

Machine learning approaches for segmentation and integration of root canal and dental crown

*2022 Orthodontic Faculty Development
Fellowships (OFDFA)*

Dr Jonas Bianchi Bianchi

jbianchi@pacific.edu
O: 734-846-9793

FollowUp Form

Award Information

In an attempt to make things a little easier for the reviewer who will read this report, please consider these two questions before this is sent for review:

- Is this an example of your very best work, in that it provides sufficient explanation and justification, and is something otherwise worthy of publication? (We do publish the Final Report on our website, so this does need to be complete and polished.)*
- Does this Final Report provide the level of detail, etc. that you would expect, if you were the reviewer?*

Title of Project*

Machine learning approaches for segmentation and integration of root canal and dental crown

Award Type

Orthodontic Faculty Development Fellowship Award (OFDFA)

Period of AAOF Support

July 1, 2022 through June 30, 2023

Institution

University of the Pacific

Names of principal advisor(s) / mentor(s), co-investigator(s) and consultant(s)

Heesoo Oh (Teaching and Education Mentor); Lucia Cevidanes (Research Mentor), and Joorok Park (Clinical Mentor). Consulting personnel: Juan Carlos, Beatriz Paniagua, Reza Soroushmehr, Kayvan Najarian, Antonio Ruellas and Joao Gonçalves.

Amount of Funding

\$20,000.00

Abstract

(add specific directions for each type here)

As a productive and highly motivated junior faculty, this 2022 Orthodontic Faculty Development Fellowships (OFDFA) proposal will foster my career development teaching, clinical, and research goals, through a detailed

mentoring plan that benefits from a strong team of committed mentors with multidisciplinary and complementary expertise.

My educational and teaching objectives include participation at departmental, university, and national level professional experiences. The specific educational and teaching plan encompasses training courses provided by the ADEA/AAL Institute for Allied Health Educators, weekly conferences with Dr. Heesoo Oh (Chair of UoP - Orthodontics) as my departmental mentor to strengthen my didactic teaching in predoctoral and graduate courses, active participation as a member of the Angle Society, as well as the continuation of my professional service as a peer-reviewer for numerous journals.

The clinical development plan will allow me to integrate the technological advances resulting from the proposed research work, through clinical implementation of tools for improved decision-making in the treatment of orthodontic patients, especially in aligners therapy. I intend to structure image analysis hands-on workshops for the students focused on their patient care, having Dr. Joorok Park (clinical director of the UoP - Orthodontics) as my clinical mentor.

In a patient-specific guided research plan, I will test innovative machine learning (ML) approaches to segment tooth crowns from digital dental models (Aim1) and root canal from CBCT images (Aim 2). This proposed work will address gaps of knowledge and the current lack of root position information when planning orthodontic movements. For this reason, I also aim to integrate crowns and roots data into a single three-dimensional digital model (Aim 3). I will work closely with the computational scientists in our research team, providing feedback for the design and training of the ML models on the multimodality imaging data that I will provide, as well as test the clinical application. Expert users' segmentations will serve as the gold standard data as part of the training sets in the validation of the machine learning models. I will meet weekly in teleconferences with Dr. Lucia Cevidanes at the University of Michigan (UofM), who will mentor this proposal research plan. As this work involves novel applications of computational imaging algorithms and processing, I have established as research collaborators: Drs. Juan Prieto (leading ML developer), Kayvan Najarian, and Reza Soroushmehr, experts in machine learning approaches at the UofM Department of Computational Medicine and Bioinformatics and the University of North Carolina, and Dr. Beatriz Paniagua, assistant director of the Medical Computing Team of the Kitware Inc. These collaborations are part of our research team dissemination efforts, in which I will lead the testing of open-source environment decision support tools.

The AAOF support for this Orthodontic Faculty Development Fellowship Award (OFDFA) will, allow me to focus on relevant educational, clinical and research goals as a dedicated orthodontic faculty. My long-term goals are to share with the profession the training that I will acquire, as well as extend the proposed patient-specific tools supported by this OFDFA as part of an application for federal funding on clinical decision support systems for Orthodontic diagnosis, treatment planning and assessment of treatment outcomes, as well as more broadly in Dentistry.

Respond to the following questions:

Detailed results and inferences:*

If the work has been published, please attach a pdf of manuscript below by clicking "Upload a file".

OR

Use the text box below to describe in detail the results of your study. The intent is to share the knowledge you have generated with the AAOF and orthodontic community specifically and other who may benefit from your study. Table, Figures, Statistical Analysis, and interpretation of results should also be attached by clicking "Upload a file".

bianchi J pdf combined (2).pdf

The publications are attached and combined as a single PDF file.

Were the original, specific aims of the proposal realized?*

Research Plan (mentor Dr. Cevidanes):

Our research team has achieved the specific aims outlined in our proposal. Using open-source software, we have developed novel tools to segment the crown from digital dental models (aim 1) and segment root canals from CBCT images (aim 2). These tools have been successfully integrated into the 3D Slicer platform, enhancing their functionality and accessibility. Additionally, we have accomplished the fusion of the root and crown into a single model (aim 3). Based on our promising pilot results and publications, we have submitted an R01 Grant proposal with the aim of developing user-friendly software and automating the proposed tools. The goal is to streamline treatment planning for orthodontists. If funded, the SMART-DOC system will provide orthodontists with an efficient and intuitive software tool to support clinical decision-making in dentistry. Overall, our research endeavors have resulted in the development of innovative tools and the pursuit of funding for a user-friendly clinical decision support system. This advancement holds significant potential for improving orthodontic workflows and enhancing patient care in the field of craniofacial treatments.

Educational and Teaching Plan (mentor Dr. Heesoo Oh):

Under the guidance of Dr. Heesoo Oh, I have made significant progress in my role as a faculty member. We have held weekly meetings to discuss my professional growth. I had the opportunity to present at the AAO 2023 - Research Category and further enhance my knowledge in machine learning and artificial intelligence by completing the course "Fundamentals of Machine Learning for Healthcare." Additionally, I successfully developed a new elective course for dental students called "Artificial Intelligence, Data Science, and Digital Orthodontics," which has been well-received and will be continued annually. In the teaching realm, I have mentored two master's theses and co-mentored two others within the Orthodontics Program. I have also guided six dental students in their research projects related to orthodontics. Furthermore, I have contributed to the academic community by reviewing 12 papers for prestigious journals such as AJODO, OCR, Progress in Orthodontics, and serving as an academic editor for PLOS ONE. I have also held the position of director for the Cephalometric and Biomedical courses. These accomplishments demonstrate that I have successfully achieved my educational and teaching goals. Furthermore, these experiences have provided me with valuable opportunities to grow as a faculty member.

Clinical Plan (mentor Dr. Park):

Under the guidance of Dr. Park, I have had the privilege of directly supervising the treatment planning of patients for 26 residents during their clinics on Fridays. Our weekly meetings have allowed us to discuss the needs of our students and patients, ensuring a positive and healthy environment. I have learned valuable approaches from Dr. Park, which I have already incorporated into my didactic courses and clinical practice. This includes utilizing digital orthodontics for treatment planning with techniques such as MARPE, aligners, and IDB. Overall, Dr. Park has been an exceptional mentor, playing a significant role in my clinical success. Through our collaboration, I have gained valuable insights and applied them to achieve my clinical goals.

In summary, under the guidance of my mentors Dr. Cevidanes, Dr. Heesoo Oh, and Dr. Park, I have achieved the specific aims of our research proposal. Additionally, I have made progress in my educational and teaching endeavors, as well as my clinical development. These achievements have positioned me for continued growth as a faculty member and have contributed to the advancement of dental imaging, treatment planning, and patient care.

Were the results published?*

Yes

Have the results of this proposal been presented?*

Yes

To what extent have you used, or how do you intend to use, AAOF funding to further your career?*

I have utilized the AAOF funding to advance my career and further the goals I have set for myself. The funding has played a crucial role in supporting my research and enabling the development of innovative tools in the field of dental imaging and treatment planning. Specifically, the funding has allowed me to successfully achieve the specific aims outlined in my proposal, including the segmentation of the crown from digital dental models and the segmentation of root canals from CBCT images. Furthermore, the funding has facilitated the submission of an R01 Grant proposal, titled "SMART-DOC Clinical Decision Support System." If successful, this grant will enable the development of a user-friendly software tool that automates the proposed tools, streamlining treatment planning for orthodontists and improving patient care in the field of dentistry. In terms of my educational and teaching goals, the AAOF funding has supported my professional growth and allowed me to expand my knowledge in machine learning and artificial intelligence. Additionally, I have been able to create a successful elective course for dental students on artificial intelligence, data science, and digital orthodontics, providing them with valuable insights into emerging technologies in the field. Lastly, the AAOF funding has contributed to my clinical development through the guidance and mentorship of Dr. Park. Our collaboration has allowed me to apply new approaches, such as digital orthodontics, in treatment planning and gain valuable experience in supervising and guiding residents. Overall, the AAOF funding has been instrumental in driving my career forward and achieving my goals in research, education, and clinical practice. It has provided the necessary resources and opportunities to make significant advancements in the field of dental imaging, treatment planning, and patient care. I am deeply grateful for the support provided by AAOF and will continue to utilize this funding to further my career and make meaningful contributions to the field of orthodontics.

Accounting: Were there any leftover funds?

\$0.00

Published

Citations*

You indicated results have been published. Please list the cited reference/s for publication/s including titles, dates, author or co-authors, journal, issue and page numbers

1. Gillot M, Miranda F, Baquero B, Ruellas A, Gurgel M, Al Turkestani N, Anchling L, Hutin N, Biggs E, Yatabe M, Paniagua B, Fillion-Robin JC, Allemang D, Bianchi J, Cevidanes L, Prieto JC. Automatic landmark identification in cone-beam computed tomography. *Orthod Craniofac Res*. 2023 Feb 21. doi: 10.1111/ocr.12642. Epub ahead of print. PMID: 36811276.

2. Baquero, B., Bianchi J, et al. (2023). Automatic Landmark Identification on IntraOralScans. In: , et al. *Clinical Image-Based Procedures. CLIP 2022. Lecture Notes in Computer Science*, vol 13746. Springer, Cham. https://doi.org/10.1007/978-3-031-23179-7_4

3. Automatic multi-anatomical skull structure segmentation of cone-beam computed tomography scans using 3D UNETR Gillot M, Baquero B, Le C, Deleat-Besson R, Bianchi J, et al. (2022) Automatic multi-anatomical skull structure segmentation of cone-beam computed tomography scans using 3D UNETR. PLOS ONE 17(10): e0275033. <https://doi.org/10.1371/journal.pone.0275033>

4. Jonas Bianchi, Gustavo Mendonca, Maxime Gillot, Heesoo Oh, Joorok Park, Najla Al Turkestani, Marcela Gurgel, Lucia Cevitanes, Three-dimensional digital applications for implant space planning in orthodontics: A narrative review, Journal of the World Federation of Orthodontists, Volume 11, Issue 6, 2022, Pages 207-215,

Was AAOF support acknowledged?

If so, please describe:

Yes the AAOF support was acknowledged.

Presented

Please list titles, author or co-authors of these presentation/s, year and locations:*

1. Accurate Segmentation and Landmark Identification on Intra Oral Scans.Hutin, Nathan, Bianchi, J , Cevitanes, L et al., 2023 AADOCR/CADR Annual Meeting (Portland, Oregon)
2. Automatic Landmark Identification on IntraOralScans. In: et al. Clinical Image-Based Procedures. MICCAI 2022 - 25th International Conference on Medical Image Computing and Computer-Assisted Intervention. Baquero, B., Bianchi J, et al. (2023).Singapore (Presented Online).
3. ALIDDM: Automatic Landmark Identification in Digital Dental Models.Baquero, Baptiste, Bianchi, J, Cevitanes L et cl,m 2022 AADOCR/CADR Annual Meeting. Hybrid, Atlanta, Georgia
4. 3D Skeletal Analysis in Fixed Orthodontics vs Invisalign in. Surgery-First Approach. Jonas Bianchi. AAO 2023 ANNUAL SESSION - Chicago - USA.
5. Automatic Landmark Identification in Cone Beam Computed Tomography Scans. Gillot, Maxime, Bianchi, J, Cevitanes L et cl,m 2022 AADOCR/CADR Annual Meeting. Hybrid, Atlanta, Georgia

Was AAOF support acknowledged?

If so, please describe:

Yes the AAOF support was acknowledged.

