network that analyzes three resolutions of the partial point cloud to create global features. With the global features, the decoder progressively deforms a template shape into the desired 3D tooth. Unlike PF-Net, the decoder predicts a 3D mesh instead of a 3D point cloud. The cascaded mesh deformation network uses deformation blocks and two graph-based unpoolings. Each deformation block takes an input mesh and predicts the new coordinates of each vertex using the global features. The graph unpooling layer increases the number of vertices while maintaining the initial connectivity. Starting from a low-resolution template, the model learns to gradually deform and add details to the mesh in a coarse-to-fine strategy. To control the deformation and ensure the smoothness of the output, a Laplacian regularization is used in addition to Chamfer Distance (CD)

in the learning process.

5.4.1 Initial surface

The model always deforms a tooth-specific template to do patient-specific completions. The template is a downsampled mesh with roughly 90 vertices and 160 edges that has been normalized and registered with the input point cloud. The oriented bounding boxes of neighboring teeth are used to register and scale the template as shown in Figure 5.3. More specifically, a manual teeth segmentation performed by a technician allows us to create bounding boxes around each tooth. A box is then defined to position and scale the template points. The box has approximately the same width and depth as that of the sliced tooth. Its height is defined by the partial tooth boundary and the adjacent teeth cusps. This box-based registration is only an initial positioning and scaling that is refined by the network. In practice, automatic teeth segmentation (Alshegri et al., 2022; Lian et al., 2020) can provide the labels to define the bounding boxes.

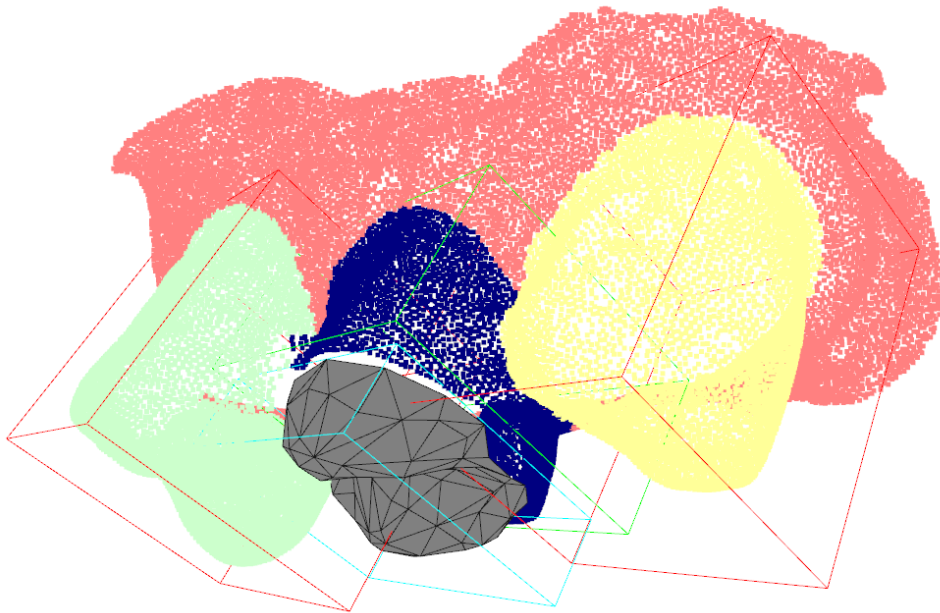


Figure 5.3 The template mesh (in gray) is positioned and scaled to fit inside a bounding box defined by the segmented input.

5.4.2 Mesh deformation block

Each mesh deformation block contains a linear layer, a 1D convolution layer and a graph convolution layer (see Figure 5.4). Graph convolutions in the deformations allow sharing of information between neighboring points, which is useful to predict appropriate displacements. In (Wang et al., 2018), the authors used 14 graph convolution layers in the deformation blocks to deform an ellipsoid into an unknown shape. However, we decided to use a single layer because the model deforms a template that already has general characteristics corresponding to those of the missing tooth. With an input mesh and some learned features, each deformation block learns displacements for each block’s input vertices. The first deformation block takes the template with M_{Low} vertices and predicts the $M_{Low} \times 3$ new coordinates using the F_{Low} features. The first intermediate mesh is then up sampled to have M_{Mid} vertices. The second deformation block takes the new upsampled mesh and predicts the $M_{Mid} \times 3$ new coordinates with the F_{Mid} features. Finally, the mesh is up sampled again, and the last deformation block predicts the high-resolution mesh with $M_{High} \times 3$ coordinates using the F_{High} features. With this coarse-to-fine deformation, the network can apply small deformations at the beginning that affect many high-resolutions vertices at once. The network thus learns to move the template vertices to the most representative locations and add details in further deformations.

5.4.3 Graph unpooling layer

The goal of the graph unpooling layers is to add new vertices in the mesh while preserving the initial connectivity. We followed the edge-based graph unpooling of Pixel2Mesh (Wang et al., 2018) to divide each triangle into 4 smaller triangles. This allows us to use two points to add new vertices instead of one like in PF-Net. This way, points are added at optimal positions and the upsampling can be very efficient since there are more edges than points. Furthermore, the edge-based unpooling maintains regular vertex degrees.

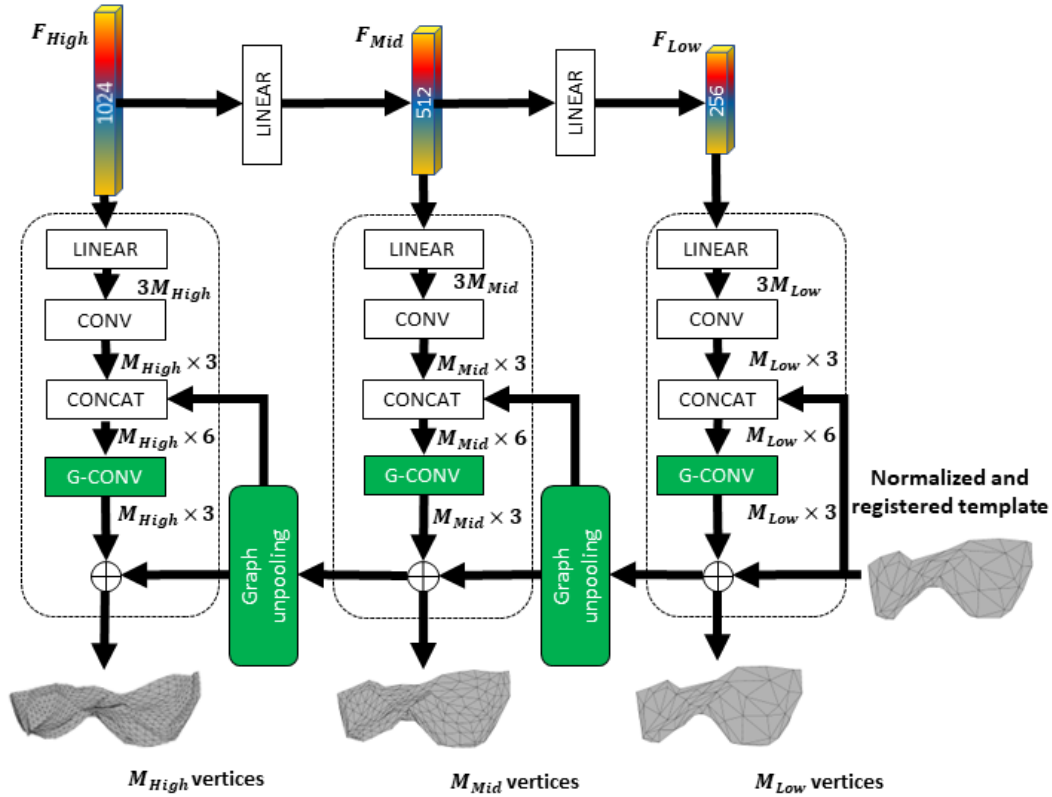


Figure 5.4 Details of the Mesh Deformation Decoder. MDD creates an output mesh along with 2 intermediate results. The graph unpooling layers preserve the connectivity during the coarse-to-fine generation. The plus circles represent add operations. Layers operating on points are shown in white; those operating on graphs are shown in green.

5.4.4 Discriminator's input

Originally in PF-Net, the discriminator receives as input the real shapes and the generated ones. It classifies them as real or fake regardless of any surrounding context. That formulation is not ideal here because the model may be able to generate shapes that look realistic by themselves but are not compatible with the full input geometry. Since the input contains many more points than the output, we propose to concatenate the sliced tooth from the input point cloud and the generated shape, as shown in Figure 5.5. This way, the discriminator's input is a complete tooth, which should be smooth everywhere and not have a gap at the intersection of the known and missing regions. A binary feature is added to let the discriminator know which are the input's points versus the

predicted ones.

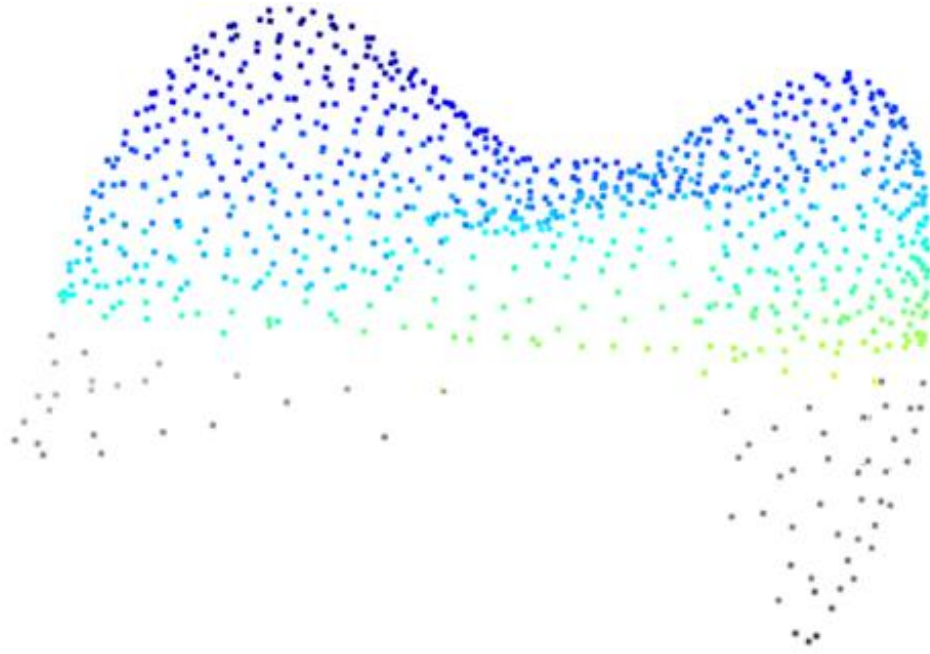


Figure 5.5 An example of the discriminator's input. The points in gray in the lower portion are sampled from the input using teeth segmentation and iterative farthest point sampling (IFPS).

5.4.5 Losses

We use three different losses to constrain the output shape and control the coarse-to-fine deformation. We use the Chamfer Distance (CD) loss (Fan et al., 2017) to constrain the vertex locations, a Laplacian regularization to maintain relative locations between neighboring vertices during the deformations, and an adversarial loss. All the losses except the adversarial loss are applied at each resolution.

The CD measures the mean squared distance between two point clouds S_1 and S_2 . The measurements are done between each point and its closest point in the other point set as described in equation (5.1):

$$d_{CD}(S_1, S_2) = \frac{1}{S_1} \sum_{x \in S_1} \min_{y \in S_2} \|x - y\|_2^2 + \frac{1}{S_2} \sum_{y \in S_2} \min_{x \in S_1} \|y - x\|_2^2, \quad (5.1)$$

where one term measures the distances from the prediction to the ground truth and the other measures the distances in the opposite direction. The completion loss uses the CD measurement

with each output resolution and specific weights as described in equation (5.2):

$$l_{comp} = d_{CD}(Y_{High}, Y_{GT}) + \alpha \cdot d_{CD}(Y_{Med}, Y'_{GT}) + \beta \cdot d_{CD}(Y_{Low}, Y''_{GT}), \quad (5.2)$$

where α and β are weights and Y_{GT} , Y'_{GT} , Y''_{GT} are the different resolutions of the ground truth point clouds obtained by iterative farthest point sampling (IFPS). l_{comp} is efficient to predict reasonably good vertex coordinates; however, it does not capture the smoothness and visual appearance of the shapes.

To preserve the template's smoothness, we propose a Laplacian regularization. This term prevents vertices from moving too freely, which can cause intersecting triangles and degrade the template characteristics. It encourages neighboring vertices to have similar movements. The Laplacian loss compares the relative positions of one point to its neighbors before and after the deformation. Each relative position is calculated using equation (5.3):

$$\delta_p = p - \sum_{k \in N(p)} \frac{1}{\|N(p)\|} k, \quad (5.3)$$

where p is a vertex in the predicted mesh and $N(p)$ are the neighboring vertices of p . The Laplacian loss is defined in equation (5.4):

$$l_{Lap} = \sum_p \|\delta'_p - \delta_p\|_2^2, \quad (5.4)$$

where δ'_p and δ_p are the relative positions before and after the deformation.

From the set of partial input point clouds X , the generator predicts shapes that should look like the missing point clouds in Y . The final predicted shapes and the ground truth are given to the discriminator network. The discriminator's goal is to classify its input shapes as real or fake. Both the generator and the discriminator are trained together with a GAN-based loss function (Goodfellow et al., 2014). The adversarial loss is formulated in equation (5.5):

$$L_{adv} = \sum_{1 \leq i \leq S} \log(D(y_i)) + \sum_{1 \leq j \leq S} \log(1 - D(G(x_j))), \quad (5.5)$$

where a partial input is represented by $x_i \in X$, a ground truth point cloud is represented by $y_i \in Y$ and S is the size of the dataset.

The final loss function is described by equation (5.6).

$$L = \lambda_{comp} l_{comp} + \lambda_{lap} l_{Lap} + \lambda_{adv} l_{adv} \quad (5.6)$$

where λ_{comp} , λ_{lap} and λ_{adv} are weights set to 0.95, 0.03 and 0.05.

5.5 Experiments

5.5.1 Dataset and preprocessing

We used different datasets for each tooth position, each one associated with a tooth-specific template. In total, we used four datasets to generate central incisors, canines, first premolars and first molars. To create each dataset, the full dental arches were first registered to a reference dental arch using oriented bounding boxes. The top part of the tooth to generate was removed using a planar slice. Afterwards, the 3D data was filtered to remove any structures that were not close to the sliced tooth. This step created a partial input that contains only partial tooth and its immediate neighborhood. Each input contained the sliced tooth, two adjacent teeth, three antagonist teeth, and the surrounding gingiva (see Figure 5.6). Because of lateral symmetry, we only generated teeth from the left side of the mouth, which reduced variability in the data. The preprocessed meshes were then converted to point clouds to be analyzed by the encoder, as shown in the lower part of Figure 5.6.

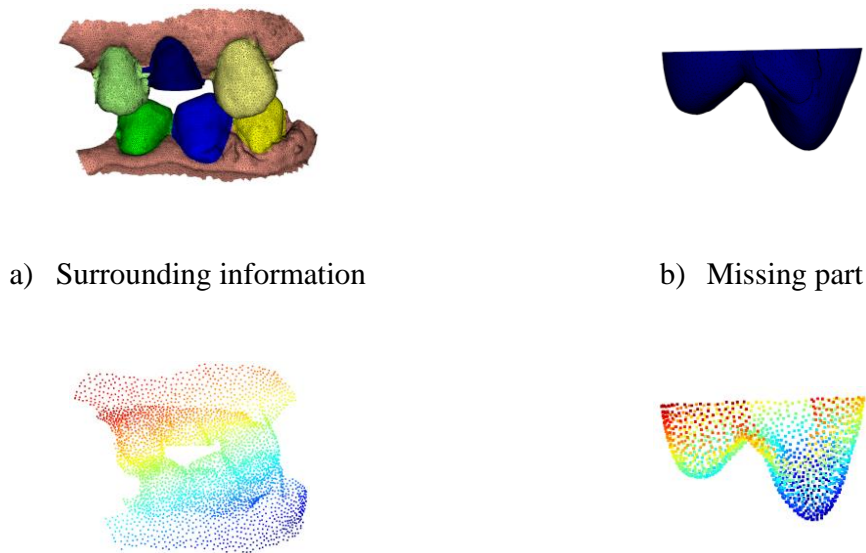


Figure 5.6 Illustration of an input-output pair. Half of the tooth to be completed is removed using a planar slice.

The input point clouds were first normalized using z-score normalization along each axis. Each of the four datasets contained roughly 40 teeth. The training, validation and test sets contained 28 teeth, 6 and 6 teeth. The training sets were augmented using small random translations, rotations, and scaling.

5.5.2 Implementation details

The input point clouds were downsampled with IFPS to 4000, 2000, and 1000 points. The intermediate resolutions M_{Low} and M_{Mid} were set to 93 and 351 vertices and the final output resolution M_{High} to 1365 vertices. 100 input points were added to the 1365 points of the missing regions to get complete teeth to provide to the discriminator. The encoder and the discriminator were the same as in PF-Net (Huang et al., 2020). The MRE predicted a 1920 feature vector. The displacements learned by the MDD were scaled down using a linear activation function with a slope of 0.01. This scaling ensured that, during training, the first deformations did not introduce distortions or artefacts that would be difficult to remove in later iterations.

We implemented our network on PyTorch. The generator and the discriminator were trained alternately using the ADAM optimizer with a learning rate of 0.001 and a batch size of 6. Batch

normalization and RELU activation units were used in the encoder and the discriminator but only RELU was used in the MDD. The model was trained for 40 epochs, which took 15.25 hours using an NVIDIA GeForce RTX 3080 graphic card. During testing, our model took 4.98ms to infer a mesh with 1365 vertices.

5.5.3 Comparison with state of the art

We evaluated our method against two baselines. The first one was the original architecture of PF-Net which is used to predict the vertex coordinates only. The second one was a slightly modified PF-Net called PF-Net*. This second baseline consists of a PF-Net with a point-based template used in the architecture to predict the low-resolution points similarly to MC-Net. The template is the same one as in MC-Net but without edges between the points. PF-Net* then learns low resolution relative coordinates instead of direct coordinates as in the original article (Huang et al., 2020). That modification makes sure the model starts its coarse-to-fine generation with a smooth and uniform point cloud. The model still uses the concatenation and addition mechanism to add new points and to deform the point template. PF-Net* is interesting experimentally because it allows us to evaluate to what extent the model can keep the smoothness and uniformity of the template when working only with point clouds. All three architectures were trained using the same datasets and training methodology. To evaluate the methods, we used the CD between the predicted shapes and the ground truth shapes. For a fair comparison, all the methods have a final output resolution of 1365 points. However, their intermediate resolutions differ due to their respective upsampling mechanisms. The CD is a bidirectional measurement, including a prediction to ground truth error $Pred \rightarrow GT$ and a ground truth to prediction error $GT \rightarrow Pred$. $Pred \rightarrow GT$ indicates how different the prediction is from the ground truth, while $GT \rightarrow Pred$ indicates how much of the ground truth surface is covered by the predicted one.

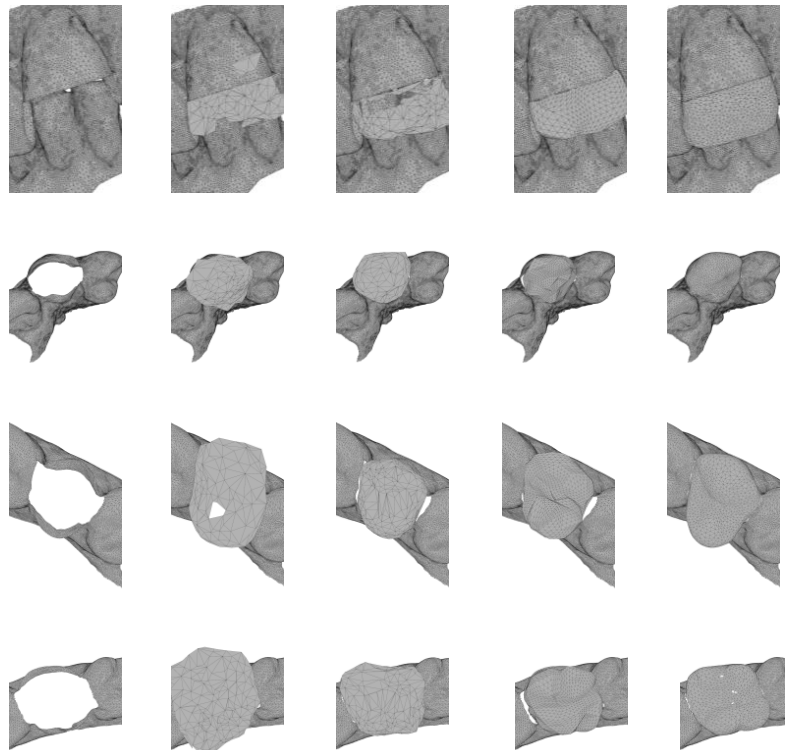
5.5.3.1 Quantitative evaluations

Table 5.1 shows the results for the three architectures with different tooth position datasets. MC-Net’s hyperparameters were determined by experiments on premolars only. Our method outperforms the two other networks for both the $GT \rightarrow Pred$ and $Pred \rightarrow GT$ errors, in the case of premolars. Our network can predict shapes that are closer to the ground truth and that cover more of the ground truth regions. The premolars generated by MC-Net have an average CD

measure of $0.85 * 10^{-2}$. By de-normalizing each CD measurement and averaging them all, the mean CD corresponds to a physical distance of 0.22 millimeters. The results of PF-Net* are consistently better than those of PF-Net. For incisors, canines and molars, PF-Net* yields the lowest errors of all three methods. However, lower CD does not necessarily lead to a better mesh model. To demonstrate this, qualitative results are shown in the next section.

Table 5.1 Tooth completion results after 6-fold cross-validation. The reported metrics are [Pred \rightarrow GT / GT \rightarrow Pred]. Resulting normalized CD values are scaled by 100.

| Category | PF-Net | PF-Net* | MC-Net |
|-------------------|-----------|------------------|------------------|
| Incisor #11, #21 | 4.92/3.38 | 2.41/1.83 | 3.08/2.71 |
| Canine #13, #23 | 2.99/1.94 | 1.59/0.99 | 1.66/ 1.32 |
| Premolar #14, #24 | 2.71/1.85 | 1.08/0.90 | 0.89/0.81 |
| Molar #16, #26 | 3.09/2.28 | 1.22/1.02 | 1.24/1.24 |
| Mean | 3.43/2.36 | 1.58/1.19 | 1.71/1.52 |



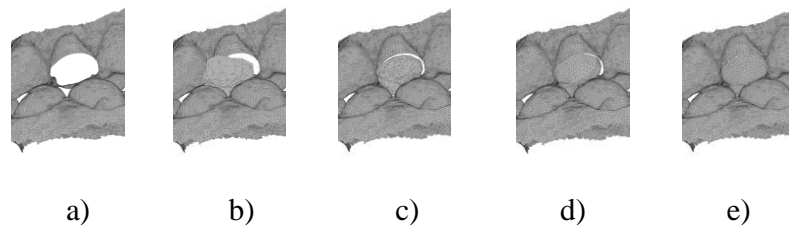


Figure 5.7 Examples of mesh completions by the different architectures. **a)** partial input meshes used to create points clouds; **b)** generated teeth with PF-Net; **c)** generated teeth with PF-Net*; **d)** generated teeth with MC-Net; **e)** ground truth meshes. In all cases except **d)**, ball pivoting reconstruction was used to create meshes from point clouds. Holes in the ground truth and the prediction meshes are due to reconstruction artefacts.

5.5.3.2 Qualitative evaluations

As mentioned above, the CD metric does not tell us about the smoothness and visual appearance of the predicted shapes. Therefore, a qualitative evaluation is necessary to judge the correctness of the tooth completions. Figure 5.7 provides several examples of visual results from the three models, as surface meshes. The figure shows the partial input, the three predictions and the ground truth in each case. To convert the point cloud results into meshes, we used a standard algorithm (ball pivoting reconstruction) (Bernardini et al., 1999).

As we can see in Figure 5.7, the mesh reconstruction from PF-Net and PF-Net*'s point clouds is non-trivial. With PF-Net, the reconstructed meshes are sometimes wrongly positioned, which significantly affects the quality of the predictions. In contrast, with PF-Net*, the meshes are better positioned and contain more details. However, with both architectures, the final meshes contain spikes, intersecting triangles, and holes. These artefacts result from the large number of degrees of freedom when no regularization is used. The resulting noise in the point clouds degrades the quality of the reconstruction results.