Internal Review

Reviewer Comments

Reviewer Status*

File Attachment Summary

Applicant File Uploads

- bianchi J pdf combined (2).pdf



Automatic Landmark Identification on IntraOralScans

Baptiste Baquero^{1,9}(✉), Maxime Gillot^{1,9}, Lucia Cevidanes¹,
Najla Al Turkestani^{1,5}, Marcela Gurgel¹, Mathieu Leclercq^{4,9},
Jonas Bianchi^{1,3}, Marilia Yatabe¹, Antonio Ruellas^{1,2}, Camila Massaro⁸,
Aron Aliaga¹, Maria Antonia Alvarez Castrillon⁶, Diego Rey⁶,
Juan Fernando Aristizabal⁷, and Juan Carlos Prieto⁴

¹ University of Michigan, Michigan, USA

baptistebaquero@gmail.com

² Federal University of Rio de Janeiro, Rio de Janeiro, Brazil

³ University of the Pacific, San Francisco, USA

⁴ University of North Carolina, Chapel Hill, USA

⁵ King Abdulaziz University, Jeddah, Saudi Arabia

⁶ CES University, Medellín, Colombia

⁷ University of Valle, Cali, Colombia

⁸ Federal University of Goiás, Goiânia, Brazil

⁹ CPE Lyon, Lyon, France

Abstract. With the advent of 3D printing and additive manufacturing of dental devices, IntraOral scanners (IOS) have gained wide adoption in dental practices and allowed for efficient workflows in clinical settings. Accurate automatic identification of dental landmarks in IOS is required to aid dental researchers and clinicians to plan and assess tooth position for crown restorations, orthodontics movements, and/or implant dentistry. In this paper, we present a new algorithm for Automatic Landmark Identification on IntraOralScans (ALIiOS), that combines image processing, image segmentation, and machine learning approaches to automatically and accurately identify commonly used landmarks on IOSs. Four hundred and five digital dental models were pre-processed by 3 clinician experts to manually annotate 5 landmarks on each dental crown in the upper and lower arches. Our approach uses the PyTorch3D rendering engine to capture 2D views of the dental arches from different viewpoints as well as the target 3D patches at the location of the landmarks. The ALIiOS algorithm synthesizes these 3D patches with a U-Net and allows accurate placement of the landmarks on the surface of each dental crown. Our results, after cross-validation, show an average distance error between the prediction and the clinicians' landmarks of 0.43 ± 0.28 mm and 0.45 ± 0.28 mm for respectively lower and upper occlusal landmarks, and 0.62 ± 0.28 mm for lower and upper cervical landmarks. There was on average a 5% error of landmarks more than 1.5 mm away from the clinicians' landmarks, due to errors in landmark nomenclature or improper segmentation. In conclusion, we present and validate a novel algorithm for accurate automated landmark identification on intraoral scans to increase efficiency and facilitate quantitative assessments in clinical practice.

Keywords: Deep learning · Automatic landmark identification · Digital dental model

1 Introduction

Digital dental models are obtained by Intraoral scans (IOS are widely used in dentistry). Even if many practices still lack this technology, conventional plaster models are now digitized as services provided by laboratories, to plan the proper placement of the dental crowns, tooth movement [3], fabrication of dental restorations [2], monitoring and maintaining periodontal health, attaining stable treatment outcomes, and the occlusal function [11]. IOSs are detailed 3D surface mesh models of the upper and lower dentition that allow clinicians to accurately evaluate the clinical crown position in three dimensions without radiation exposure to the patient [1]. Time efficiency increased patient comfort, and data fusion options within a computer-aided design and manufacturing technologies increasingly used in dentistry are among the multiple advantages of IOS systems [5]. Given that intraoral scanning and digitization of tooth geometries is a fundamental step in the dental digital workflow, the accuracy of measurements in IOS must be evaluated critically. Dentists need to segment each tooth in the IOS and annotate the corresponding anatomical landmarks to analyze, rearrange and/or restore tooth position. Manual performance of these tasks is time-consuming and prone to inconsistency. There is a clinical need to develop fully automatic methods instead of manual operation. The development of an artificial intelligence tool for landmark localization of dental crown surfaces is challenging, mainly due to variability of the anatomical structures of different teeth, abnormal, disarranged, and/or missing teeth for some patients. Compared with the individual tooth segmentation and labeling, the localization of anatomical landmarks is typically more sensitive to the variable shape appearance of each patient’s teeth, as each tooth’s landmarks are just small points encoding local geometric details. Facing this challenge, in this paper, we present an algorithm for Automatic Landmark Identification on IntraOralScans (ALIOS) to predict 3 occlusal landmarks and 2 cervical landmarks on the upper and lower dental arches in a total of 140 landmarks, based on the segmentation of precise patch locations. In the following sections, we describe the materials, briefly review the most relevant related work, describe the study datasets, the proposed algorithm with the training and testing steps, and the results.

2 Related Work

Landmark localization remains a crucial task in both computer vision and medical imaging analysis, and the computer vision community has collectively attempted numerous approaches to address this task. Occlusion-net [7] implements an approach that encourages occlusions, where a camera can only view one side of an object (left or right, front, or back), and part of the object is outside the field of view. The framework then predicts 2D and 3D locations of

occlusal key points for objects, in a largely self-supervised manner, using an off-the-shelf detector as input that is trained only on visible key point annotations. Then a graph encoder network explicitly classifies invisible edges, and a graph decoder network corrects the occluded key point locations from the initial detector. Another method uses a heatmap regression-based landmark localization on IOS datasets [10]. It incorporates the spatial configuration of anatomical landmarks at the region of interest of individual teeth to improve the robustness of the regression. Other approaches for IOS processing have determined the IOS orientation, then used the local maxima in the vertical direction for an initial approximation of the landmarks, followed by an extraction of surface gradient and curvature information to identify the shape and boundaries of each tooth. [9].

3 Method

3.1 Data

The dataset consisted of four hundred and five IOSs of the upper and lower dental arches acquired at 2 clinical centers: Universidad Corporación para Estudios en la Salud (CES) in Medellin Colombia and University of Michigan. These scans were acquired using 3Shape Trios and iTero® intraoral scanners. The scanners utilize ultrafast optical sectioning and confocal microscopy to generate 3D images from multiple 2D images with an accuracy of $6.9 \pm 0.9 \mu\text{m}$. The dataset was composed of individual anatomic shapes, patients could present one or more missing teeth and a third molar and dental appliances (braces). For each IOS, 70 maxillary dental landmarks were placed by 3 experienced clinicians on each arch, using the markups module in 3D Slicer 4.11 [6]. For each IOS, we recovered important information useful for the training steps and included the vertices, faces of the mesh, label of each face (or the positions of the landmarks), and the normal vector for each vertex.

3.2 Pre-processing

The IOSs were pre-processed using an open-source tool in 3D Slicer 4.11 [6], DentalModelSeg [4], to segment and assign the universal numbering to each tooth respectively. The scan pre-processing allowed the selection of each tooth to predict landmark placement (Fig. 1).

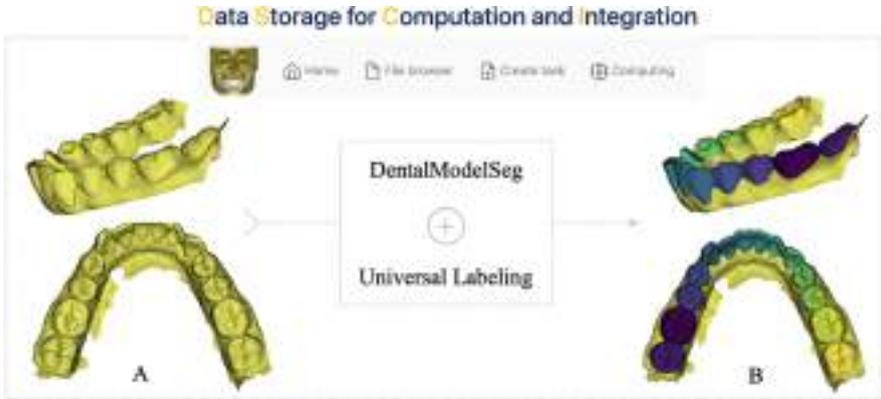


Fig. 1. IOS pre-processing A. IOS acquired using iTero® or 3Shape scanners B. Segmentation of the dental crowns with the two open-source tools: DentalModelSeg and Universal Labelling [4] to segment and assign the universal numbering.

3.3 Rendering 2D Views

We used PyTorch3D framework that allows fast 3D data representation and batching, *i.e.*, there are no intermediate pre-processing steps on the input meshes. This library was used to perform end-to-end training by rendering images of the IOS meshes that are fed to the ALIOS convolutional neural network (CNN). PyTorch3D renderers are designed to be modular, extensible, and ready to perform gradient computation. The renderers are based on two principles:

- **Rasterizer:** The rasterization consists of projecting a 3D object on a 2D image. It uses a camera such as the FoVPerspectiveCamera which was used following the OpenGL convention for perspective and orthographic cameras. This camera is by default in the NDC coordinate system, which is a normalized coordinate system that confines in a volume the rendered part of the object/scene. In this work, we used 224×224 pixels images with a 120° FOV which allowed us to take perfect 2D views of the target area. As well as the image, the rasterizer outputs a look-up table that links each pixel on the rendered image to a corresponding face on the mesh.
- **Shaders:** The shaders are used to apply texturing/shading/blending on the rasterized images. It needs a light source as well as textures on the meshes. In this work, to generate the input images, we placed a light source in front of the 3D model, and used the normal at each vertex (encoded in RGB components) as the texture of the mesh. The mesh renderer is a Pytorch3D “HardPhong-Shading” shader.

3.4 Training

We trained a residual U-Net [8] architecture from Monai, with 4 down-sampling steps and 4 up-samplings, kernel 3×3 and stride 2, with an increasing number of

features starting at 64 up to 512. This implementation used residual units during training. The objective of ALIOS is to segment patches on the tooth surface around the landmark defined by the clinicians. To do that, we first centered and scaled the meshes to be in a unit sphere. We trained one model to identify landmarks in the upper and one model for the lower arch. Using the universal labeling for each crown, [4] we moved the cameras tooth by tooth, located the region of interest and rendered the surface of the crown. Each camera was placed on a sphere with a defined radius and the camera was oriented to look at the center of the tooth (this view is determined by taking the average of all coordinates of the tooth’s vertex). Depending on the position of the landmarks, the cameras’ positions will be different (top views of the crowns for occlusal landmarks and side views of the crowns for cervical landmarks) to make predictions. For all views, we rendered 2D RGB images (normal vectors encoded in RGB components) and a depth map as a fourth channel. These images are then fed to the ALIOS U-Net (Fig. 2 A). These depth maps were grayscale representations of the distance of the faces to the cameras. For the ground truth, we used the pix-to-face lookup table to retrieve the corresponding labeled images of uniform patches with unique colors for each type of landmark (Fig. 2 B). We used DiceCEloss to compare similarities between the output and the ground truth and the ADAM optimization algorithm for stochastic gradient descent. The learning rate was $1e-4$, with a batch size of ten for the occlusal landmarks and a batch size of one for the cervical landmarks. To train each model, 6 GB of the GPU was used and the training took an average of 5 h. The training was done on a workstation with 2 NVIDIA Corporation GP102 [TITAN Xp] graphic cards, Intel® Core™ i7-8700K CPU @ 3.70 GHz \times 12 processors, and 2 TB disk capacity.

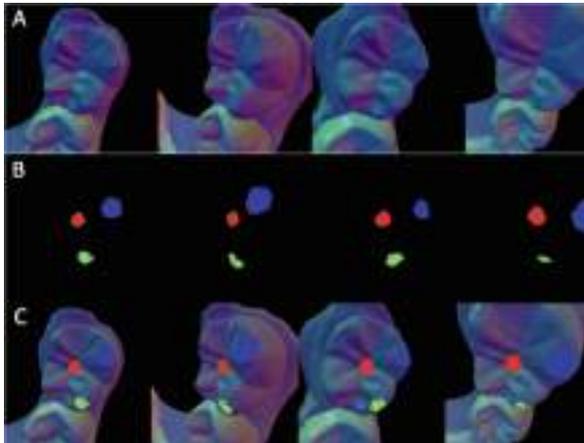


Fig. 2. A.U-Net input images. B.U-Net output patches. C.Identification of the surface meshes vertices using the U-Net output.

3.5 Prediction

To predict the landmarks on an IOS, after the segmentation with universal labeling, we moved the cameras to adequate positions. The 2D images generated by the renderer were set as input of the ALIiOS U-Net. To improve accuracy, post-processing steps were applied to the output to clean the pixels of the patches incorrectly placed on adjacent teeth. Only the faces that belong to the target tooth remained. Using the “pix to face” function on the segmented patches, we identified the faces corresponding to each pixel of the patches. For each patch color, we collected all the corresponding faces and averaged all vertices coordinates to find an approximated position. The final predicted landmark position was identified as the closest point to the approximated point on the mesh surface, saved as a fiducial list that contains all the landmarks.

4 Results

To test the performance of the ALIiOS approach in our entire dataset, we performed a 5-fold cross-validation, each time using a different 20% portion of the available data as a test set that was not included in the training. A fiducial list was generated with the predicted positions of the landmarks in about 1 min. To compute the prediction accuracy, we compared the distance between the clinicians’ landmarks and the predicted landmarks. The clinically acceptable distance range that landmark prediction should not exceed is 1 mm. Figure 3 shows the average accuracy for each tooth in the lower jaw. Table 1 summarizes the accuracy of each different model. A violin plot of each type of model is presented below in Fig. 4 to 7.

Table 1. Accuracy results table for occlusal and cervical landmarks on lower and upper arches

	Upper	Lower
Occlusal	0.45 ± 0.28 mm	0.43 ± 0.28 mm
Cervical	0.62 ± 0.28 mm	0.62 ± 0.28 mm

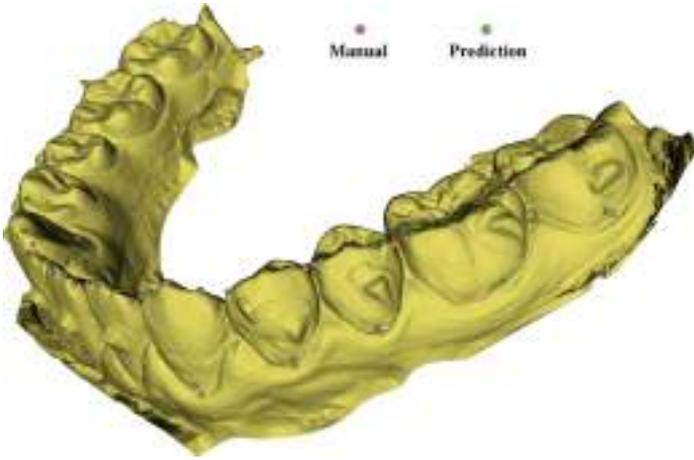


Fig. 3. Comparison between manual landmarks and predicted on the lower arch. The red spheres represent the clinician’s landmarks (manual) and the green spheres are the ones predicted by ALIIOS. The diagram displays the average error (mm) for each landmark. (Color figure online)

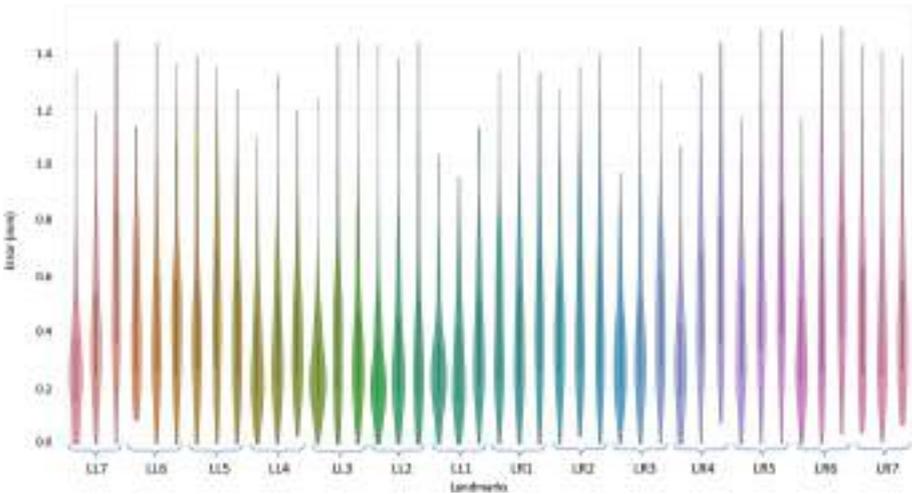


Fig. 4. Accuracy for the lower occlusal landmarks.

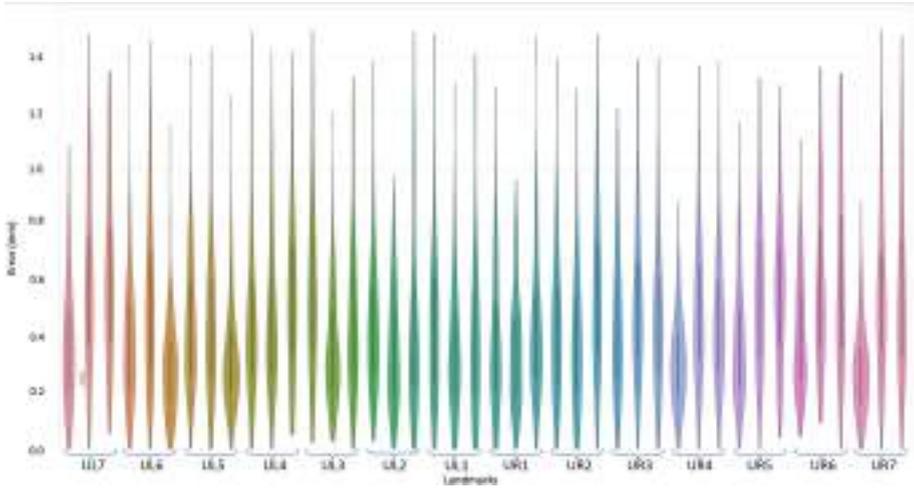


Fig. 5. Accuracy for the upper occlusal landmarks.

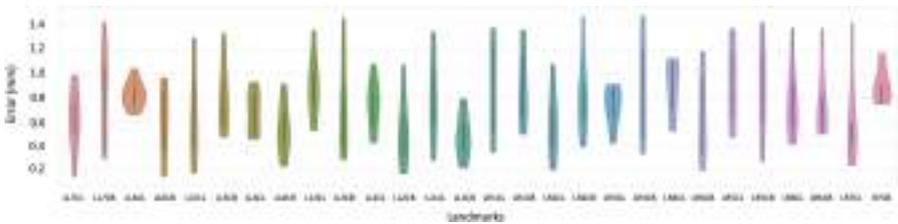


Fig. 6. Accuracy for the lower cervical landmarks.

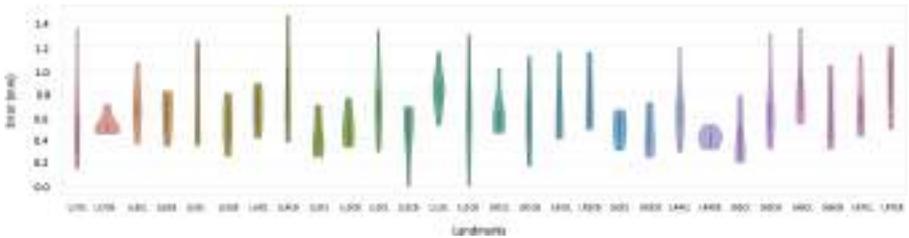


Fig. 7. Accuracy for the upper cervical landmarks.

Landmark name description: The first letter represents upper or lower. The second character represents left or right. Then the tooth number and finally the landmark type. In Fig. 4 and 5, each tooth has, from left to right, occlusal, mesial buccal, and distal buccal landmarks. In Fig. 6 and 7, “CL” stands for cervical lingual and “CB” stands for cervical buccal.

5 Discussion

This paper presents the ALIOS algorithm, a novel method for robust and accurate automatic landmark identification on IOS. The ALIOS approach is more precise (Table 1 and Fig. 3) and outperforms previously published approaches: the accuracy of the landmarks predicted with the iMeshSegNet+PointNet-Reg algorithm was 0.597 ± 0.761 mm [10], and the automatic landmark recognition (ALR) algorithm was 0.389 mm [9]. A potential limitation of the ALR approach proposed by Woodsend et al. [9] is that it is based on the local maxima, detecting only landmarks in the tips of the cusp. The ALIOS approach is also more flexible and allows for variability in positioning the cameras according to the location and clinical needs to place landmarks. Furthermore, the ALIOS method is time efficient, as it takes less than one minute to predict all the landmarks on each dental in comparison to manual landmark placement which is time-consuming and prone to inconsistencies. To facilitate its use by clinicians and researchers in dentistry, the ALIOS tool has been deployed as a 3D Slicer extension [6] and the open-source code is available on Github (<https://github.com/baptistebaquero/ALIDDM.git>). The ALIOS intuitive interface allows users to predict occlusal or cervical landmarks on the selected tooth. Additionally, to allow users to automatically compute measurements between the ALIOS landmarks, the work in progress will be to implement another Slicer extension called AQ3DC (Automatic Quantification 3D Components). AQ3DC automatically computes lists of measurements selected by users for a single case or a whole study sample, at one or more time points. This user-friendly tool aims to decrease users' time for the extraction of quantitative image analysis features. The AQ3DC implementation is aimed at the automatic computation of 3D components of the directionality of distances (Anteroposterior, Right/Left, Supeoinferior) between points, point to line, the midpoint between two points, or angles (Pitch, Roll, and Yaw), which can be further extended to any type of desired computation/quantitative image analysis. The design of the user interface is currently aimed at the quantification of craniofacial dental, skeletal and soft tissue structures. The ALIOS tool has been developed as part of a learning health system in dentistry that integrates root canal surface meshes to IOS dental crowns toward detecting the tooth long axes that is clinically relevant for restorations, implant placement, and tooth movement [4] (Fig. 8). The present study lays the groundwork for machine learning approaches that synthesize crown information for quantitative assessments. Future studies will utilize multi-modality merging and annotation of cone-beam CT and IOS scans for challenging craniofacial applications that require both imaging modalities.

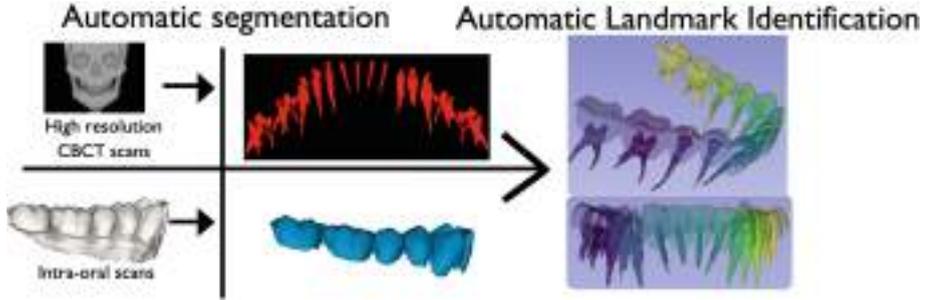


Fig. 8. Proposed future work: Segmentation of root canal and prediction of the tooth long axis.

6 Conclusion

We developed and validated a novel ALIIOS algorithm to automatically identify teeth landmarks on IOS. Our algorithm is optimized using Monai and PyTorch libraries. The ALIIOS predicts the location of landmark patches and identifies the final precise landmark position following post-processing steps. Our method has a precision of 0.43 ± 0.28 mm and 0.45 ± 0.28 mm for respectively lower and upper occlusal landmarks, and 0.62 ± 0.28 mm for lower and upper cervical landmarks. Overall, these findings demonstrate the clinical application of ALIIOS to more automated quantitative 3D imaging assessments in dental research and practice.

Acknowledgments. Supported by NIDCR R01 024450, AA0F Dewel Memorial Biomedical Research award by Research Enhancement Award Activity 141 from the University of the Pacific, Arthur A. Dugoni School of Dentistry, and the American Association of Orthodontists Foundation (AAOF).

References

1. Burhardt, L., Livas, C., Kerdijk, W., van der Meer, W.J., Ren, Y.: Treatment comfort, time perception, and preference for conventional and digital impression techniques: a comparative study in young patients. *Am. J. Orthod. Dentofac. Orthop.* **150**(2), 261–267 (2016)
2. Chiu, A., Chen, Y.W., Hayashi, J., Sadr, A.: Accuracy of cad/cam digital impressions with different intraoral scanner parameters. *Sensors* **20**(4), 1157 (2020)
3. Cong, A., et al.: Dental long axes using digital dental models compared to cone-beam computed tomography. *Orthod. Craniofac. Res.* **25**(1), 64–72 (2022)
4. Deleat-Besson, R., et al.: Automatic segmentation of dental root canal and merging with crown shape. In: 2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), pp. 2948–2951. IEEE (2021)

5. Patzelt, S.B., Lamprinos, C., Stampf, S., Att, W.: The time efficiency of intraoral scanners: an in vitro comparative study. *J. Am. Dent. Assoc.* **145**(6), 542–551 (2014)
6. Pieper, S., Halle, M., Kikinis, R.: 3D slicer. In: 2004 2nd IEEE International Symposium on Biomedical Imaging: Nano to Macro (IEEE Cat No. 04EX821), pp. 632–635. IEEE (2004)
7. Reddy, N.D., Vo, M., Narasimhan, S.G.: Occlusion-net: 2d/3d occluded keypoint localization using graph networks. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 7326–7335 (2019)
8. Ronneberger, O., Fischer, P., Brox, T.: U-Net: convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) MICCAI 2015. LNCS, vol. 9351, pp. 234–241. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24574-4_28
9. Woodsend, B., et al.: Development of intra-oral automated landmark recognition (ALR) for dental and occlusal outcome measurements. *European J. Orthod.* **44**(1), 43–50 (2022)
10. Wu, T.H., et al.: Two-stage mesh deep learning for automated tooth segmentation and landmark localization on 3D intraoral scans. arXiv preprint [arXiv:2109.11941](https://arxiv.org/abs/2109.11941) (2021)
11. Zimmermann, M., Ender, A., Mehl, A.: Local accuracy of actual intraoral scanning systems for single-tooth preparations in vitro. *J. Am. Dent. Assoc.* **151**(2), 127–135 (2020)

Automatic landmark identification in cone-beam computed tomography

Maxime Gillot^{1,2} | Felicia Miranda^{1,3}  | Baptiste Baquero^{1,2} | Antonio Ruellas⁴ |
 Marcela Gurgel¹  | Najla Al Turkestani^{1,5}  | Luc Anchling^{1,2} | Nathan Hutin^{1,2} |
 Elizabeth Biggs¹ | Marilia Yatabe¹  | Beatriz Paniagua⁶ |
 Jean-Christophe Fillion-Robin⁶ | David Allemang⁶ | Jonas Bianchi⁷  |
 Lucia Cevidanes¹  | Juan Carlos Prieto⁸ 

¹Department of Orthodontics and Pediatric Dentistry, University of Michigan School of Dentistry, MI, Ann Arbor, USA

²CPE Lyon, Lyon, France

³Department of Orthodontics, Bauru Dental School, University of São Paulo, Bauru, Brazil

⁴Department of Orthodontics and Pediatric Dentistry, School of Dentistry, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil

⁵Department of Restorative and Aesthetic Dentistry, Faculty of Dentistry, King Abdulaziz University, Jeddah, Saudi Arabia

⁶Kitware Inc., Chapel Hill, NC, USA

⁷Department of Orthodontics, University of the Pacific, San Francisco, CA, USA

⁸Department of Psychiatry, University of North Carolina, Chapel Hill, NC, USA

Correspondence

Felicia Miranda, Department of Orthodontics and Pediatric Dentistry, University of Michigan School of Dentistry, 1011 N University Ave, Ann Arbor, MI 48109, USA.
 Email: felmiran@umich.edu

Funding information

American Association of Orthodontists Foundation; National Institute of Dental and Craniofacial Research

Abstract

Objective: To present and validate an open-source fully automated landmark placement (ALICBCT) tool for cone-beam computed tomography scans.