Unlike the baselines, our model predicts meshes directly. Its components allow us to carefully deform the template into a smooth continuous surface. Indeed, as revealed in Figure 5.7, MC-Net's predictions are smoother, more uniform and detailed than the baseline shapes from PF-Net and PF-Net*.

5.5.4 Ablation study

We performed an ablation study to evaluate the effectiveness of different components of MC-Net.

The ablation study is done using the premolar dataset. The validation set was used to report results. We introduced a new metric called ICP-CD. It measures the CD after an Iterative Closest Point (ICP) registration of the prediction with respect to the ground truth. This metric evaluates the shape without any positional bias.

5.5.4.1 Effectiveness of the tooth template

Using a half sphere template instead of a tooth template implies learning a more challenging deformation. Indeed, the generated shapes starting from a half sphere are smoother and contain fewer curvatures. As seen in Figure 5.8, the model deforms the half sphere into teeth that lack grooves that are characteristic of normal teeth. This is problematic because the designed shape should have similar grooves as the other patient’s teeth to be realistic. To get grooves using a half sphere template, the deformation blocks would need to be more complex, and the model would have to learn all the characteristics of dental shapes from scratch. As shown in Table 5.2 and Figure 5.8, our full model can generate correct teeth with simple deformation blocks because the tooth template provides important knowledge of the intended shape.

5.5.4.2 Effectiveness of the Laplacian regularization

As shown in Figure 5.8, the Laplacian loss is important for the smoothness of the predictions. Without it, the generated shapes are much less realistic. The noise introduced causes overlaps with the surrounding teeth, which is not desirable. The noise also increases the CD as shown in Table 5.2. The Laplacian regularization is hence a key component of MC-Net, ensuring that the predictions remain smooth.

5.5.4.3 Effectiveness of graph convolutions

As shown by the qualitative and quantitative results, the single graph convolution layer in each mesh deformation block contributes very little to the results. It changes the shape slightly, especially on the occlusion surface. The predictions using graph convolutions have smoother grooves compared to those without them. Indeed, with only one layer, a point receives only information from its direct neighbors. This helps the model to predict better displacements locally to avoid undesired surface folds. By adding more graph convolution layers, each point would receive information from a larger neighborhood, but this would increase the model complexity.

5.5.4.4 Effectiveness of the new discriminator’s input

According to Table 5.2, the points describing the sliced tooth added to the discriminator’s input help to reduce the CD. This result is not surprising because the predicted shape is supposed to form a realistic whole tooth when joined with the partial input. Therefore, the positions of the boundary points are an important aspect of the prediction. Using the new input formulation, the adversarial loss can provide a more meaningful gradient. The cost function implicitly encourages the predicted boundary points to be close to the partial input’s boundary, i.e. the base of the tooth. As shown in Figure 5.8, the predictions have a better fit with the partial input when the new input formulation is used.

Table 5.2 Ablation study to evaluate the effect of different components of our MC-Net model. Each row reports the average error metrics when removing the corresponding component. Resulting normalized CD values are scaled by 100.

| Experiment | CD | ICP-CD |
|------------------------------------|--------------|--------------|
| - Tooth template | 1.119 | 0.562 |
| - Laplacian | 0.971 | 0.421 |
| - Graph convolution | 0.659 | 0.302 |
| - Discriminator input modification | 0.670 | 0.307 |
| Full model | 0.659 | 0.291 |

5.6 Discussion

The improvements of PF-Net* over PF-Net as shown by the last row in Figure 5.7 are mainly due to the positioning of the point cloud template. The initial positioning helps the model understand precisely where the predicted mesh should be located with respect to the partial input. Thus, the template positioning allows both PF-Net* and MC-Net to reduce the CD error metric. PF-Net* produces a point cloud with fewer constraints than MC-Net. This greater freedom allows the model to reduce the CD, but it does not improve the visual appearance of the meshes, as revealed in our qualitative evaluation (see Figure 5.7).

With PF-Net*, even if the provided point template is smooth and uniform, the decoder still introduces noise and artefacts that affect the CD and the smoothness. The template characteristics

are lost and the model is not able to preserve its smoothness. This problem is due to the concatenation and addition mechanism that does not exploit connectivity information when adding new points. Therefore, new points in PF-Net* and PF-Net have a large degree of freedom, which creates noise. Since we want to complete a surface of a known shape category, the flexibility of the point cloud approach does not provide any advantage; instead, it creates a smoothness problem.

By contrast, MC-Net's predictions are very smooth. They benefit from the edge-based graph unpooling and the Laplacian regularization. The unpooling preserves the initial connectivity and maintains the smoothness during the upsampling. Meanwhile, the Laplacian regularization provides a meaningful gradient to generate smooth meshes with relevant details as demonstrated in the ablation study. That regularization reduces the noise and the CD at the same time. Its effect is more apparent in the qualitative results as shown in Figure 5.7. By finetuning the Laplacian regularization weight for each category of teeth, as we did for the premolars, MC-Net could provide optimal results for any tooth position.

We made several design choices with the aim of having more control over the mesh generation process and producing appropriate shapes for tooth completion. However, our method may produce meshes with incorrect boundaries. The boundary seems to be a difficult region for the model to understand. This is unexpected because the partial input contains a boundary that tells the network where the generated mesh must begin. However, our results indicate that that constraint is not explicit enough yet. This may be due to the standard CD formulation, in which all the points are equally important. A weighted CD giving more importance to the boundary points could help with this problem. A better template initialization could also be explored. Furthermore, a transformer-based encoder could help to capture dependencies between the boundary points at a broader scale.

Other improvements to our current approach could be considered. Among them, a mesh encoder could potentially provide better features to guide the shape generation than the currently used point encoder. Additionally, a functionality-aware loss function could help the predicted tooth to have better contact points with adjacent teeth. Finally, adding more graph convolution layers would allow information sharing between more distant points in the mesh generation.

MC-Net addresses the task of dental mesh completion, which is closely related but not identical to the task of dental crown generation. One difference is that tooth preparations have quite variable shapes, and the crown must fit "around" the base as opposed to extending a previously sliced tooth.

Another difference lies in the crown itself, which must be a closed shape rather than the completion of one. On the other hand, real teeth (that our model learns to complete) have greater variability than crowns; the latter are typically adapted from templates using standard techniques for sizing and contact points. Therefore, the two problems are of comparable complexity, and an appropriate adaptation of our methodology could lead to dental crowns with better personalization than existing techniques.

5.7 Conclusion

We have presented an approach to perform mesh completion from a partial input point cloud. Our model, called MC-Net, controls the mesh completion process using different mesh-related losses, graph unpooling, information propagation along edges and a mesh template. Our results show that MC-Net can complete single surfaces accurately compared to the state-of-the-art PF-Net. The predicted surfaces are always smooth and are sized accurately. Thus, we believe that this work will be beneficial to subsequent research on tooth completion and to the related problem of dental crown generation.

5.8 Acknowledgments

This work was funded by Kerenor Dental Studio, Intellident Dentaire Inc., iMD Research Inc., the Natural Science and Engineering Research Council of Canada (NSERC), and MEDTEQ. We thank Compute Canada for providing computational resources for our experimentation. We also acknowledge the help provided by JACOB Inc. and Object Research Systems Inc. We thank Philippe Debanné for revising this manuscript. This work was approved by the research ethics committee of Polytechnique Montréal. It has not been submitted elsewhere for publication.

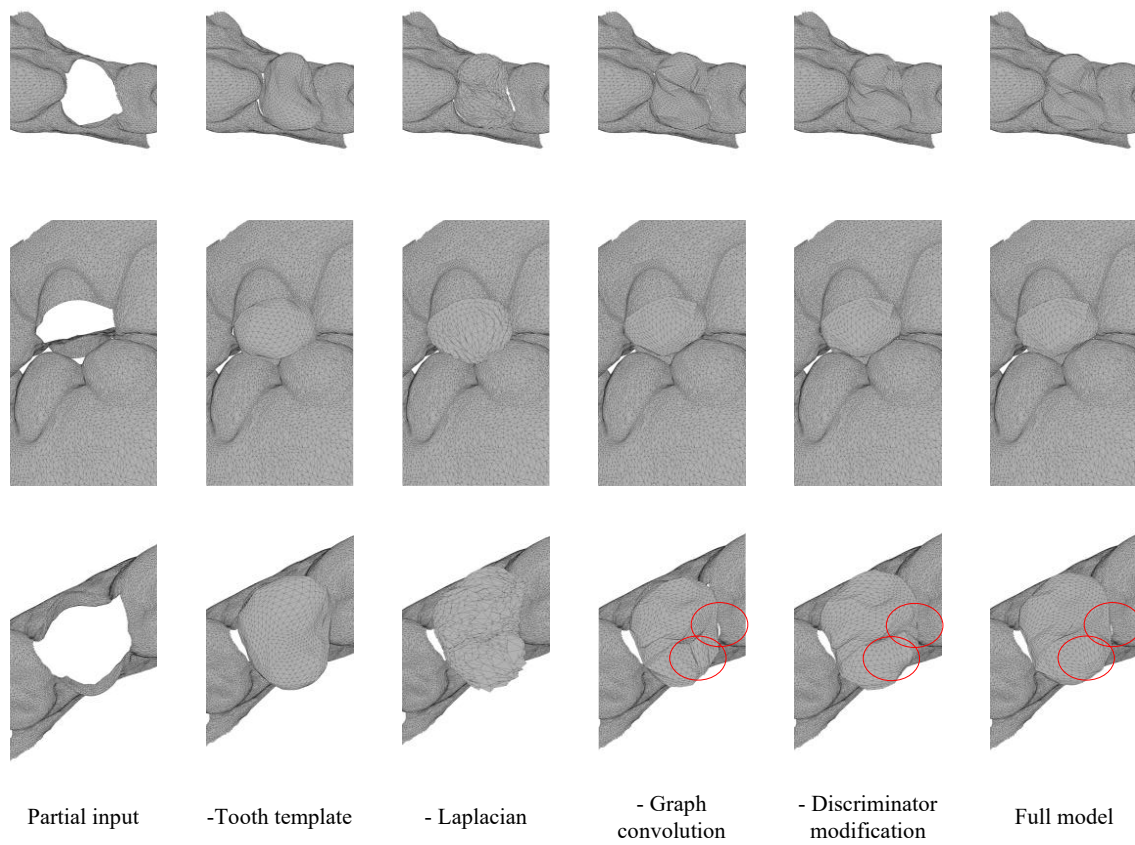


Figure 5.8 Qualitative results of the ablation study. The images in each column reflect the contribution of each component of MC-Net by showing the resulting mesh without that component. The red circles overlaid on the images at bottom right show subtle difference

CHAPITRE 6 EXPÉRIMENTATION COMPLÉMENTAIRE

Dans cette section, des résultats complémentaires obtenus à la suite de la réalisation de l'article 1 seront présentés. Cette section permet de mieux illustrer les limitations de PF-Net avec 768 points en sortie. Ces résultats ont servi d'inspiration pour l'article 2.

6.1 Analyse complémentaire des résultats l'article 1

Dans l'article 1, PF-Net génère plusieurs formes différentes. Afin de comprendre la complexité des dents de l'ensemble de test, un graphique mettant en relation l'erreur de test et la quantité de données dans l'ensemble d'entraînement de chaque position de dents peut être fait. Ce graphique est obtenu avec les mêmes résultats de test que ceux présentés dans l'article 1.

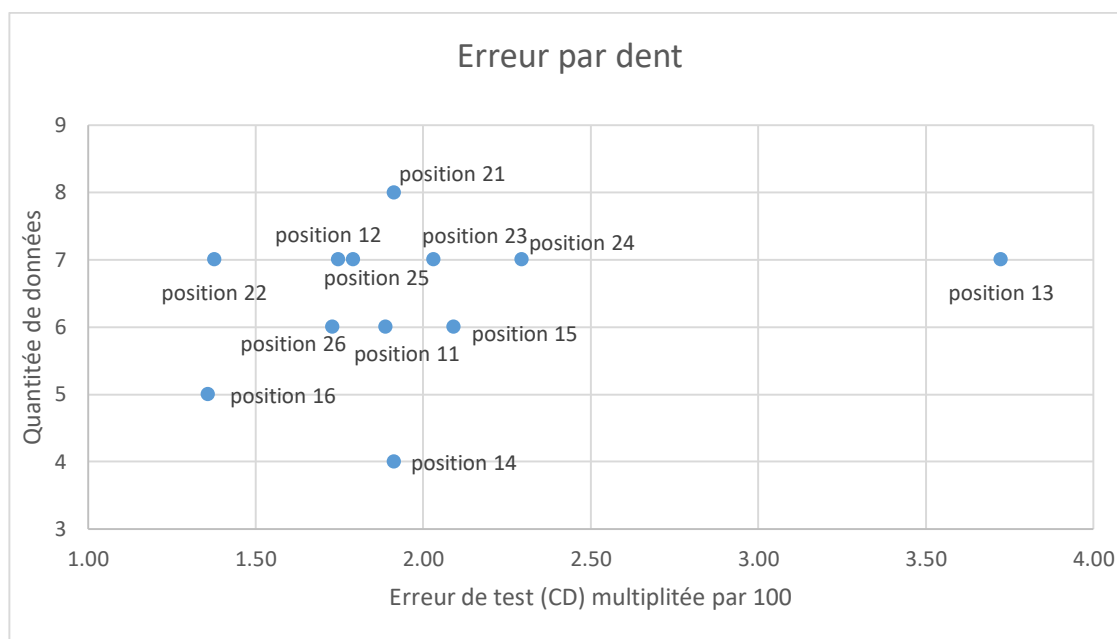


Figure 6.1 Relation entre la quantité de données d'entraînement et les erreurs de l'ensemble de test par dent de l'article 1

La Figure 6.1 permet d'illustrer la relation entre la quantité de données d'entraînement et l'erreur mesurée (CD). Les différentes dents possèdent chacune entre 4 et 8 échantillons dans l'ensemble d'entraînement. Théoriquement, leurs erreurs moyennes respectives devraient être similaires. Cependant, on remarque que l'erreur pour la position 13 est significativement plus grande. Cette position correspond à aux canines gauches. Ce résultat démontre la complexité de la forme des

canines. Cette complexité est due à la forme spécifique des canines qui se distingue des autres dents. Par opposition, les canines droites sont associées à la position 23 et leur erreur est similaire à celles des autres positions. Néanmoins, en calculant les erreurs en regroupant les dents sous les catégories d'incisives, de canines et de molaires, l'erreur sur les canines est plus grande que sur les autres catégories. La Figure 6.2 démontre que les canines sont moins représentées que les autres catégories. En considérant la symétrie, les canines sont représentées par une seule position contrairement aux incisives et aux molaires. Ces deux dernières peuvent bénéficier de caractéristiques communes entre leurs différentes positions, ce qui facilite leur génération.

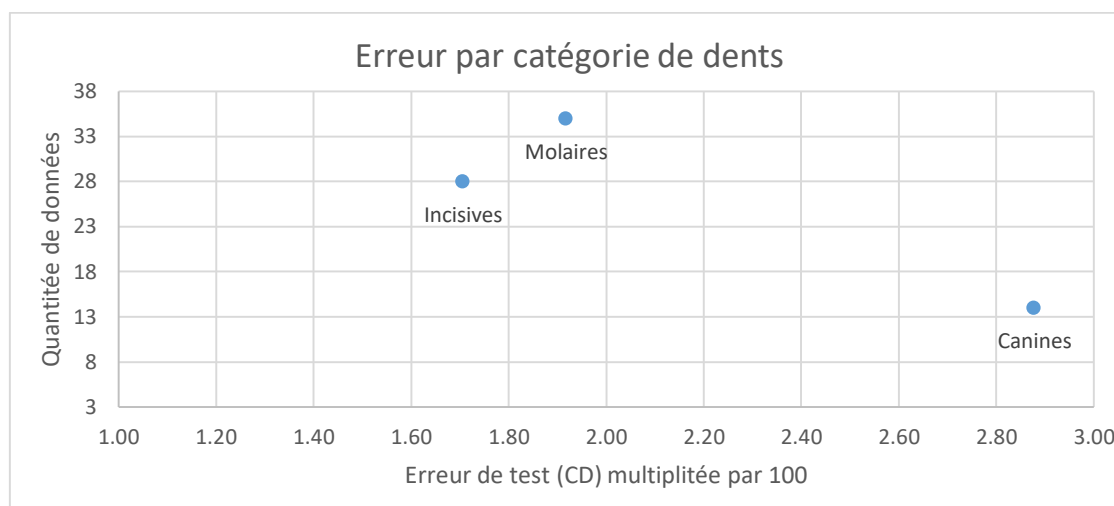


Figure 6.2 Relation entre la quantité de données dans l'ensemble d'entraînement et les erreurs de l'ensemble de test par catégorie de dents de l'article 1

L'utilisation de la distance de chanfrein permet de comparer deux formes mais elle ne permet pas d'indiquer les régions plus difficiles à générer. Pour mieux comprendre les sources d'erreurs de l'article 1, les prédictions peuvent être superposées aux formes de références. Dans la Figure 6.3, les prédictions les plus difficiles sont colorées selon leur plus petite distance au carré par rapport aux références montrées en noir. L'échelle de couleur va de vert vers rouge, rouge représentant une plus grande distance. Les régions en rouge représentent les régions éloignées des références. Ces régions correspondent parfois aux cuspidés ou aux rainures. La couleur rouge représente une erreur par point d'un millimètre ou plus tandis que la couleur verte représente une erreur presque nulle. De plus, les points prédits ont des normales pointant vers le plus proche voisin dans la référence.

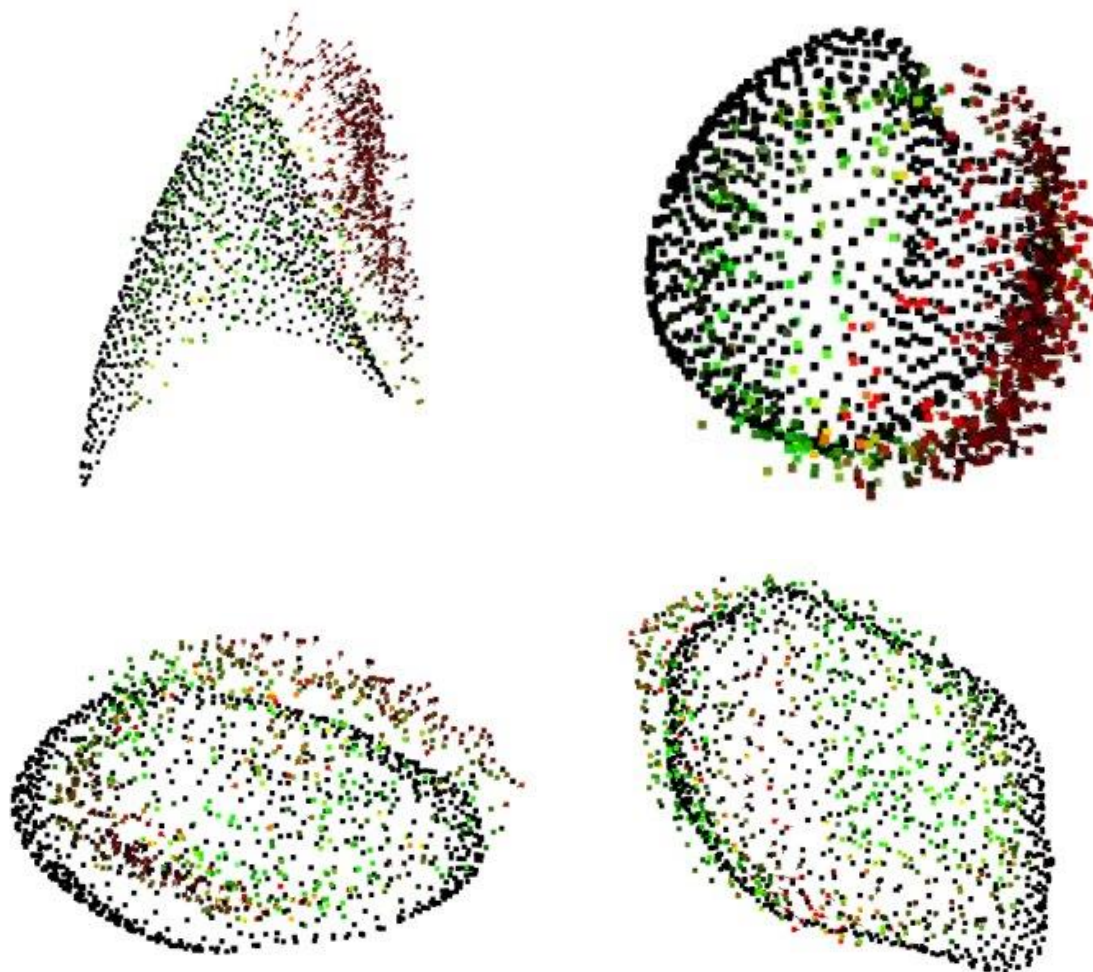


Figure 6.3 Erreurs des prédictions (en couleur) par rapport aux références (en noire).

Grâce à la Figure 6.3, on remarque que les prédictions possèdent des positionnements différents de ceux des formes de références. Ces positionnements augmentent sévèrement les distances de Chanfrein mesurées.

Positionner une dent sur une partie de gencive est une tâche difficile. Un recalage de la prédiction sur la référence a été appliqué afin de mesurer la CD sans être biaisée par le positionnement. Les résultats avec et sans recalage sont illustrés dans le tableau 6.1. En appliquant le recalage, la CD est diminuée presque de moitié. Cette différence démontre que la position buccal-lingual est une source d'erreur importante dans les prédictions. Cette source d'erreur est due à la formulation du problème qui n'inclue pas de contour guidant la génération.

Tableau 6.1 Résultats avec et sans recalage pour une génération de 768 points

| Méthode | CD de test |
|----------------|------------------------------------|
| sans ICP | $2.03 * 10^{-2}$ |
| avec ICP | $1.04 * 10^{-2}$ |

En entraînant le modèle avec plusieurs catégories de formes, l'ensemble d'entraînement comporte une très grande variabilité. Les dents qui s'y retrouvent peuvent être des incisives, des canines, des prémolaires ou même des molaires. En entraînant un modèle avec toutes ces dents, celui-ci doit comprendre le type de dents qui doit être générer en plus du positionnement de la prédiction. Ce besoin entraîne une baisse de détails dans les prédictions spécialement au niveau des rainures et cuspides.

6.2 Résultats complémentaires de l'ajout d'un gabarit de points dans le décodeur de PF-Net

Dans cette section, des expériences utilisant PF-Net* avec des hyperparamètres différents seront présentées et discutées. Pour ces expériences, la base de données comprenant uniquement des prémolaires supérieures a été utilisée. Les résolutions en sorties ont été fixées à 128, 256 et 768 points. De plus, le post traitement des nuages de points présentés à l'article 1 a été utilisé pour éliminer les anomalies. Tel qu'expliqué dans l'article 2, PF-Net* utilise un gabarit de points illustré à la Figure 6.4 qui est intégré dans le décodeur comme à la Figure 6.5.

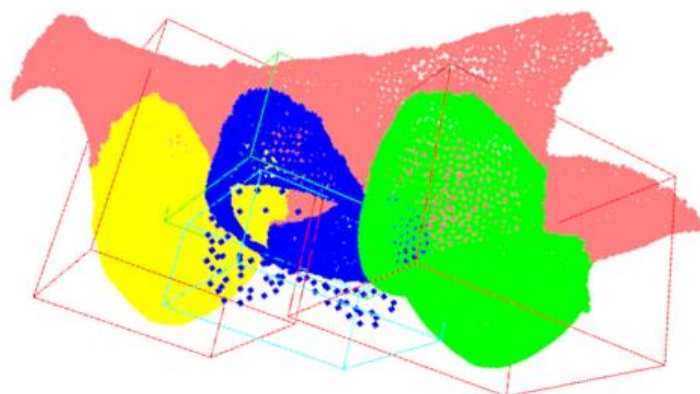


Figure 6.4 Positionnement d'un gabarit pour la dent 14.