Materials and Methods: One hundred and forty-three large and medium field of view cone-beam computed tomography (CBCT) were used to train and test a novel approach, called ALICBCT that reformulates landmark detection as a classification problem through a virtual agent placed inside volumetric images. The landmark agents were trained to navigate in a multi-scale volumetric space to reach the estimated landmark position. The agent movements decision relies on a combination of DenseNet feature network and fully connected layers. For each CBCT, 32 ground truth landmark positions were identified by 2 clinician experts. After validation of the 32 landmarks, new models were trained to identify a total of 119 landmarks that are commonly used in clinical studies for the quantification of changes in bone morphology and tooth position.

Results: Our method achieved a high accuracy with an average of 1.54 ± 0.87 mm error for the 32 landmark positions with rare failures, taking an average of 4.2 second computation time to identify each landmark in one large 3D-CBCT scan using a conventional GPU.

Conclusion: The ALICBCT algorithm is a robust automatic identification tool that has been deployed for clinical and research use as an extension in the 3D Slicer platform allowing continuous updates for increased precision.

KEYWORDS

anatomic landmarks, fiducial markers, machine learning

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2023 The Authors. *Orthodontics & Craniofacial Research* published by John Wiley & Sons Ltd.

1 | INTRODUCTION

As artificial intelligence (AI) technology is evolving and being adopted in clinical practice and new workflows, the greatest challenges are to evaluate and monitor the agility, stability, adaptability and robustness of AI algorithms, towards ensuring clinical quality and patient safety. The best practices for AI infrastructure include clinical imaging data access and security, integration across platforms and domains, clinical translation and delivery, and a culture of inclusive participation and continuous updates. However, the rapid increase in the number of commercially available algorithms and the variety of ways in which each algorithm can affect clinical workflows adds complexity to the AI implementation process.^{1,2}

The accurate anatomical landmark localization for medical imaging data is a challenging problem due to the frequent ambiguity of their appearance and the rich variability of the anatomical structures. Landmark detection represents a prerequisite for medical image analysis. It supports entire clinical workflows from diagnosis,³ treatment planning,⁴ intervention, follow-up of anatomical changes, or disease conditions,⁵ and simulations.⁶ Landmark identification may serve as initialization to other algorithms such as segmentation algorithms,⁷ or image-to-image registration.^{8,9} Most of the available solutions for landmark detection rely on machine learning,^{10–12} however, previous methods have been proposed for other image modalities and have not been validated for Cone-beam computed tomography (CBCT) scans with various imaging acquisition protocols to lower radiation dose in dentistry. Other approaches for landmark identification rely on sub-optimal search strategies, i.e., exhaustive scanning,^{11,12} one-shot displacement estimation,^{13,14} or end-to-end image mapping techniques.^{15,16} In many cases, these methods can lead to false-positive detection results and excessively high computation times.

The application of AI technology for the automatic landmark identification in CBCT can help to promote precise and more efficient landmark location in different craniofacial structures of interest for oral research and clinical aspects.² In the present study, the landmark detection task is set up as a behaviour classification problem for an artificial agent that navigates through the voxel grid of the image at different spatial resolutions. The aim of this study was to present and validate a new automated landmark identification method for CBCT (ALICBCT) inspired by a deep reinforcement learning system (DRL) technique.

2 | MATERIALS AND METHODS

This secondary data analysis was approved by the Institutional Review Board of the University of Michigan, School of Dentistry (HUM00217436). The sample was composed by 143 de-identified CBCT scans of patients acquired in 6 different University Centers (University of Michigan - School of Dentistry, University of University of Pacific - School of Dentistry, Scientific University of the South in Peru, National University of Colombia, CES University and Federal University of Ceará). The inclusion criteria were permanent dentition

and image availability acquired for dental clinical purposes. The exclusion criteria were patients with craniofacial anomalies or syndromes and scans with artefacts produced by orthodontic appliances.

Two open-source software packages, ITK-SNAP, version 3.8 (www.itksnap.org)¹⁷ and 3D Slicer, version 4.11 (www.slicer.org)¹⁸ were used by clinician experts to orient the scans and place the landmarks. Head orientation was performed accordingly with a previous study.¹⁹ For the large field of view, CBCT scans orientation was standardized across patients with Frankfort horizontal plane matching the axial plane, and the midsagittal plane matching the sagittal plane in a common coordinate system. For the small/medium field of view, the axial plane orientation was determined by the occlusal plane and the midsagittal plane by the midpalatal suture. A set of 32 landmarks located in different anatomical structures, including the cranial base, maxilla, mandible, an teeth (Table 1) was created by the clinician experts, which was considered the ground truth (GT) fiducial list.

The present method relies on two principles: a multi-scale environment and a search agent inspired by the behavioural problem solved as described in DRL Systems.²⁰

2.1 | Environment

The sample consisted of 77 large field of view CBCTs with voxel sizes varying from 0.3 to 0.4 mm, and 66 small/medium field of view CBCTs with voxel sizes varying from 0.08 to 0.16 mm. In order to obtain environments with the finest scale level, the large field of view scans were re-sampled to an isotropic resolution of 0.3 mm and the small/medium field of view scans were re-sampled to 0.08 mm. We wanted the agent to learn different scales of the structures of interest. For our multi scale-space, we used an additional low-resolution level at an isotropic spatial resolution of 1 mm. The image histograms were re-scaled to have better contrast and the data was normalized to a [-1.0, 1.0] interval. A multi-scale environment can be seen in Figure 1A. For each CBCT, the 32 landmark were marked by clinicians and stored as a fiducial list. During the training, the landmark's position from the list was mapped to the discrete image coordinates for each resolution and stored in the environment memory.

2.2 | Agent

The protagonist of this work was the agent. The agent is a virtual object whose goal is to reach a target position (the landmark) by moving inside an environment. The agent has a set of 6 possible actions, to move from one voxel to another by going superiorly, inferiorly, anteriorly, posteriorly, left or right.

The agent state is a 3D box around the agent position that has been cropped inside the environment (Figure 1B). The size of the FOV is an important parameter, and we have to make sure that enough relevant image features can be extracted at the current location while limiting memory usage. The agents use deep networks for feature extraction (FeatNet), followed by fully connected layers

TABLE 1 Landmarks definition.

Description of the landmarks	
Cranial base	
Ba	Placed at the most posteroinferior point of the anterior margin of the foramen magnum in the midsagittal plane
S	Placed on the most central point of sella turcica from supero-inferior, antero-posterior, and transversal aspects
N	Placed at the most anterosuperior junction of the nasofrontal suture
Maxilla	
A	The most posterior point of the concavity of the anterior region of the maxilla
ANS	Placed at the anterior nasal spine
PNS	Placed at the posterior nasal spine
UR6DB	Placed at the distal buccal cusp of the maxillary right permanent first molars
UR6MB	Placed at the mesial buccal cusp of the maxillary right permanent first molars
UR6R	Placed at the center of the pulp chamber floor of the maxillary right permanent first molars
UR3O	Placed at the cusp tip of the maxillary right permanent canine
UR3R	Placed at the center portion of the root canal at the axial level of the cemento-enamel junction of the maxillary right permanent canine
UR1R	Placed at the center portion of the root canal at the axial level of the cemento-enamel junction of the maxillary right permanent central incisor
UL3O	Placed at the cusp tip of the maxillary left permanent canine
UL3R	Placed at the center portion of the root canal at the axial level of the cemento-enamel junction of the maxillary left permanent canine
UL6MB	Placed at the mesial buccal cusp of the maxillary left permanent first molars
UL6R	Placed at the center of the pulp chamber floor of the maxillary left permanent first molars
UR1O	Placed in the middle of the incisal edge of the maxillary right permanent central incisor
Mandible	
LR6MB	Placed at the mesial buccal cusp of the mandibular right permanent first molars
LR6R	Placed at the center of the pulp chamber floor of the mandibular right permanent first molars
LR1R	Placed at the center portion of the root canal at the axial level of the cemento-enamel junction of the mandibular right permanent central incisor
B	Placed at the most posterior point of the concavity of the anterior region of the symphysis
Pog	Placed at the most anterior point of the symphysis
Gn	Placed in the projection of a virtual bisector of a line adjacent to the Pog and Me landmarks
Me	Placed at the most inferior point of the chin
RGo	Placed in the projection of a virtual bisector of a line adjacent to the right mandibular base and right posterior border of mandible
RCo	Placed at the most superior and central point of right condyle
LGo	Placed in the projection of a virtual bisector of a line adjacent to the left mandibular base and left posterior border of mandible
LCo	Placed at the most superior and central point of left condyle
LL6MB	Placed at the mesial buccal cusp of the mandibular left permanent first molars
LL6DB	Placed at the distal buccal cusp of the mandibular left permanent first molars
LL6R	Placed at the center of the pulp chamber floor of the mandibular right permanent first molars
LR1O	Placed in the middle of the incisal edge of the mandibular right permanent central incisor

that predict the best action to take at any given step. A Densely connected convolution network (DenseNet) was used.²¹ The FeatNet is made of convolution layers that are trained to capture the different image features. It takes as input the agent's state and outputs a vector describing relevant image features. This vector is then fed into the fully connected dense layers that output a probability vector ($P \in \mathbb{R}^6$) of the best movement to reach the final landmark position. The agent moves following the highest probability.

2.3 | Training the agents

Our data was split by scans, 70% for training, 10% for validation and 20% for testing. An environment was generated for each scan, and the position of corresponding landmarks was loaded. One agent was created for each landmark, and their network weights were initialized using a Xavier uniform function. For minimizing the distance between the agent and the landmark, each agent was trained using

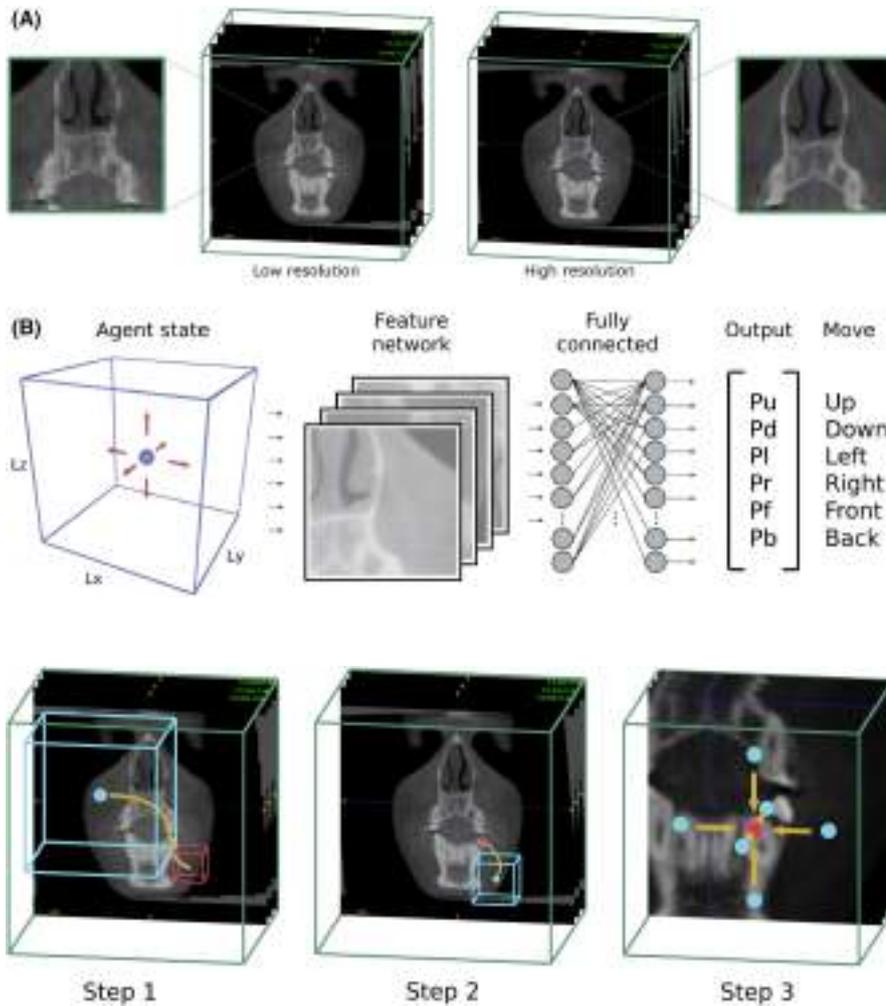


FIGURE 1 A, Visualization of an environment. On the left, the low 1 mm resolution was re-scaled from the high-resolution 0.3 mm scan on the right. B, Visualization of the agent's $Lx \times Ly \times Lz$ field of view (blue box), and the 6 possible moves (red arrows) after the network prediction.

FIGURE 2 Visualization of the agent (blue) in the multi-scale environment (green) searching for the target (red).

a combination of a state and the best action to take from the 6 possible movements described before.

The high-resolution and low-resolution scans had an average size of $180 \times 180 \times 180$ and $600 \times 600 \times 600$ voxels, respectively. It means that for each environment, we had more than 200 000 000 possible states that could be used to train the agent. However, the higher the number of states an agent needs to be trained, the higher would be the memory usage and the training time. We used the following strategy to generate outputs for each agent and limit the memory usage and training time:

- At the low-resolution level, we initialized K random position with a 20% chance to be within a radius R_{low} of a ground truth landmark (a region where more precision is needed). The remaining 80% could be anywhere in the scan. The agent is supposed to find the landmark from any starting point at this level.
- At the high-resolution level, we initialized K random position within a radius R_{high} of a ground truth landmark, knowing that the agent should be in this radius after the search at the low resolution.