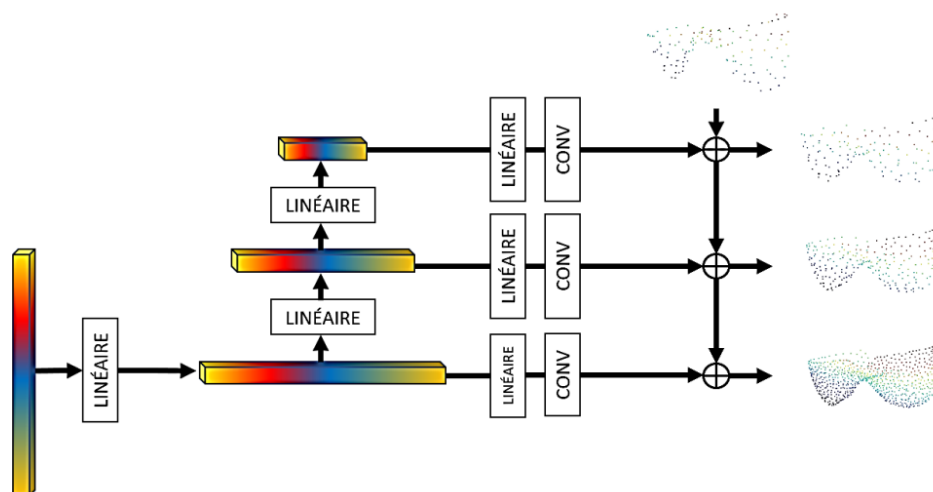


Figure 6.5 PPD avec gabarit.

Le Tableau 6.2 présente les résultats obtenus avec PF-Net et PF-Net* et ces hyperparamètres. On peut remarquer que ces deux architectures obtiennent une distance de chanfrein de 0.0087. Ces résultats permettent de conclure que le gabarit n'amène pas d'amélioration significative pour ces paramètres. Ce résultat démontre encore le manque de contrôle au niveau du lissage du décodeur de PF-Net. Avec une basse résolution augmentée et une haute résolution diminuée, l'effet du gabarit de points est moins important, contrairement aux résultats de l'article 2.

Tableau 6.2 Comparaison des méthodes pour une génération de prémolaires de 768 points

| Méthodes | Moyenne $\pm \sigma$ (résultats multipliés par 100) | |
|----------|---|-----------------|
| | CD | ICP-CD |
| PF-Net | 0.87 \pm 0.41 | 0.47 \pm 0.24 |
| PF-Net* | 0.87 \pm 0.43 | 0.45 \pm 0.23 |

6.3 Ajout d'une fonction de coût de fonctionnalité

La génération conditionnelle ne rend pas les contraintes dentaires explicites. Le générateur pourrait produire des couronnes plausibles qui seraient inutilisables. Afin de formuler les contraintes dentaires, l'espace entre une dent générée et les dents opposées doit être mesuré. Pour ce faire, nous proposons de mesurer des distances selon l'axe z entre les points d'une dent et les points des dents opposées. Ces distances sont nommées distances d'occlusion. Une dent générée idéale devrait posséder quelques points de contacts avec les dents opposées afin de bien broyer les aliments. Ces points de contacts n'ont pas besoin d'être exactement au même endroit que ceux des dents de références. De plus, la dent ne devrait pas chevaucher les dents voisines pour permettre une bonne fermeture de la mâchoire. L'ensemble des distances d'occlusion mesurées pour une dent doit donc suivre une distribution spécifique pour répondre à ces critères de fonctionnalité.

Avec ces mesures de distances, des histogrammes de distances d'occlusion peuvent être créés pour chaque dent réelle et artificielle. Nous proposons d'utiliser une fonction de coût afin de faire correspondre les distributions de ces histogrammes. La fonction développée s'inspire de celle de Hwang et al. (Hwang et al., 2018). Ce type de fonction est intéressant puisqu'elle permet plus de flexibilité lors de l'entraînement. En effet, les régions avec point de contact n'ont pas besoin d'être

exactement au même endroit pour assurer une bonne fonctionnalité et obtenir un même histogramme.

Les distances d'occlusion sont calculées entre une dent artificielle \hat{y} créée par un générateur G ou une dent réelle y et les dents opposées x_{occ} . La Figure 6.6 illustre une dent réelle en couleur et une partie des dents opposées en noir. Chaque dent réelle ou artificielle est représentée par un nuage de points de N points appartenant à l'ensemble \mathbb{R}^3 où les caractéristiques correspondent aux coordonnées x-y-z des points. Puisque les modèles dentaires sont préalablement recallés dans le plan x-y, nous assumons que les distances d'occlusion se mesurent selon l'axe z. Pour chaque point p d'une dent du haut, une distance d'occlusion est mesurée avec le plus proche voisin q appartenant aux dents du bas x_{occ} après une projection dans le plan x-y comme suit:

$$d_{occ} = f(p, x_{occ}) = p_3 - q_3 \quad (6.1)$$

où

$$d_{xy} = \min_{q \in x_{occ}} \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2} \quad (6.2)$$

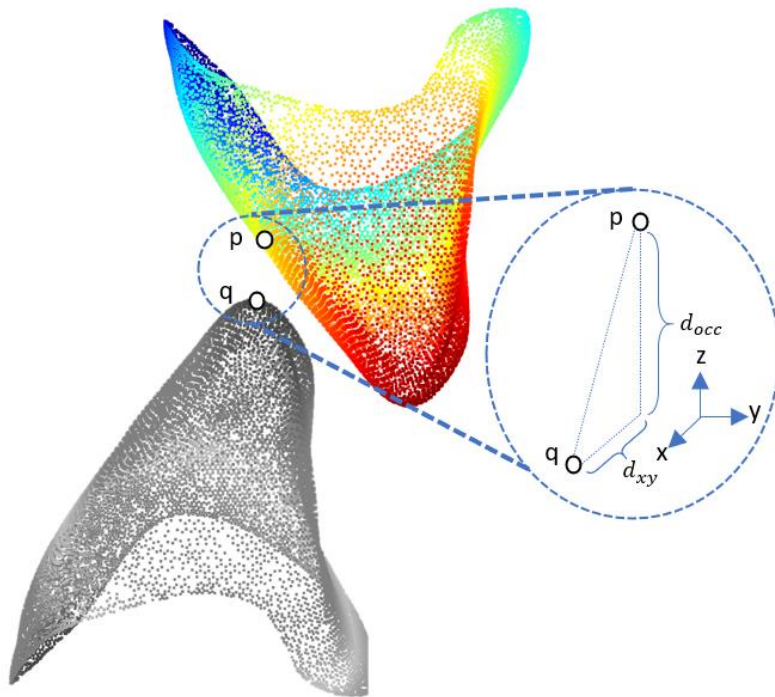


Figure 6.6 Mesure de distance d'occlusion avec nuage de points

Ces distances sont illustrées à la Figure 6.6. En utilisant cette mesure d'espace, un ensemble de distance d'occlusion noté D_y ou $D_{\hat{y}}$ compris dans \mathbb{R}^1 est créé pour chaque forme réelle et artificielle. Cet ensemble est un vecteur N par 1. Ces distances doivent être toujours être positive pour éviter tout chevauchement avec les dents opposées. Les ensembles sont utilisées pour construire des histogrammes (Wang et al., 2016) et faire correspondre les distributions de distances. La correspondance entre ceux-ci est évaluée avec la fonction de coût suivante:

$$L_H(G) = \mathbb{E}_{D_y, D_{\hat{y}}} \left[\sum_i \frac{(h_i(D_{\hat{y}}) - h_i(D_y))^2}{\max\{1, h_i(D_y)\}} \right], \quad (6.3)$$

où la bande i de l'histogramme est modélisée par une fonction linéaire par morceaux $h_i(d_{occ})$. Cette fonction calcule tous les scores des distances par rapport à une classe de l'histogramme comme suit :

$$h_i(d_{occ}) = \text{sum}\{\max\{0, 1 - |d_{occ} - c_i| \times l_i\}\}, \quad (6.4)$$

où c_i et l_i représentent le center et la largeur de la bande i . Seules les distances comprises dans la bande $(c_i - l_i, c_i + l_i)$ obtiennent des scores non-nuls. Lors de la rétropropagation d'erreur, ces distances reçoivent un gradient avec une pente de $\pm l_i$. La fonction de coût finale devient :

$$L = \lambda_{com} L_{com} + \lambda_{adv} L_{adv} + \lambda_H L_H \quad (6.5)$$

où λ_H est un poids de 0.01 pour balancer la nouvelle fonction de coût d'histogramme. Dans le tableau suivant, l'effet de l'ajout de la fonction de coût est présenté. On remarque que la fonction de coût permet d'améliorer les résultats. La complétion de plusieurs dents avec une base de données est très difficile. En ajoutant cette fonction, le modèle devient plus complexe et arrive à mieux modéliser les complétions.

Tableau 6.3 Effet de la fonction d'occlusion sur la base de données de l'article 1

| Méthode | Distance de Chanfrein (CD) |
|----------------------------------|------------------------------------|
| PF-Net | $2.27 * 10^{-2}$ |
| PF-Net avec fonction d'occlusion | $2.03 * 10^{-2}$ |

Cette fonction de coût a également été testée avec l'architecture PF-Net* entraîné à générer uniquement des prémolaires comme présenté à la section 6.2. L'ajout de cette fonction n'améliore pas significativement les prédictions, et les métriques restent sensiblement les mêmes comme montré au Tableau 6.4. Les erreurs sont encore principalement au niveau de la bordure. Étant donné que la fonction de coût utilise des mesures de distances selon un axe parallèle aux surfaces touchant la bordure, celle-ci n'est pas idéale pour aider la génération à cet endroit. Elle permet surtout de procurer de l'information utile aux cuspidés et creux, puisque les surfaces locales de ces régions sont en général perpendiculaires à l'axe utilisé.

Tableau 6.4 Effet de la fonction d'occlusion sur la base de données de prémolaires

| Méthodes | Moyenne \pm σ (résultats multipliés par 100) | |
|--|---|---------------------------------|
| | CD | ICP-CD |
| PF-Net* | 0.87\pm0.43 | 0.45 \pm 0.23 |
| PF-Net* avec fonction de coût d'occlusion | 0.91 \pm 0.53 | 0.39\pm0.36 |

En créant une base de données comportant une seule position, la variabilité dans les données est beaucoup réduite. Pour cette raison, l'utilisation de la fonction de coût avec une petite base de données spécialisée entraîne un sur-apprentissage plutôt qu'une généralisation lors de l'entraînement. Cet effet ne permet donc pas d'améliorer les performances du modèle, comme montré dans le Tableau 6.4.

6.4 Résultats complémentaires de l'ajout d'information conditionnelle au discriminateur

Dans cette section, l'effet de la nouvelle formulation de l'entrée du discriminateur sera présenté sans autre modification d'architecture. Le modèle comprenant cet ajout est intitulé PF-Net-D+. Comme présenté dans l'article 2, l'ajout d'information conditionnelle dans l'entrée du discriminateur permet d'obtenir des complétions avec un meilleur positionnement. Le Tableau 6.5 illustre les résultats obtenus avec des résolutions de sorties de 128, 256 et 768 points et le post traitement. On remarque que la distance de chanfrein est réduite d'environ deux écarts-type.