The K positions for both resolution level were generated evenly in the N environments selected for the training. At every training epoch,

we updated 50% of the K positions with new randomly selected ones. This is one of the most important parts of our training strategy because it allowed the agent to be trained in most of the scan regions while reducing memory usage. The agent had a different network for each scale. These networks were optimized using the PyTorch library using a combination of algorithms and optimizer. The training was done on an NVIDIA Quadro RTX 6000/8000 GPU with a batch size of 100, $Lx = Ly = Lz = 64$, $K = 10\,000$, $N = 2$, and $R_{low} = R_{high} = 30$ voxels. It took about 4 h for one agent to be trained and reach a good accuracy.

2.4 | Prediction of the landmarks' position

To predict the landmark positions in a CBCT, we rescaled it to the resolutions used during training. The landmark location is predicted through the following steps (Figure 2):

- **Step 1:** The prediction begins at the low-resolution level. The agent is placed in the middle of the scan to optimize the search time. Once the agent reaches a confident zone, it goes to the high-resolution layer.
- **Step 2:** The agent moves in the high-resolution layer until it sets a preliminary estimation of the landmark location.

- **Step 3:** Now, a verification step is applied. This step consists of another search in the high-resolution layer starting from the 6 possible positions in a small radius from the predicted location in Step 2. The final result is an average of the 6 predicted positions.

During landmark position prediction, the stopping criteria is active and was implemented using a visitation map. The agent stops if it tries to reach a previously visited voxel. Fiducial lists are generated with the predicted positions of the landmarks and saved as JSON files.

After the initial training and validation, the agents were trained to predict a list of 119 landmark located in the cranial base, maxilla, mandible and dental structures commonly used for quantification of skeletal and dental changes in clinical studies (Table S1).

2.5 | Statistical analysis

To assess the prediction accuracy, the distance between each landmark in the ground truth fiducial list and the predicted one was computed by using the root mean square error.^{22,23} A 5-folds cross-validation was performed. For each landmark group, the placement errors and percentage of fails are presented. The error is the distance of the predicted landmark to the ground truth in mm and the distribution of the prediction error for each landmark was tested. A prediction was considered failed when the agent did not find the landmark or when the error was greater than 5 mm.

3 | RESULTS

The results are summarized in Table 2 that shows the errors (in mm) and fails (in %) for each landmark group. An average error of 1.54 mm was found for the landmark's prediction. A prediction is considered a failure when the agent did not reach the ground truth landmark region. Most of the landmarks have a 0% fail rate. Figure 3 shows the distribution of the prediction error (in mm) for each landmark. Only landmarks that presented percentages of failures are shown in Figure 3.

It took approximately 4.2 second on GPU for each landmark prediction. The prediction on the testing scans required 8.8 GB of cache memory and 2.1 GB of GPU memory. Each agent did 90 moves on average to reach the landmark position using a DenseNet.²⁴ In addition, new agents were trained to locate an extended list of 119

landmarks in different craniofacial structures. Adequate landmark identification was observed with the final trained model with 119 landmarks (Figure 4 and Table S1).

4 | DISCUSSION

This study presents a novel method for robust and accurate anatomical landmarks localization for 3D medical imaging data. The addition of dental records into healthcare data ecosystems and infrastructure is challenging, time-consuming, and dependent on clinician expertise. To leverage unstructured information in imaging data, we proposed and validated a method for automatic landmark identification in CBCT targeting clinical applications for dental, oral, and craniofacial clinical conditions that require quantitative landmark-based phenotyping. Previous studies using CBCT scans have demonstrated that manual landmark placement is a precise but time-consuming process.^{25,26} Landmark placement using both surface models and MPR images took an average of 10:41 ± 4:01 minutes to trace each patient.²⁵ In this study, the proposed open-source method combined the concept of scanning-based systems with smart displacement inside the scan using an agent. The training on a multi-resolution image enabled the artificial agent to systematically learn to find the targeted anatomical structures. The behaviour classification was solved using imitation learning, as this approach is easier to implement and train. It allows the use of deeper neural networks that encode a wider range of image features. The average automatic detection speed of 4.2 second for landmark was adequate considering the size of the CBCT volumes used.

Our results showed that this novel approach is robust and finds landmarks in CBCT scans accurately and automatically. An average error of 1.54 ± 0.87 mm was found for the assessed landmarks. This error is below the clinician's average error limit of 2 mm.²⁷⁻²⁹ In contrast, a previous study that tested an automatic landmark identification tool for CBCT showed a mean error distance of 3.19 ± 2.6 mm.² Cranial base landmarks showed better accuracy than the maxillary and mandibular landmarks (Table 2 and Figure 3). Also, a smaller failure rate was found for the cranial base landmarks that included only landmarks placed in the midsagittal plane (Table 2). Previous studies with manual landmark identification have shown that variables located in edges, crests or apices and between structures with different densities were easier to identify and therefore can present higher levels of accuracy.²⁹ Conversely, landmarks located on flat surfaces, curved bone structures, areas of low density, neighbouring areas of two dense structures or dental restorations are subjected to greater levels of error.²⁹ This can explain the smaller accuracy found for some dental landmarks, which also includes greater individual variation in tooth position that may require additional training. Additionally, 4 dental landmarks in the present study had errors greater than 5 mm in 25% of the cases, probably due to crowns and restoration artefacts or impacted/ectopic teeth in the CBCTs scans. This is

TABLE 2 Cross-validation prediction accuracy.

Bone group	Mean error ± SD	Maximum error	Fail percentage (%)
Maxillary	1.53 ± 0.85	4.83	4.7
Mandibular	1.61 ± 0.93	4.92	8.3
Cranial base	1.22 ± 0.51	2.29	2.7
All	1.54 ± 0.87	4.92	6.1

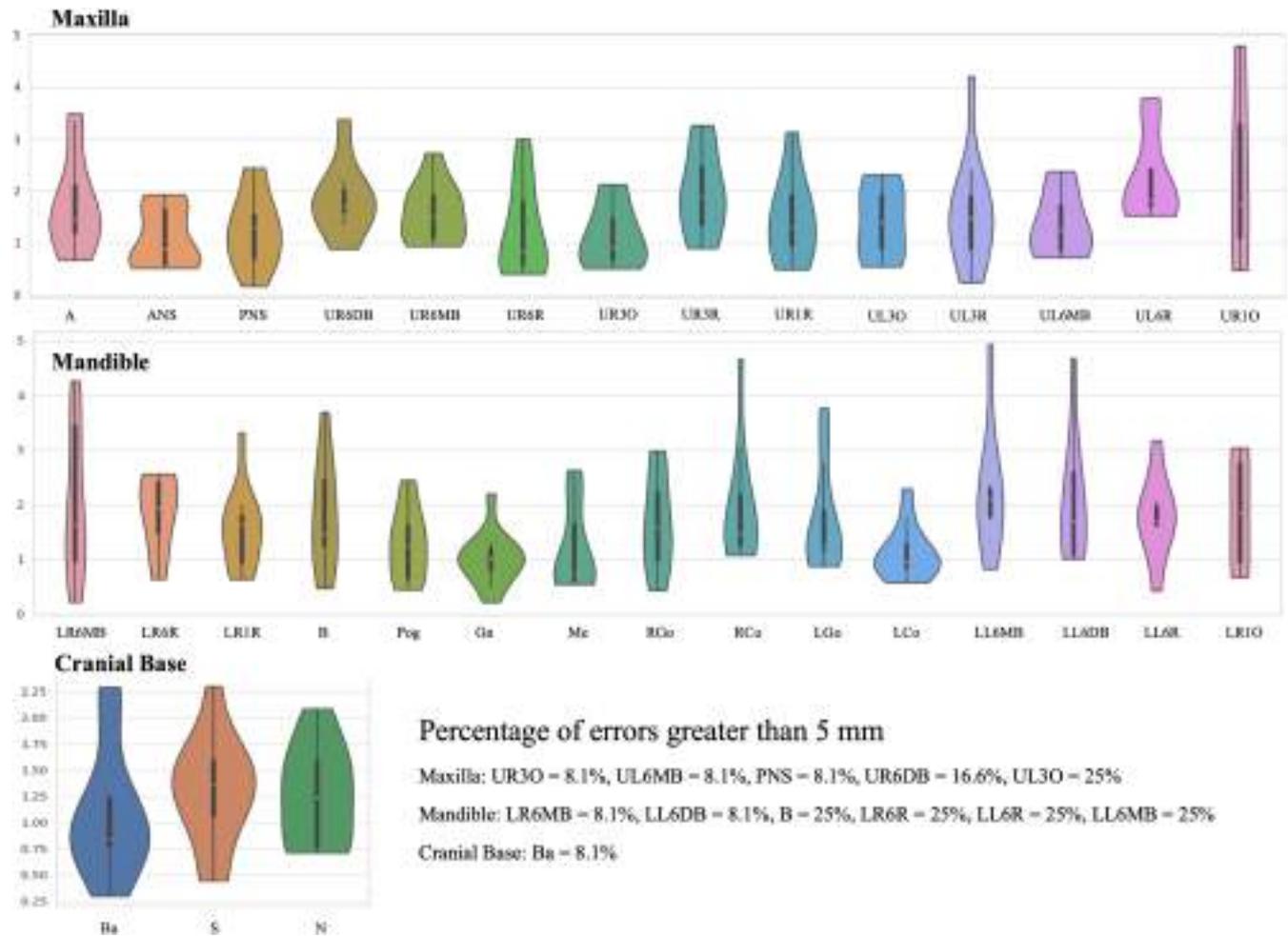


FIGURE 3 Violin plot of the cross-validation result on the cranial base (top) maxilla (middle) and mandible (bottom) and a summary of the fails (top right). Each landmark is represented with its error distribution in mm. The white dot and the black strip are respectively the mean error and the standard deviation.

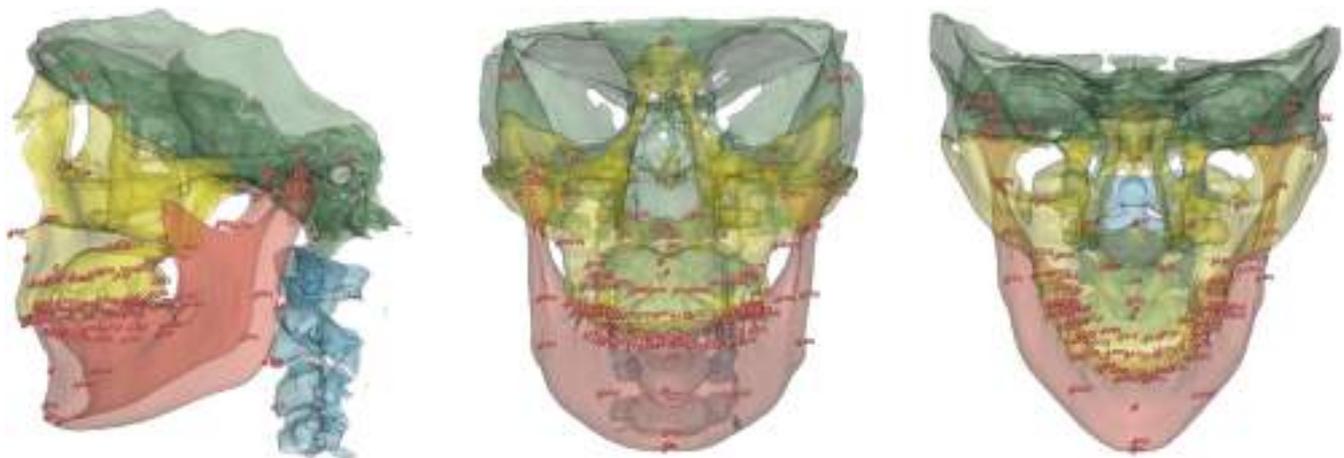


FIGURE 4 All 119 landmark that can be identify by the latest trained agents.

in accordance with a previous study that found greater errors for automatic landmark identification in dental landmarks when compared to manual landmark placement using computed tomography scans.³⁰

Previous studies in 2D and 3D have demonstrated that landmark placement can present different level of errors in the three spatial dimensions (x, y and z).^{31,32} Landmarks placed on curved structures, such as Gonion, can present greater levels of errors during placement

and these errors can differ in the three spatial dimensions.³¹ In a 3D evaluation, a higher reliability along the X direction was found for Gonion, whereas this landmark presented poor reliability along the Y and Z directions.³² In the present study, the landmark placement accuracy was assessed using the root mean square error. The root mean square error is one of the top performance metric systems used to assess the precision of machine learning approaches.^{22,23} However, this method does not allow to assess the differences in error in the three spatial dimensions.

A limitation of this study is the small sample size. However, this is an open-source code and software, where the machine learning models can be continuously updated to better assist clinicians and researchers in this crucial but time-consuming task. After validation of the proposed method, the clinical application was extended to 119 landmarks (Figure 4) annotated by clinician experts. Twenty-seven additional large field of view scans were used for initial location training of 119 landmarks, commonly required for 3D quantification of skeletal and dental structures. Future studies are needed to train a more robust generation of agents with larger datasets towards refining, testing and improving landmark placement accuracy with decreased failures in landmark location.

Given the preliminary robustness and good timing performance, the algorithm has been deployed for clinical and research use in an open-source web-based clinical decision support system (<https://dsci.dent.umich.edu>), and in a user-friendly open-source 3D Slicer module, with the code and detailed read me files available in Github (<https://github.com/DCBIA-OrthoLab/SlicerAutomatedDentalTools>) and video tutorials posted in Youtube (<https://www.youtube.com/@DCBIA/playlists>). Our models are developed with the Pytorch* framework and monai† library which facilitates reusability of the code and continuous improvement of the models. The current robustness of ALICBCT tool still requires clinical adjustments/verification from the users after the AI prediction. Continuous training of the ALICBCT will help increment the performance of the agents in future versions. We train separate model for each landmark, once that having separate machine learning models for each agent allows the clinician to make custom lists of landmarks and facilitates the periodic retraining of new agents separately without compromising the previously trained models.

5 | CONCLUSION

The ALICBCT algorithm presented an adequate level of accuracy in automatic landmark placement in CBCT scans. The precision and performance of this novel automated tool make it an important contribution to 3D imaging analysis in clinical and research studies. ALICBCT's open-source code and machine learning models offer the capability of continuous retraining with additional datasets to improve its performance. We expect to continue adding landmarks for future studies that require automated measurement and/or diagnostics.

AUTHOR CONTRIBUTION

MG (first author): conceptualization, formal analysis, methodology, writing original draft. FM: writing original draft, investigation and resources. BB, LA, NH: software engineering, validation, investigation, resources, original draft review and editing. AR, MG (fifth author), NA, EB, MY, JB: investigation, resources, original draft review and editing. BP, JCFR, DA: software engineering, original draft review and editing. LC: conceptualization, formal analysis, resources, original draft preparation and editing, supervision, project administration, funding acquisition. JCP: conceptualization, formal analysis, resources, original draft preparation and editing, supervision.

ACKNOWLEDGEMENTS

Supported by NIDCR R01 024450 and AAOF - Graber Family Teaching and Research Award - Orthodontic Faculty Development Fellowship.

CONFLICT OF INTEREST STATEMENT

The authors have no conflict of interest to declare.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available upon reasonable request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

ORCID

Felicia Miranda  <https://orcid.org/0000-0002-4015-0623>
 Marcela Gurgel  <https://orcid.org/0000-0001-9978-0542>
 Najla Al Turkestani  <https://orcid.org/0000-0002-7650-3638>
 Marilia Yatabe  <https://orcid.org/0000-0002-8748-6714>
 Jonas Bianchi  <https://orcid.org/0000-0002-3749-0918>
 Lucia Cevidanes  <https://orcid.org/0000-0001-9786-2253>
 Juan Carlos Prieto  <https://orcid.org/0000-0002-3778-9098>

ENDNOTES

* <https://pytorch.org/>
 † <https://monai.io/>

REFERENCES

1. Daye D, Wiggins WF, Lungren MP, et al. Implementation of clinical artificial intelligence in radiology: who decides and how? *Radiology*. 2022;305(1):212151.
2. Ghowsi A, Hatcher D, Suh H, et al. Automated landmark identification on cone-beam computed tomography: accuracy and reliability. *Angle Orthod*. 2022;92(5):642-654.
3. Zhang J, Liu M, Le A, Gao Y, Shen D. Alzheimer's disease diagnosis using landmark-based features from longitudinal structural MR images. *IEEE J Biomed Health Inform*. 2017;21(6):1607-1616.
4. Yu JI, Kim JS, Park HC, et al. Evaluation of anatomical landmark position differences between respiration-gated MRI and four-dimensional CT for radiation therapy in patients with hepatocellular carcinoma. *Br J Radiol*. 2013;86(1021):20120221.
5. Yang L, Georgescu B, Zheng Y, Wang Y, Meer P, Comaniciu D. Prediction based collaborative trackers (PCT): a robust and accurate

- approach toward 3D medical object tracking. *IEEE Trans Med Imaging*. 2011;30(11):1921-1932.
6. Cebra JR, Lohner R. Efficient simulation of blood flow past complex endovascular devices using an adaptive embedding technique. *IEEE Trans Med Imaging*. 2005;24(4):468-476.
 7. Pouch AM, Wang H, Takabe M, et al. Fully automatic segmentation of the mitral leaflets in 3D transesophageal echocardiographic images using multi-atlas joint label fusion and deformable medial modeling. *Med Image Anal*. 2014;18(1):118-129.
 8. Glocker B, Zikic D, Haynor DR. Robust registration of longitudinal spine CT. *Med Image Comput Comput Assist Interv*. 2014;17(Pt 1):251-258.
 9. Lüthi M, Jud C, Vetter T. Using landmarks as a deformation prior for hybrid image registration. In: Mester R, Felsberg M, eds. *Pattern Recognition. DAGM*. Vol 6835. Springer; 2011. Lecture Notes in Computer Science.
 10. Cuingnet R, Prevost R, Lesage D, Cohen LD, Mory B, Ardon R. Automatic detection and segmentation of kidneys in 3D CT images using random forests. *Med Image Comput Comput Assist Interv*. 2012;15:66-74.
 11. Donner R, Menze BH, Bischof H, Langs G. Global localization of 3D anatomical structures by pre-filtered Hough forests and discrete optimization. *Med Image Anal*. 2013;17(8):1304-1314.
 12. Ghesu FC, Krubasik E, Georgescu B, et al. Marginal space deep learning: efficient architecture for volumetric image parsing. *IEEE Trans Med Imaging*. 2016;35(5):1217-1228.
 13. Criminisi A, Shotton J, Robertson D, Konukoglu E. Regression forests for efficient anatomy detection and localization in CT studies. In: Menze B, Langs G, Tu Z, Criminisi A, eds. *Medical Computer Vision. Recognition Techniques and Applications in Medical Imaging. MCV 2010*. Lecture Notes in Computer Science. Vol 6533. Springer; 2011.
 14. Štern D, Ebner T, Urschler M. From local to global random regression forests: exploring anatomical landmark localization. In: Ourselin S, Joskowicz L, Sabuncu M, Unal G, Wells W, eds. *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2016*. Lecture Notes in Computer Science. Vol 9901. Springer; 2016:221-229.
 15. Dai J, Li Y, He K, Sun J. R-FCN: Object detection via region-based fully convolutional networks. *Advances in Neural Information Processing Systems NIPS*. 2016; 379-387.
 16. Payer C, Štern D, Bischof H, Urschler M. Regressing heatmaps for multiple landmark localization using CNNs. In: Ourselin S, Joskowicz L, Sabuncu M, Unal G, Wells W, eds. *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2016*. Lecture Notes in Computer Science. Vol 9901. Springer; 2016:230-238.
 17. Yushkevich PA, Piven J, Hazlett HC, et al. User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability. *Neuroimage*. 2006;31(3):1116-1128.
 18. Fedorov A, Beichel R, Kalpathy-Cramer J, et al. 3D slicer as an image computing platform for the quantitative imaging network. *Magn Reson Imaging*. 2012;30(9):1323-1341.
 19. Ruellas AC, Tonello C, Gomes LR, et al. Common 3-dimensional coordinate system for assessment of directional changes. *Am J Orthod Dentofacial Orthop*. 2016;149(5):645-656.
 20. Ghesu FC, Georgescu B, Zheng Y, et al. Multi-scale deep reinforcement learning for real-time 3D-landmark detection in CT scans. *IEEE Trans Pattern Anal Mach Intell*. 2019;41(1):176-189.
 21. Huang G, Liu Z, Van Der Maaten L, Weinberger KQ. Densely connected convolutional networks. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017:4700-4708.
 22. Lang Y, Lian C, Xiao D, et al. Automatic localization of landmarks in craniomaxillofacial CBCT images using a local attention-based graph convolution network. *Med Image Comput Comput Assist Interv*. 2020;12264:817-826.
 23. Botchkarev A. A new typology design of performance metrics to measure errors in machine learning regression algorithms. *Interdiscip J Inf Knowl Manag*. 2019;14:45-76.
 24. Boroumand M, Chen M, Fridrich J. Deep residual network for Steganalysis of digital images. *IEEE Trans Inf Forens Sec*. 2019;14(5):1181-1193.
 25. Hassan B, Nijkamp P, Verheij H, et al. Precision of identifying cephalometric landmarks with cone beam computed tomography in vivo. *Eur J Orthod*. 2013;35(1):38-44.
 26. Ludlow JB, Gubler M, Cevidanes L, Mol A. Precision of cephalometric landmark identification: cone-beam computed tomography vs conventional cephalometric views. *Am J Orthod Dentofacial Orthop*. 2009;136(3):312e1-10.
 27. Kragkov J, Bosch C, Gyldensted C, Sindet-Pedersen S. Comparison of the reliability of craniofacial anatomic landmarks based on cephalometric radiographs and three-dimensional CT scans. *Cleft Palate Craniofac J*. 1997;34(2):111-116.
 28. Periago DR, Scarfe WC, Moshiri M, Scheetz JP, Silveira AM, Farman AG. Linear accuracy and reliability of cone beam CT derived 3-dimensional images constructed using an orthodontic volumetric rendering program. *Angle Orthod*. 2008;78(3):387-395.
 29. Lagravere MO, Low C, Flores-Mir C, et al. Intraexaminer and interexaminer reliabilities of landmark identification on digitized lateral cephalograms and formatted 3-dimensional cone-beam computerized tomography images. *Am J Orthod Dentofacial Orthop*. 2010;137(5):598-604.
 30. Dot G, Schouman T, Chang S, et al. Automatic 3-dimensional cephalometric landmarking via deep learning. *J Dent Res*. 2022;101(11):1380-1387.
 31. Baumrind S, Frantz RC. The reliability of head film measurements. 1. Landmark identification. *Am J Orthod*. 1971;60(2):111-127.
 32. Cattaneo PM, Yung AKC, Holm A, Mashaly OM, Cornelis MA. 3D landmarks of craniofacial imaging and subsequent considerations on superimpositions in orthodontics-the Aarhus perspective. *Orthod Craniofac Res*. 2019;22(Suppl 1):21-29.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Gillot M, Miranda F, Baquero B, et al. Automatic landmark identification in cone-beam computed tomography. *Orthod Craniofac Res*. 2023;00:1-8. doi:[10.1111/ocr.12642](https://doi.org/10.1111/ocr.12642)

RESEARCH ARTICLE

Automatic multi-anatomical skull structure segmentation of cone-beam computed tomography scans using 3D UNETR

Maxime Gillot^{1,4*}, Baptiste Baquero^{1,4}, Celia Le^{1,4}, Romain Deleat-Besson^{1,4}, Jonas Bianchi¹, Antonio Ruellas¹, Marcela Gurgel¹, Marilia Yatabe¹, Najla Al Turkestani^{1,10}, Kayvan Najarian¹, Reza Soroushmehr¹, Steve Pieper², Ron Kikinis³, Beatriz Paniagua⁵, Jonathan Gryak¹, Marcos Ioshida¹, Camila Massaro¹, Lilliane Gomes¹, Heesoo Oh⁶, Karine Evangelista¹, Cauby Maia Chaves Junior⁷, Daniela Garib⁸, Fábio Costa¹, Erika Benavides¹, Fabiana Soki¹, Jean-Christophe Fillion-Robin⁵, Hina Joshi⁹, Lucia Cevidanes¹, Juan Carlos Prieto⁹

1 University of Michigan, Ann Arbor, Michigan, United States of America, **2** ISOMICS, Cambridge, Massachusetts, United States of America, **3** Harvard Medical School, Boston, Massachusetts, United States of America, **4** CPE Lyon, Lyon, France, **5** Kitware, Clifton Park, New York, United States of America, **6** University of Pacific, Stockton, California, United States of America, **7** University of Ceara, Fortaleza, Brazil, **8** University of São Paulo, São Paulo, Brazil, **9** University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, United States of America, **10** King Abdulaziz University, Jeddah, Saudi Arabia

* maxime.gillot@cpe.fr



OPEN ACCESS

Citation: Gillot M, Baquero B, Le C, Deleat-Besson R, Bianchi J, Ruellas A, et al. (2022) Automatic multi-anatomical skull structure segmentation of cone-beam computed tomography scans using 3D UNETR. *PLoS ONE* 17(10): e0275033. <https://doi.org/10.1371/journal.pone.0275033>

Editor: Sathishkumar V E, Hanyang University, KOREA, REPUBLIC OF

Received: May 10, 2022

Accepted: September 9, 2022

Published: October 12, 2022

Peer Review History: PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: <https://doi.org/10.1371/journal.pone.0275033>

Copyright: © 2022 Gillot et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: We have made all the code available at "https://github.com/Maxlo24/AMASSS_CBCT" that can be used to replicate this study. The database cannot be shared publicly

Abstract

The segmentation of medical and dental images is a fundamental step in automated clinical decision support systems. It supports the entire clinical workflow from diagnosis, therapy planning, intervention, and follow-up. In this paper, we propose a novel tool to accurately process a full-face segmentation in about 5 minutes that would otherwise require an average of 7h of manual work by experienced clinicians. This work focuses on the integration of the state-of-the-art UNETR Transformers (UNETR) of the Medical Open Network for Artificial Intelligence (MONAI) framework. We trained and tested our models using 618 de-identified Cone-Beam Computed Tomography (CBCT) volumetric images of the head acquired with several parameters from different centers for a generalized clinical application. Our results on a 5-fold cross-validation showed high accuracy and robustness with a Dice score up to 0.962±0.02. Our code is available on our [public GitHub repository](#).

1 Introduction

Segmentation of medical and dental images is a visual task that aims to identify the voxels of organs or lesions from background grey-level scans. It represents a prerequisite for medical image analysis and supports entire clinical workflows from computer-aided diagnosis [1] to therapy planning [2], intervention [3], and follow-up [4]. Particularly for challenging dental and craniofacial conditions, such as dentofacial deformities, craniofacial anomalies, and tooth impaction, quantitative image analysis requires efficient solutions to solve the time-consuming

because the scans contain the patient facial skin that could allow facial recognition. Data are available from the University of Michigan Institutional Review Board (Contact Robert Eber, email reber@umich.edu) for researchers who meet the criteria for access to confidential data.