Tableau 6.5 Effet de la nouvelle formulation de l'entrée du discriminateur

| Méthodes | Moyenne $\pm \sigma$ (résultats multipliés par 100) | |
|-----------|---|--------------------------------|
| | CD | ICP-CD |
| PF-Net | 0.87 \pm 0.41 | 0.47 \pm 0.24 |
| PF-Net-D+ | 0,79\pm0,32 | 0,44\pm0,2 |

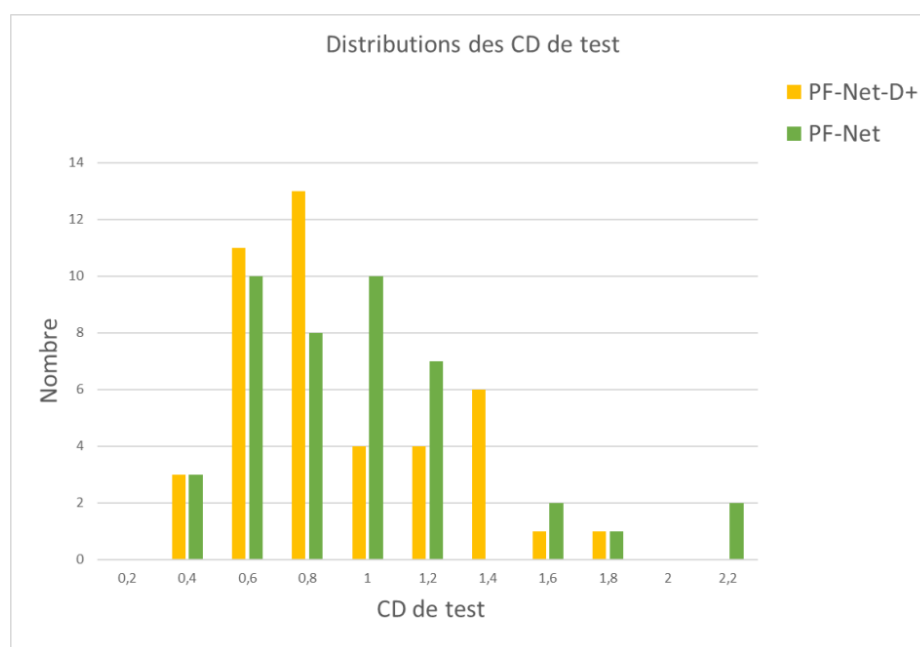


Figure 6.7 Histogramme illustrant le changement de distribution de CD de test

La Figure 6.7 montre le changement de distribution des CD de test en effectuant la modification. On peut tout d'abord remarquer que la distribution en jaune possède une moyenne plus basse que celle en vert. De plus, PF-Net a de la difficulté à générer les formes manquantes pour deux cas représentés en vert à la droite de l'histogramme. PF-Net-D+ obtient de meilleurs résultats sur ces mêmes données, ce qui permet de réduire la CD moyenne. Un de ces cas difficile est montré dans la prochaine image pour faire une évaluation qualitative.

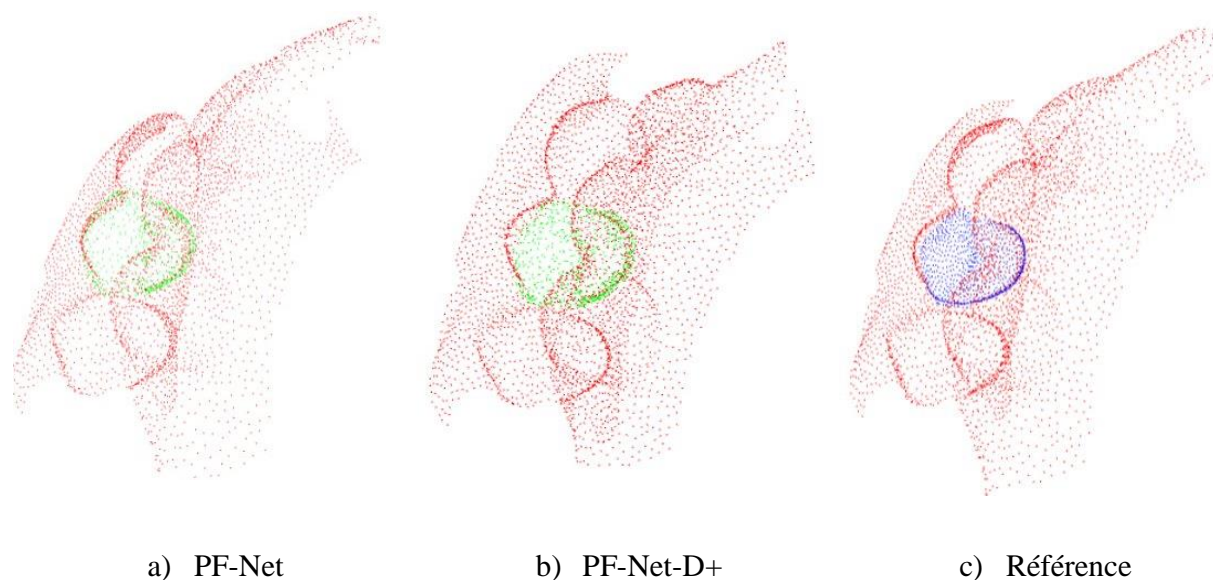
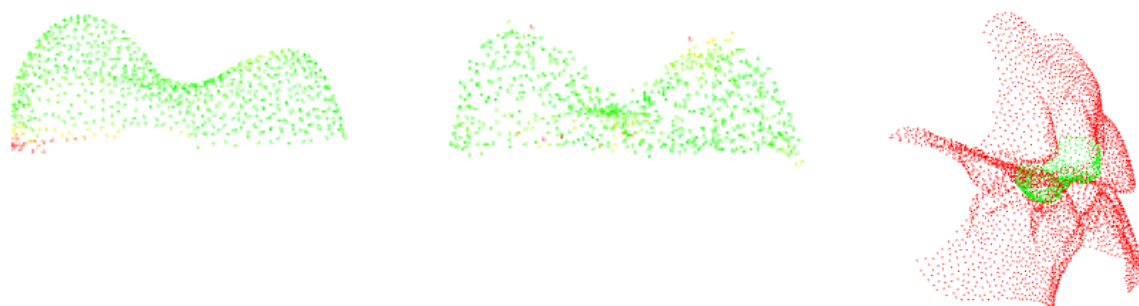


Figure 6.8 Analyse qualitative de l'amélioration avec une vue du dessous des dents.



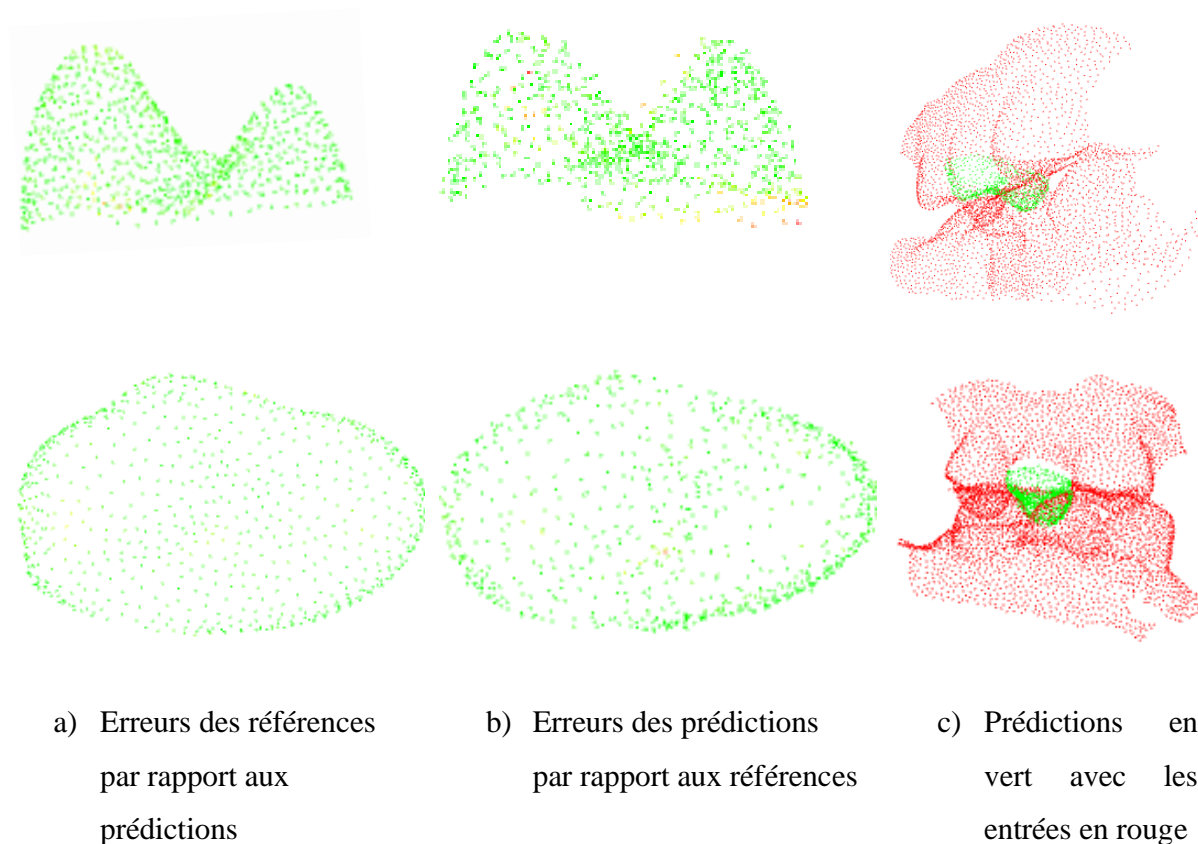


Figure 6.9 Autres résultats pour PF-Net-D+

Dans la Figure 6.8, on voit que la référence possède une ligne marginale précise et qu'elle suit la préparation. De son côté, PF-Net prédit une forme réaliste manquant de détails avec un contour bruité. Sa prédiction possède une bordure qui va au-delà de la préparation. Pour sa part, PF-Net-D+ prédit également une forme réaliste manquant de détails mais avec une bordure mieux positionnée. Dans les prédictions à la Figure 6.9, on peut remarquer aussi des erreurs plus importantes sur les côtés des formes. Ces erreurs proviennent du fait que le modèle est entraîné avec certaines dents possédant des points de contacts avec leurs voisines contrairement à d'autres. Cette variabilité à ce niveau rend la complétion difficile dans cette région en raison de la formulation. L'ajout d'information conditionnelle au discriminateur permet donc de réduire les erreurs au niveau de la ligne de marge pour différentes résolutions de sorties.

CHAPITRE 7 DISCUSSION GÉNÉRALE

Dans ce chapitre, l'ensemble de cette thèse sera discuté et les limitations de l'approche proposée seront présentées davantage.

7.1 Synthèse des travaux

Les travaux présentés dans ce mémoire ont permis de démontrer les limitations des méthodes courantes de complétion de nuage de points et la pertinence des approches de déformations de maillage pour la conception de formes 3D. La base de données comportant des dents manquantes a permis d'entraîner l'architecture PF-Net afin compléter des parties de mâchoires. Tel que démontré dans l'article 1, il est difficile de générer des dents lisses avec un positionnement adéquat. Dans cet article, le modèle arrive à générer plusieurs formes de dents différentes grâce à la flexibilité des nuages de points. En revanche, cette flexibilité entraîne un bruit qui rend les reconstructions de surfaces difficiles. Étant donné la faible quantité de données, le modèle est devenu biaisé vers des objets moyens manquant de détails. Cette perte est due à l'encodage fait dans le décodeur. Tel que mentionné dans PF-Net (Huang et al., 2020), le modèle peut arriver à générer des formes détaillées si la base de données est suffisamment grande. Cependant, lorsque ce n'est pas le cas, le modèle risque de générer des formes moyennes en raison de l'encodage. La base de données n'était donc pas suffisamment grande par rapport à la complexité des formes qu'elle comprenait.

Afin d'aider le positionnement des prédictions, une seconde base de données a été conçue pour entraîner PF-Net à compléter des dents plutôt que générer des dents complètes. Celle-ci comportait uniquement des prémolaires supérieures afin de réduire la variabilité présente dans les entrées et de faciliter l'apprentissage des détails. La génération n'avait donc plus besoin d'une grande flexibilité puisque les formes à générer sont toujours connues. Cette formulation permet au modèle de recevoir en entrée une dent coupée indiquant le contour que la forme devrait posséder. Comme démontré dans l'article 2, cette formulation améliore les prédictions mais PF-Net a encore de la difficulté à produire des formes lisses bien positionnées.

Pour mieux comprendre la provenance du bruit dans les prédictions de PF-Net, un gabarit uniforme et lisse a été ajouté dans le décodeur. La basse résolution est apprise en apprenant des coordonnées relatives par rapport à ce gabarit. Cependant, cet ajout n'a pas permis d'améliorer les prédictions.

Cette expérience démontre le manque de contrôle dans l'architecture pour générer des formes lisses. En utilisant les nuages de points, les connections entre des points voisins ne sont pas considérées lors des opérations. Cette négligence rend la réduction du bruit difficile puisqu'il n'est pas possible d'utiliser des fonctions de coût basées sur les normales ou sur le Laplacien. Pour la conception de couronne dentaire, la réduction du bruit est très importante. Elle permet d'obtenir des formes réalistes et d'améliorer les reconstructions de surface s'il y a lieu. De plus, les approches de complétion de nuages de points utilisent une résolution relativement basse contrairement aux approches de déformation de maillages. Cette résolution est souvent moins que 2000 points.

De leur côté, les approches de reconstructions par apprentissage assument la plupart du temps que le nuage de points d'entrée possède très peu de bruit et une grande résolution. La génération de maillage par apprentissage pourrait être formulée comme une combinaison de complétion de nuage de points suivi d'une reconstruction de surface. Cependant, cette formulation est très limitée pour l'instant puisque les résultats obtenus avec les approches de complétion de nuages de points n'arrivent pas à répondre aux hypothèses des méthodes de reconstructions. Comme démontré dans l'article 2, en raison du bruit dans les prédictions de PF-Net, la reconstruction avec « ball pivoting » doit être faite avec de grands rayons pour réduire le bruit et éviter d'obtenir des trous. Ce besoin entraîne une perte de détails importante lors de la reconstruction.

Un nouveau réseau de complétion de maillage MC-Net a été proposé pour générer des mailles lisses avec détails directement. Ce réseau tire avantage de la connaissance du type de forme à concevoir pour effectuer une génération plus spécialisée. Contrairement à PF-Net, MC-Net offre moins de flexibilité et plus de contrôle sur les complétions. MC-Net utilise une déformation de maillage avec un gabarit de dent. Le gabarit permet d'utiliser une fonction de coût de Laplacien et des opérations de graphes aidant la génération. MC-Net est préférable pour la conception de couronne dentaire puisque le type de dent à concevoir est toujours connu et la conception demande beaucoup de précision et peu de flexibilité. Par sa formulation, MC-Net s'assure de prédire des formes lisses, ce qui est essentiel pour une génération de couronne dentaire.

7.2 Limitations de la solution proposée

Les complétions de dents effectuées par MC-Net sont très prometteuses pour la conception de couronne dentaire. L'architecture proposée arrive à générer des maillages de dents lisses avec

détails. Cependant, les contours des surfaces générées sont souvent erronés. Les fonctions de coûts de complétion et antagoniste sont utilisées lors de l'entraînement pour encourager MC-Net à prédire de bons contours. Cependant, ces fonctions ne sont pas suffisantes pour assurer le respect du besoin mentionné précédemment. Afin de remédier à ce problème, le contour du gabarit pourrait être fixé à la ligne de marge tout au long de l'entraînement. De cette façon, le gabarit épouserait toujours la bordure donnée en entrée et le modèle apprendrait à déplacer les autres sommets du gabarit. La fonction de coût de Laplacien pourrait tout de même être calculée en utilisant le gabarit original sans modification de la bordure. L'article de Wang et al. (Wang et al., 2021) peut servir d'inspiration afin d'intégrer la ligne de marge dans l'apprentissage en plus de servir au gabarit.

Dans Pixel2Mesh et ses améliorations (Wang et al., 2018) (Shi et al., 2021), une projection du maillage sur le plan de caméra permet d'agréger les caractéristiques des sommets. Cette agrégation est intéressante puisqu'elle permet d'établir une correspondance entre les régions locales de l'image et les régions locales du maillage. La correspondance permet d'éviter la perte d'information présente dans les architectures de type encodeur-décodeur. Une mise en commun géométrique permettrait d'éviter cette perte d'information. Cette mise en commun pourrait être inspirée par la mise en commun perceptive utilisée dans Pixel2Mesh. Pour la conception d'une couronne, l'information conditionnelle est un maillage représentant la bouche du patient. Une correspondance de surfaces dentaires serait intéressante. Chaque sommet d'une dent pourrait être associée à une ou plusieurs surfaces dentaires. Pour continuer, les surfaces dentaires des dents voisines pourraient être mis en correspondance avec celles de la couronne. Par exemple, la surface d'occlusion de la dent opposée à la préparation pourrait guider la déformation de la surface d'occlusion de la couronne. De cette façon, les caractéristiques locales des surfaces dentaires voisines seraient transmises avec peu de perte d'information à la couronne. Un encodeur basé sur les Transformers pourrait également permettre de modéliser ces longues dépendances entre surfaces.

La solution proposée comporte une seule couche de convolution de graphe. Cette couche permet de l'information entre des points voisins. En rajoutant d'avantage de couche comme celle-ci, le modèle gagnerait en complexité et chaque point recevrait de l'information d'un plus grand voisinage. Ceci permettrait sûrement d'ajouter des détails dans les formes prédites.

Afin d'améliorer les prédictions de MC-Net, les formes de références pourraient être des maillages. Cela permettrait de définir des fonctions de coût basée sur les normales (Wang et al., 2018) et de rééchantillonner les surfaces avant de les comparer avec CD. Ce type d'échantillonnage est utilisé dans Pixel2mesh++ (Wen et al., 2019) et il permet une meilleur analyse quantitative.

CHAPITRE 8 CONCLUSION

Le travail effectué au long de cette maîtrise est une étape essentielle vers l'automatisation de génération de couronne dentaire. Contrairement aux autres approches existantes, la solution présentée permet de compléter des maillages directement en 3D sans besoin de reconstruction ni d'optimisation pendant l'inférence. Ces avantages sont très intéressants pour le déploiement d'un tel algorithme.

L'architecture proposée MC-Net pourra servir de référence pour de futurs travaux en génération de forme 3D. La prochaine étape pour continuer ce travail serait d'utiliser MC-Net pour générer des surfaces externes de couronne plutôt que des surfaces de dents. Pour ce faire, une nouvelle base de données comprenant des couronnes et des lignes de marge doit être créée. Avec quelques modifications, l'architecture pourra générer des formes de dents détaillées lisses avec un bon contour qui pourront être directement utilisables en production. De plus, avec l'agrandissement des bases de données, les performances des méthodes d'apprentissage profond telle que MC-Net s'approcheront à un niveau de performance similaire à celle d'un technicien dentaire.

8.1 Nouvelles avenues de recherche

En utilisant une base de données comprenant des surfaces de couronnes complètes plutôt que des surfaces de dents partielles, le modèle apprendrait à toujours générer des dents complètes. Cette nouvelle tâche serait sûrement plus facile puisque la variabilité dans les formes à générer serait réduite. Les formes à générer possèderaient toujours des côtés fermés et elles partiraient toujours de la ligne de marge jusqu'aux cuspidés.

La conception de couronne dentaire pourrait également être vue comme une séquence d'opérations CAD. Récemment, Wu et al. ont proposé DeepCAD (Wu et al., 2021). Dans cet article, les auteurs utilisent un modèle basé sur les Transformers pour prédire une séquence d'opérations CAD permettant de générer une forme 3D désirée. Ils arrivent à générer des formes intéressantes avec contours précis. Cependant, leur modèle a plus de difficulté lorsque la séquence d'opération à prédire est longue. Les techniciens dentaires utilisent souvent les mêmes opérations pour concevoir leurs couronnes. Ils utilisent des translations, rotations, mises à l'échelle pour rendre un gabarit réaliste. Par la suite, ils utilisent des opérations d'extrusion sur des surfaces dentaires complètes ou partielles pour améliorer la fonctionnalité de la dent. DeepCAD pourrait être utilisé afin

d'apprendre la suite d'opérations à effectuer sur le gabarit pour obtenir le résultat final désiré. La suite d'opérations devrait être définie uniquement à partir du modèle dentaire et d'un gabarit et elle pourrait être évaluée en comparant le résultat final et celui désiré. Cette approche aurait l'avantage de préserver les détails du gabarit mais elle peut être limitée par le choix des opérations possible et le nombre d'opération requis. Il faudrait donc s'assurer que la suite d'opérations du technicien respecte ces limites.

RÉFÉRENCES

- Achlioptas, P., Diamanti, O., Mitliagkas, I. et Guibas, L. (2018). Learning Representations and Generative Models for 3D Point Clouds. *International conference on machine learning*, 40-49. <http://proceedings.mlr.press/v80/achlioptas18a.html>
- Alshegri, A., Ghadiri, F., Zhang, Y., Lessard, O., Keren, J., Cheriet, F. et Guibault, F. (2022). Semi-supervised segmentation of tooth from 3D scanned dental arches. *Medical Imaging 2022: Image Processing, 12032*, 766-771. <https://doi.org/10.1117/12.2612655>
- Atwood, J. et Towsley, D. (2016). Diffusion-Convolutional Neural Networks. *Advances in Neural Information Processing Systems*, 29. <https://arxiv.org/abs/1511.02136v6>
- Berger, M., Tagliasacchi, A., Seversky, L., Alliez, P., Levine, J., Sharf, A., Silva, C., Seversky, L. M., Levine, J. A. et Silva, C. T. (2014). State of the Art in Surface Reconstruction from Point Clouds. *Eurographics 2014-State of the Art Reports*, 1(1), 161-185. <https://doi.org/10.2312/egst.20141040i>
- Berger, M., Tagliasacchi, A., Seversky, L. M., Alliez, P., Guennebaud, G., Levine, J. A., Sharf, A. et Silva, C. T. (2017). A Survey of Surface Reconstruction from Point Clouds. *Computer Graphics Forum*, 36(1), 301-329. <https://doi.org/10.1111/cgf.12802>
- Bernardini, F., Mittleman, J., Rushmeier, H., Silva, C., Taubin, G. et Member, S. (1999). The Ball Pivoting Algorithm for Surface Reconstruction. *IEEE Transactions on Visualization and Computer Graphics*, 5(4), 349-359.
- Blanz, V. et Vetter, T. (1999). A Morphable Model For The Synthesis Of 3D Faces. *Proceedings of the 26th annual conference on Computer graphics and interactive techniques*, 187-194.
- Cazals, F. et Giesen, J. (2004). *Delaunay triangulation based surface reconstruction: ideas and algorithms*.
- Chen, X., Ravikumar, N., Xia, Y., Attar, R., Diaz-Pinto, A., Piechnik, S. K., Neubauer, S., Petersen, S. E. et Frangi, A. F. (2021). Shape registration with learned deformations for 3D shape reconstruction from sparse and incomplete point clouds. *Medical Image Analysis*, 74, 102228. <https://doi.org/10.1016/J.MEDIA.2021.102228>
- Chen, X., Chen, B. et Mitra, N. J. (2019). Unpaired point cloud completion on real scans using

- adversarial training. *arXiv preprint arXiv:1904.00069*.
- Cheng, S., Bronstein, M., Zhou, Y., Kotsia, I., Pantic, M. et Zafeiriou, S. (2019). MeshGAN: Non-linear 3D Morphable Models of Faces. *arXiv preprint arXiv:1903.10384*. <http://arxiv.org/abs/1903.10384>
- Choy, C. B., Xu, D., Gwak, J., Chen, K. et Savarese, S. (2016). 3D-R2N2: A Unified Approach for Single and Multi-view 3D Object Reconstruction. *European conference on computer vision*, 628-644.
- Cocias, T., Razvant, A. et Grigorescu, S. (2020). GFPNet: A Deep Network for Learning Shape Completion in Generic Fitted Primitives. *IEEE Robotics and Automation Letters*, 5(3), 4493-4500.
- Davis, J., Marschner, S. R., Garr, M. et Levoy, M. (2002). Filling holes in complex surfaces using volumetric diffusion. *Proceedings - 1st International Symposium on 3D Data Processing Visualization and Transmission, 3DPVT 2002*, 428-441. <https://doi.org/10.1109/TDPVT.2002.1024098>
- Donahue, C., McAuley, J. et Puckette, M. (2018). Adversarial audio synthesis. *arXiv preprint arXiv:1802.04208*.
- Elani, H. W., Starr, J. R., Da Silva, J. D. et Gallucci, G. O. (2018). Trends in Dental Implant Use in the U.S., 1999–2016, and Projections to 2026. *Journal of Dental Research*, 97(13), 1424-1430. <https://doi.org/10.1177/0022034518792567>
- Fan, H., Su, H. et Guibas, L. (2017). A Point Set Generation Network for 3D Object Reconstruction from a Single Image. *Proceedings of the IEEE conference on computer vision and pattern recognition*, 605-613. <https://github.com/fanhqme/PointSetGeneration>.
- Fei, B., Yang, W., Chen, W., Li, Z., Li, Y., Ma, T., Hu, X. et Ma, L. (2022). Comprehensive Review of Deep Learning-Based 3D Point Clouds Completion Processing and Analysis. *arXiv preprint arXiv:2203.03311*.
- Foti, S., Koo, B., Dowrick, T., Ramalhinho, J., Allam, M., Davidson, B., Stoyanov, D. et Clarkson, M. J. (2020). Intraoperative Liver Surface Completion with Graph Convolutional VAE. *Uncertainty for Safe Utilization of Machine Learning in Medical Imaging, and Graphs in Biomedical Image Analysis*, 198-207.