Funding: This work was supported by NIDCR R01 024450, American Association of Orthodontists Foundation Grabber Family Teaching and Research Award and by Research Enhancement Award Activity 141 from the University of the Pacific, Arthur A. Dugoni School of Dentistry. The funders had no role in the study design, data collection and analysis, decision to publish or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

and user-dependent task of image segmentation. With medical and dental images being acquired at multiple scales and/or with multiple imaging modalities, automated image analysis techniques are needed to integrate patient data across scales of observation.

Due to the low signal/noise ratio of Cone-Beam CT (CBCT) images used in Dentistry, the current open-source tools for anatomic segmentation, such as ITK-SNAP [5] and 3D-Slicer [6] are challenging for clinicians and researchers. The large field of view CBCT images commonly used for Orthodontics and Oral Maxillofacial Surgery clinical applications require on average to perform detailed segmentation by experienced clinicians: (Fig 1) 7 hours of work for full face, 1.5h for the mandible, 2h for the maxilla, 2h for the cranial base (CB), 1h for the cervical vertebra (CV), and 30min for the skin. Additional challenges for accurate and robust automatic anatomical segmentation are the rich variety of anatomical structures morphology and the differences in imaging acquisition protocols and scanners from one center to another. Furthermore, patients that present with facial bone defects pose additional challenges for automatic segmentation because of unexpected anatomical abnormalities and variability. For this reason, the training of the machine learning models in the present study also included gold standard (ground-truth) clinicians' expert segmentations of CBCT images from patients with craniofacial large bone defects such as cleft lip and palate (CLP). Being able to accurately segment those maxillary deformities (Fig 1) is for the diagnosis and treatment planning of correction of the bone defects and craniomaxillofacial anomalies.

Although in the last decades, automatic approaches such as region seed growing [7], clustering methods, random forests [8], atlas-based system [9], and deep convolutional neural network (CNN) [10] have been proposed to segment the mandible, the maxilla, and the teeth, CBCT image segmentation remains challenging. Those previous studies focused on small samples from a single acquisition protocol; however, scans acquired at different clinical centers with different acquisition protocols, scales, and orientations require laborious manual



Fig 1. Multi-anatomical skull structure manual segmentation of the full-face by combining the mandible, the maxilla, the cranial base, the cervical vertebra, and the skin segmentation. Patient has written consent on file for the use of the images.

<https://doi.org/10.1371/journal.pone.0275033.g001>

correction in clinical settings to achieve accurate segmentation. Hence, methods for generalizable automatic image segmentation are sought.

The present study objective is to offer a free open-source tool to facilitate medical and dental image segmentation for clinics and research. We focused on the best practices for Artificial Intelligence in healthcare imaging across academia and enterprise researchers. Hence, the use of the new Medical Open Network for Artificial Intelligence (MONAI) framework that implements state-of-the-art machine learning algorithms such as the UNet Transformers (UNETR) [11]. In the following sections, we describe the data used to train our machine learning models, followed by related work on approaches to segment medical images, testing the performance of the proposed methods compared to the clinician's expert segmentation, and discussion of the novel results.

2 Materials

A total of 618 DICOM-formatted CBCT images of the head were used in this work. The images were acquired from 7 clinical centers with various scanners, image acquisition protocols, and field of views. All patient HIPAA identifiable personal information was removed from the DICOM files metadata through an anonymization process in the 3D Slicer platform [6]. The anonymization was performed before the clinical centers shared the data for this retrospective study. The University of Michigan Institutional Review Board HUM00217585 waived the requirement for informed consent and granted IRB exemption. Patients' skin was not removed from the large field of scans; however, those files are used only for the training of the proposed machine learning model.

Two open-source software packages, ITK-SNAP 3.8 [5] and 3D Slicer 4.11 were used by clinical experts to perform user interactive manual segmentation of the volumetric images and common spatial orientation of the head as the ground-truth to train our deep learning models.

All the 618 scans don't come with a full-face segmentation, the dataset was composed of 446 patients with mandible segmentation, 132 with the maxilla, 116 of the cranial base, 80 with the skin, and 14 patients with the cervical vertebra. The image spatial resolution varied from 0.16 to 0.5 mm³ voxels. To test the robustness of the proposed method, patients with CLP were included in the dataset. Those patients have large bone defects in the jaw that varies a lot from one patient to another.

3 Related work

3.1 Region seed growing [7]

This method needs to place the seed inside the region of interest. The grayscale intensity grid and spatial distances from the seed to all the other voxels are computed to estimate a segmentation of similar features. This method showed less accuracy than the following methods and can require the clinicians to place the seeds.

3.2 Atlas-based system [9]

An atlas is defined as the combination of an intensity image and its segmentation to generate a template. From this point, 2 steps occur: label transfer which transfers segmentation labels from pre-labeled atlases to a novel image and label fusion which combines the label transfer results. The main con of this method is the lack of flexibility when exposed to high changes in the data such as in patients with CLP.

3.3 Random forests [8]

A probability grid is made to estimate the initial segmentation based on multiple expert-segmented CBCT images. The appearance features from CBCTs and the context features from the initial probability maps are both extracted to train a first-layer of random forest classifiers. A sequence of classifiers can segment CBCT images by iteratively training the subsequent random forest classifier using both the original CBCT features and the updated segmentation probability maps. Those methods are slow to train, computing-intensive and the prediction time can be high.

3.4 CNN

Previous methods were mostly using 2D [12] or 2.5D UNet [13], limited by computer power. Recent progress in GPU power and network architecture allowed the appearance of 3D CNN architectures showing better results than their 2/2.5D analogs. 3D UNet [10], TransUNet [14], and nnU-Net [15] showed high performance for medical imaging tasks including segmentation. However, the new UNETR architecture showed better results than all the previously cited CNN for CT segmentation.

4 Proposed method

Thanks to recent advances in deep learning, this study proposes a convolutional neural network (CNN) to extract a hierarchical feature representation for segmentation, which is robust to image degradation such as noise, blur, and contrast. Our algorithm requires Python 3.9 and uses various libraries to perform image processing. For the post-processing and the pre-processing, we are using ITK, SimpleITK, VTK, and connected-components-3d libraries. For the data augmentation and the segmentation, we used the MONAI library which simplifies the UNETR implementation and is optimized to process medical images in Python.

4.1 Pre-processing

Depending on the scanner and the image acquisition protocol, the CBCT scans are gray-scaled images with high contrast variation from one patient to another and the image spacing can be different. Among all the different spacing, 0.4 mm^3 is the most frequent. It's also a resolution that keeps enough details of the skull structure to segment while limiting memory usage with reasonable image size. From one center to another, the manual segmentation method can change. Different labels are used and the skull structure can be filled or not. From this point, to have more consistency in the dataset, all the data go through the following pre-processing steps:

- All the CBCTs and segmentations are re-sampled with a 0.4-mm^3 isometric voxel size using respectively a linear and a nearest-neighbor interpolation function.
- The scans go through a contrast adjustment function Fig 2. A cumulative graph is made from the image histogram ignoring the background color. The new minimum and maximum intensity are selected when reaching an X_{min} and X_{max} percentage on the cumulative graph. The intensity is then re-scaled in the $[0, 1]$ interval.
- A “fill hole” morphological operation is applied to the segmentation and the label is set at 1.

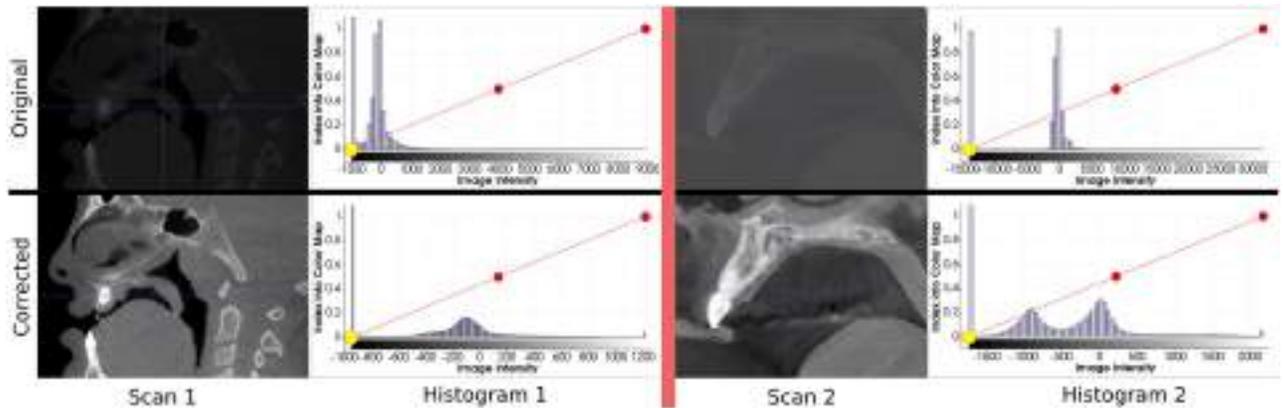


Fig 2. Visualization of the contrast adjustment steps on two different scans. This result is obtained by keeping the data between $X_{min} = 1\%$ and $X_{max} = 99\%$ on the cumulative graph.

<https://doi.org/10.1371/journal.pone.0275033.g002>

4.2 UNETR

For this machine learning tool, we decided to use the new state-of-the-art model in 3D scan segmentation, the UNETR. Its architecture utilizes a transformer as the encoder to learn sequence representations of the input volume and effectively capture the global multi-scale information. The network design follows the successful “U-shape” for the encoder and decoder. The transformer encoder is directly connected to a decoder via skip connections at different resolutions to compute the final semantic segmentation output. The size of the scans to segment is not consistent and tends to be large (up to 600x600x600 voxels). No GPU is powerful enough to take this voxel grid size as input. We decided to shape our UNETR classifier with a 128x128x128 voxels input (Fig 3). To segment the entire image, the classifier moves across the scan to perform predictions in different locations. Once the entire image has been processed, segmented crops are merged to match the original input image size. Individual UNETR models were trained for different segmentation needs. All the models share common

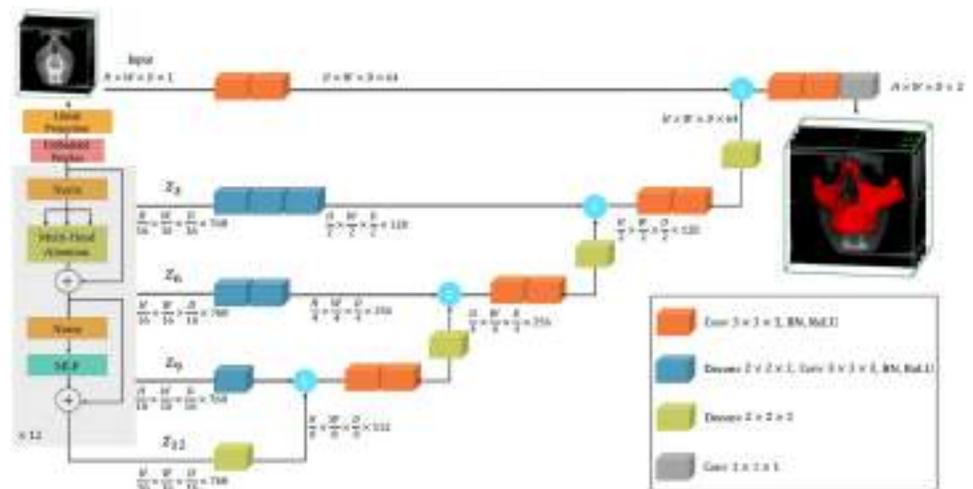


Fig 3. Overview of the UNETR used. A 128x128x128x1 cropped volume of the input CBCT is divided into a sequence of 16 patches and projected into an embedding space using a linear layer. A transformer model is fed with the sequence added with 768 position embedding. Via skip connections, the decoder will extract and merge the final 128x128x128x2 crop segmentation from the encoded representations of different layers in the transformer.

<https://doi.org/10.1371/journal.pone.0275033.g003>

parameters: feature size = 16, hidden layer = 768, feedforward layer = 3072, number of attention heads = 12, and a dropout rate of 5%.

4.3 Training

For each skull structure to segment, the patients were sorted by separated folders based on the clinical center they were coming from. The dataset was then split into 3: 70% for the training, 10% for the validation, and 20% for testing. The data was split evenly from each folder to avoid overfitting to any specific center.

We used the MONAI “CacheDataset” tool to load the pre-processed data. Those datasets allow the use of transformers for data augmentations. Every time an image and its segmentation are loaded for the training, a number of N_s cube samples are randomly cropped in the voxel grid. Those cubes all have the same $L_x \times L_y \times L_z$ shape to match the UNETR input size. For data augmentation (Table 1), random flip and 90° rotation are applied in each direction along with a random shift in intensity and contrast for the scans.

This step is applied to N_i images to generate a batch of size $N_i \times N_s$. This batch is then fed into the UNETR the training. For the validation, data augmentation is also applied by only ignoring the cropping step, a prediction occurs on the full image using MONAI sliding window inference to move the UNETR classifier across the image. This network is optimized using the PyTorch library by a combination of a back-propagation algorithm to compute the network gradients and the Adam optimizer with weight decay. In this work, we used the weighted average of both the Dice loss (Table 1) and Cross Entropy Loss (Table 2) function.

$$DL = 1 - \frac{2 \sum_{i=1}^N p_i g_i}{\sum_{i=1}^N p_i^2 + \sum_{i=1}^N g_i^2}, \tag{1}$$

where $p_i \in P$ is the predicted probability of the i -th voxel and $g_i \in G$ is the ground truth of the i -th voxel.

$$\ell(x, y) = L = \{l_1, \dots, l_N\}^T, \quad l_n = -w_{y_n} \log \frac{\exp(x_{n,y_n})}{\sum_{c=1}^C \exp(x_{n,c})}, \tag{2}$$

Where x is the input, y is the target, w is the weight, C is the number of classes, and N spans the minibatch dimension as well as d_1, \dots, d_k for the K -dimensional case.