- Girdhar, R., Fouhey, D. F., Rodriguez, M. et Gupta, A. (2016). Learning a Predictable and Generative Vector Representation for Objects. *European Conference on Computer Vision*, 484-499.
- Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. et Bengio, Y. (2014). Generative Adversarial Nets. *Advances in Neural Information Processing Systems*. <http://www.github.com/goodfeli/adversarial>
- Guillard, B. B., Remelli, E., Lukoianov, A., Richter, S., Bagautdinov, T., Baque, P. et Fua, P. (2021). DeepMesh: Differentiable Iso-Surface Extraction. *arXiv preprint arXiv:2106.11795*.
- Gupta, S., Arbeláez, P., Girshick, R. et Malik, J. (2015). Aligning 3D Models to RGB-D Images of Cluttered Scenes. *Proceedings of the IEEE conference on computer vision and pattern recognition*, 4731-4740.
- pixel2
- Hwang, J.-J., Azernikov, S., Efros, A. A. et Yu, S. X. (2018). Learning Beyond Human Expertise with Generative Models for Dental Restorations. *arXiv preprint arXiv:1804.00064*.
- Isola, P., Zhu, J.-Y., Zhou, T. et Efros, A. A. (2017). Image-to-Image Translation with Conditional Adversarial Networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 1125-1134. <https://github.com/phillipi/pix2pix>.
- Oord, A., Kalchbrenner, N., Com, K. et Kavukcuoglu, K., (2016). Pixel Recurrent Neural Networks. *International conference on machine learning*, 1747-1756.
- Kazhdan, M., Bolitho, M. et Hoppe, H. (2006). Poisson Surface Reconstruction. *Proceedings of the fourth Eurographics symposium on Geometry processing*.
- Kazhdan, M. et Hoppe, H. (2013). Screened Poisson Surface Reconstruction. *ACM Transactions on Graphics (ToG)*, 32(3), 1-13.
- La couronne dentaire : quand? pourquoi? comment? | Clinique dentaire Laetitia Trudel Baribault.* (s. d.). Consulté 11 juin 2021, à l'adresse <https://www.cliniquedentaireltb.com/la-couronne-dentaire-quand-pourquoi-comment/>
- Lessard, O., Guibault, F., Keren, J. et Cheriet, F. (2022). Dental Restoration using a Multi-Resolution Deep Learning Approach. *2022 IEEE 19th International Symposium on*

Biomedical Imaging (ISBI), 1-4. <https://doi.org/10.1109/ISBI52829.2022.9761622>

- Lian, C., Wang, L., Wu, T. H., Wang, F., Yap, P. T., Ko, C. C. et Shen, D. (2020). Deep Multi-Scale Mesh Feature Learning for Automated Labeling of Raw Dental Surfaces from 3D Intraoral Scanners. *IEEE Transactions on Medical Imaging*, 39(7), 2440-2450. <https://doi.org/10.1109/TMI.2020.2971730>
- Liao, S.-H., Liu, S.-J., Zou, B.-J., Ding, X., Liang, Y. et Huang, J.-H. (2015). Automatic Tooth Segmentation of Dental Mesh Based on Harmonic Fields. *BioMed research international*. <https://doi.org/10.1155/2015/187173>
- Litany, O., Bronstein, A., Bronstein, M. et Makadia, A. (2017). Deformable Shape Completion with Graph Convolutional Autoencoders. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1886-1895. <http://arxiv.org/abs/1712.00268>
- Tchapmi, L., Kosaraju, V., Hamid RezaTofighi, S., Reid, I. et Savarese, S. (2019). TopNet: Structural Point Cloud Decoder. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 383-392. <http://completion3d.stanford.edu>
- Mitra, N. J., Guibas, L. J., Pauly, M. et Zürich, E. (2006). Partial and Approximate Symmetry Detection for 3D Geometry. *ACM Transactions on Graphics (ToG)*, 25(3), 560-568. <https://doi.org/10.1145/1179352>
- Mitra, N. J., Pauly, M., Wand, M. et Ceylan, D. (2013). Symmetry in 3D Geometry: Extraction and Applications. *Computer Graphics Forum*, 32(6), 1-23.
- Ping, Y., Wei, G., Yang, L., Cui, Z. et Wang, W. (2021). Self-attention implicit function networks for 3D dental data completion. *Computer Aided Geometric Design*, 90, 102026. <https://doi.org/10.1016/J.CAGD.2021.102026>
- Qi, C. R., Su, H., Mo, K. et Guibas, L. J. (2017). PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation. *Proceedings of the IEEE conference on computer vision and pattern recognition*, 652-660.
- Qi, C. R., Yi, L., Su, H. et Guibas, L. J. (2017). PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space. *Advances in Neural Information Processing Systems*, 30. <https://arxiv.org/abs/1706.02413v1>

- Ranjan, A., Bolkart, T., Sanyal, S. et Black, M. J. (2018). Generating 3D faces using Convolutional Mesh Autoencoders. *Proceedings of the European conference on computer vision (ECCV)*, 704-720. <http://arxiv.org/abs/1807.10267>
- Sarkar, K., Varanasi, K. et Stricker, D. (2017). Learning quadrangulated patches for 3D shape parameterization and completion. *International Conference on 3D Vision (3DV)*, 383-392.
- Sarmad, M., Korea, S., Lee, H. J. et Kim, Y. M. (2019). RL-GAN-Net: A Reinforcement Learning Agent Controlled GAN Network for Real-Time Point Cloud Shape Completion. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 5898-5907.
- Shi, J., Xu, L., Heng, L. et Shen, S. (2021). Graph-Guided Deformation for Point Cloud Completion. *IEEE Robotics and Automation Letters*, 6(4), 7081-7088. <https://doi.org/10.1109/LRA.2021.3097081>
- Shi, Y., Ni, B., Liu, J., Rong, D., Qian, Y. et Zhang, W. (2021). Geometric Granularity Aware Pixel-to-Mesh. *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 13097-13106.
- Smith, E. J., Fujimoto, S., Romero, A. et Meger, D. (2019). GEOMETRICS: Exploiting Geometric Structure for Graph-Encoded Objects. *arXiv preprint arXiv:1901.11461*. <https://github.com/EdwardSmith1884/GEOMETRICS>
- Xie, H., Yao, H., Zhou S., Mao, J., Zhang, S. et Sun W. (2020). GRNet: Gridding Residual Network for Dense Point Cloud Completion. *European Conference on Computer Vision*, 365-381. <https://haozhxie.com/project/grnet>
- Valsesia, D., Fracastoro, G. et Magli, E. (2018). Learning localized generative models for 3d point clouds via graph convolution. *International conference on learning representations*.
- Van Den Oord, A., Dieleman, S., Zen, H., Simonyan, K., Vinyals, O., Graves, A., Kalchbrenner, N., Senior, A. et Kavukcuoglu, K. (2016). Wavenet: A generative model for raw audio. *arXiv preprint arXiv:1609.03499*.
- Vaswani, A., Brain, G., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł. et Polosukhin, I. (2017). Attention Is All You Need. *Advances in Neural Information Processing Systems*, 5998-6008.

- Wang, X., H. Ang Jr, M. et Hee Lee, G. (2020). Cascaded Refinement Network for Point Cloud Completion. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 790-799.
- Wang, X., H. Ang Jr, M. et Hee Lee, G. (2021). Voxel-based Network for Shape Completion by Leveraging Edge Generation. *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 13189-13198.
- Wang, N., Zhang, Y., Li, Z., Fu, Y., Liu, W. et Jiang, Y.-G. (2018). Pixel2Mesh: Generating 3D Mesh Models from Single RGB Images. *Proceedings of the European conference on computer vision (ECCV)*, 52-67.
- Wang, X., H. Ang Jr, M. et Hee Lee, G. (2021). Cascaded Refinement Network for Point Cloud Completion with Self-supervision. *IEEE Transactions on Pattern Analysis and Machine Intelligence*.
- Wang, X., Xu, D. et Gu, F. (2020). 3D model inpainting based on 3D deep convolutional generative adversarial network. *IEEE Access*, 8, 170355-170363. <https://doi.org/10.1109/ACCESS.2020.3024288>
- Wang, Z., Li, H., Ouyang, W. et Wang, X. (2016). Learnable Histogram: Statistical Context Features for Deep Neural Networks. *European Conference on Computer Vision*, 246-262.
- Wen, C., Zhang, Y., Li, Z. et Fu, Y. (2019). Pixel2Mesh++: Multi-View 3D Mesh Generation via Deformation. *Proceedings of the IEEE International Conference on Computer Vision*, 1042-1051. <https://doi.org/10.48550/arxiv.1908.01491>
- Wen, X., Li, T., Han, Z. et Liu, Y.-S. (2020). Point Cloud Completion by Skip-attention Network with Hierarchical Folding. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 1939-1948.
- Wook Shu, D., Woo Park, S. et Kwon, J. (2019). 3D Point Cloud Generative Adversarial Network Based on Tree Structured Graph Convolutions. *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 3859-3868.
- Wu, R., Xiao, C. et Zheng, C. (2021). DeepCAD: A Deep Generative Network for Computer-Aided Design Models. *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 6772-6782.

- Xia, Y., Xia, Y., Li Inceptio, W., Song, R., Cao, K. et Stilla, U. (2021). *ASFM-Net: Asymmetrical Siamese Feature Matching Network for Point Completion*. arXiv preprint arXiv:2104.09587. <https://completion3d.stanford.edu/>.
- Xiang, P., Wen, X., Liu, Y.-S., Cao, Y.-P., Wan, P., Zheng, W. et Han, Z. (2021). SnowflakeNet: Point Cloud Completion by Snowflake Point Deconvolution with Skip-Transformer. *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 5499-5509. <https://doi.org/10.48550/arxiv.2108.04444>
- Yang, Y., Feng, C., Shen, Y. et Tian, D. (2018). FoldingNet: Point Cloud Auto-encoder via Deep Grid Deformation. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 206-215. <http://arxiv.org/abs/1712.07262>
- Yu, X., Rao, Y., Wang, Z., Liu, Z., Lu, J. et Zhou, J. (2021). PoinTr: Diverse Point Cloud Completion with Geometry-Aware Transformers. *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 12498-12507. <https://arxiv.org/abs/2108.08839v1>
- Yu, Y., Huang, Z., Li, F., Zhang, H. et Le, X. (2020). Point Encoder GAN: A deep learning model for 3D point cloud inpainting. *Neurocomputing*, 384, 192-199. <https://doi.org/10.1016/j.neucom.2019.12.032>
- Yuan, F., Dai, N., Tian, S., Zhang, B., Sun, Y., Yu, Q. et Liu, H. (2020). Personalized design technique for the dental occlusal surface based on conditional generative adversarial networks. *International Journal for Numerical Methods in Biomedical Engineering*, 36(5), e3321. <https://doi.org/10.1002/cnm.3321>
- Yuan, W., Khot, T., Held, D., Mertz, C. et Hebert, M. (2018). PCN: Point Completion Network. *2018 International Conference on 3D Vision (3DV)*, 728-737. <http://arxiv.org/abs/1808.00671>
- Zhang, J., Chen, X., Cai, Z., Pan, L., Zhao, H., Yi, S., Yeo, C. K., Dai, B. et Loy, C. C. (2021). Unsupervised 3D Shape Completion through GAN Inversion. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 1768-1777. <https://junzhezhang.github.io/projects/ShapeInversion/>
- Zhu, H., Jia, X., Zhang, C. et Liu, T. (2022). ToothCR: A Two-Stage Completion and Reconstruction Approach on 3D Dental Model. *Advances in Knowledge Discovery and Data*

Mining: 26th Pacific-Asia Conference, PAKDD 2022, Chengdu, China, 161-172.
https://doi.org/10.1007/978-3-031-05981-0_13

Zhu, L., Wang, B., Tian, G., Wang, W. et Li, C. (2021). Towards point cloud completion: Point Rank Sampling and Cross-Cascade Graph CNN. *Neurocomputing*, 461, 1-16.
<https://doi.org/10.1016/J.NEUCOM.2021.07.035>