Table 1. Data augmentation transformations for the training.

Data	Random crop	Random flip and rotation	Random shift in intensity	Random contrast adjustment
Images	Anywhere in the scan N_s times	Along X, Y and Z-axis with a 25% probability for each axis for each axis	50% chances of a 0.1 intensity shift	80% chances to change image gamma in a [0.5,2] interval
Segmentation			N/A	N/A

<https://doi.org/10.1371/journal.pone.0275033.t001>

Table 2. Comparison of manual and automatic segmentation using AUPRC, AUPRC-Baseline, Dice, F2 Score, Accuracy, Recall, and Precision of the 5-fold cross-validation for the 5 skull structures segmentation.

Structure	AUPRC	AUPRC Baseline	Dice	F2 Score	Accuracy	Recall	Precision
Mandible	0.926 ± 0.037	0.011 ± 0.003	0.962 ± 0.020	0.961 ± 0.026	0.9992 ± 0.0005	0.960 ± 0.031	0.965 ± 0.026
Maxilla	0.738 ± 0.096	0.011 ± 0.003	0.853 ± 0.064	0.857 ± 0.061	0.996 ± 0.001	0.862 ± 0.073	0.855 ± 0.099
Cranial base	0.642 ± 0.127	0.018 ± 0.006	0.788 ± 0.103	0.804 ± 0.109	0.992 ± 0.004	0.824 ± 0.099	0.774 ± 0.135
Cervical vertebra	0.602 ± 0.145	0.008 ± 0.006	0.760 ± 0.113	0.723 ± 0.164	0.995 ± 0.004	0.704 ± 0.192	0.854 ± 0.033
Skin	0.947 ± 0.035	0.425 ± 0.72	0.971 ± 0.018	0.982 ± 0.009	0.974 ± 0.018	0.989 ± 0.009	0.954 ± 0.037

<https://doi.org/10.1371/journal.pone.0275033.t002>

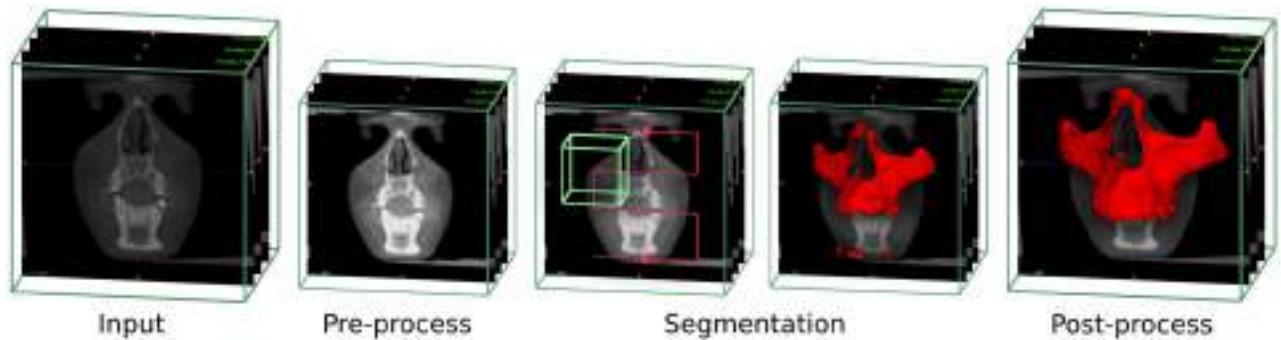


Fig 4. Visualization of the automatic maxilla segmentation steps. Re-sample and contrast adjustment of the input image, segmentation with the sliding window using UNETR, and finally, re-sampling of the cleaned-up segmentation to the input size.

<https://doi.org/10.1371/journal.pone.0275033.g004>

The training was done on an NVIDIA Quadro RTX 6000/8000 GPU. With $X_{min} = 1\%$, $X_{max} = 99\%$, $L_x = L_y = L_z = 128$, $N_i = N_s = 10$ (batch size of 100), a dropout rate of 0.05, a learning rate of 10^{-4} and a weight decay of 10^{-5} it takes around 4h and 22GB of GPU memory for one model to be trained.

4.4 Segmentation and post-processing

Once we have a trained model, the challenge is to segment new scans that possibly have a different contrast and spacing than the ones used for the training. For the prediction, we create a new temporary file to work on and preserve the original. We apply the 2 first pre-processing steps (re-sample in a 0.4mm^3 spacing if needed and adjust the contrast). The sliding window inference is then used to segment the whole image. We get as an output a voxel grid of probability on which we apply an argmax function. The segmentation can have some artifacts and unwanted elements. Therefore, we used the connected-components-3d 3.9.1 library [16] to keep the biggest segmented object only. A morphological operation is then applied to the segmentation to fill the holes. The final result is re-sampled to match the original image, orientation, spacing, origin, and size. All the steps are summarized in Fig 4.

5 Results

We performed a 5-fold cross-validation, each fold with a different 20% portion of the available data for the test. It allows testing the models on the entirety of the dataset.

The MONAI sliding window inference allows overlapping of the classifier for more precision but has a drastic impact on the computation time. During the validation step of the training, a prediction takes about 4s with 20% of overlap. To compute the metrics we used a 50% overlap to segment the test scans and it takes around 24s on GPU for each CBCT to be segmented. The prediction goes up to 1 min with an 80% overlap for even more precision.

To compare the clinician experts' manual segmentation and the AMASSS automatic segmentation, we used the Area Under the Precision-Recall Curve (AUPRC Eq 8) metric for class imbalance. Most of the bone groups represent about 10% of the volume only. Other metrics such as the recall (6), precision (7), Dice coefficient (DC Eq 3), and F2 (4) score were also computed to know how efficient the model is.

$$DC(M, A) = \frac{2|A \cap M|}{|A| + |M|}, \quad (3)$$

Where M and A are respectively the binary image of the ground truth segmentation and the

AMASSS output.

$$F_2 = \frac{TP}{TP + 0.2FP + 0.8FN}, \quad (4)$$

$$A = \frac{TP + TN}{TP + TN + FP + FN}, \quad (5)$$

$$R = \frac{TP}{TP + FN}, \quad (6)$$

$$P = \frac{TP}{TP + FP}, \quad (7)$$

Where TP stand for the number of true positive in the AMASSS output voxel grid, TN true negative, FP false positive and FN false negative.

$$AUPCR = \sum_{n=1}^{N-1} \frac{(R[n] - R[n - 1]) \times (P[n] - P[n - 1])}{2}, \quad (8)$$

Where R and P are the recall and the precision values from N confusion matrices for different thresholds. All these measurements (Table 2) vary from zero to one, where zero means no superposition between the two volumes, and one shows a perfect superposition between both. All metrics were performed on the binarized 3D images resulting from the post-processing. From a clinical point of view, it is better to have over-segmented images rather than under-segmented ones, and hence the F2 score was computed considering recall as twice as important as precision.

The average results for the mandible and the skin show the high precision of the automatic segmentations with a Dice above 0.96. Additionally, the standard deviation is quite low, indicating that the predictions are robust, consistent, and generalizable to unseen patients. Maxilla and cranial base showed similar results. The lower Dice compared to the mandible can be explained by fewer data used to train, but more importantly because of inconsistency from one ground-truth segmentation to another. The separation between the maxilla and the mandible can change, those regions have very thin bones and the amount of details segmented is different depending on the center. With only 14 segmentation available from one center, the cervical vertebra results are promising, showing the potential to be generalizable in future training with a larger sample.

We processed a full-face segmentation (Fig 5) of the patient Fig 1 that was kept out of all training. The CLP and even the cervical vertebra were successfully segmented, showing the robustness of the UNETR.

6 Discussion

This is the first study to our knowledge to use the new 3D UNETR architecture to segment multiple anatomic skeletal, dental, and soft tissue structures in the craniofacial complex of CBCT scans. Recent studies have focused on only one specific facial structure such as the maxilla, [17], mandible [18] or airway [12], and used smaller samples from a single CBCT acquisition protocol, thus, those algorithms are not yet generalizable like the proposed AMASSS.

Traditional image processing methods, such as super-voxels and graph clustering [19], atlas-based segmentation [8, 20], watershed methods [21] are available tools that presented good accuracy for segmentation, however, due to image artifacts and noise, that can be caused



Fig 5. Visualization of the automatic full-face segmentation results. In red, the prediction is superposed with the manual segmentation in transparent green. On the full-face, we can see that the models managed to average the separation line between the maxilla and the mandible. The separation on the manual segmentation is different. It also explains why the metrics are lower than the mandible for those two skull structures.

<https://doi.org/10.1371/journal.pone.0275033.g005>

by intercuspatation of the dentition and the presence of metallic crowns, it is still a challenge to segment the images properly and also to segment different tissues such as bone with different densities (boundaries) and soft tissues. Due to these limitations, machine learning methods for image segmentation in dentistry have become popular, and the major limitation in training AI models such as the proposed AMASS is to have a gold standard to serve as training models [22]. To overcome this limitation, in this study, manual annotations were performed for each scan used in the training, provided by clinicians with expertise and experience in 3D CBCT segmentations.

Moreover, AMASSS showed better and similar accuracy when compared to Si Chen et al.'s Maxilla segmentation with a dice score of 0.800 ± 0.029 and Verhelst et al.'s with a dice of 0.9722 ± 0.006 for the mandible segmentation, respectively. Commercial companies such as Materialise [23], Relu [24], and Diagnocat [25] have recently marketed AI-based segmentation for CBCT scans, but they are expensive and the precision of their algorithms require validation by clinicians.

Another important challenge in automated systems in dentistry, explained by Schwendicke et al. [26], is to provide solutions that can be largely entered into dental routine practice, and also follows principles such as demonstrating clinical value, protecting patient data, individual privacy, maintaining trustworthiness, and ensuring robustness and generalizability of the tools. Towards these goals, the proposed open-source AMASSS algorithm was deployed as a free 3D Slicer extension "Automated dental tools". The software interface allows users to select the most updated trained model for increased precision of anatomic structures segmentation, continuously updating toward improved identification of patient facial structures and clinical applications [27].

Regarding the advantages and limitations, this study has the capacity of performing the segmentation of multi-structures in approximately 5 minutes; however, to achieve the necessary precision the ground-truth data can take several hours to be manually produced by the clinicians, which makes the addition of new structures of interest challenging and still human-dependent. Also, automated tools such as AMASS focus on future clinical decision support systems, to improve the human-computer interface rather than interrupt the clinical workflow [28], and for this reason, human interaction is still required, but less time-consuming.

Future work will continue to increase the databases for cervical vertebra, maxilla, and cranial base as well as add detailed anatomic structures such as the teeth roots and mandibular canals segmentation. Additional potential applications may be generalizable to other imaging

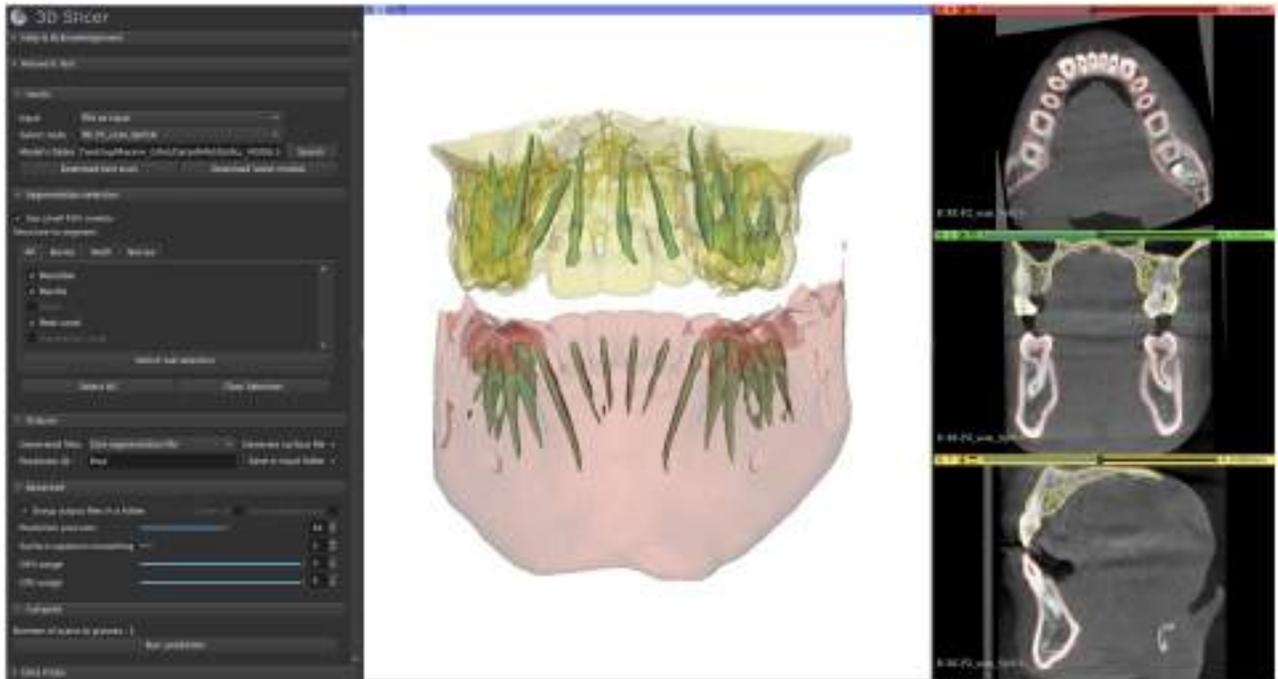


Fig 6. 3D Slicer module in development for AMASSS-CBCT. On the left, we can see the module with the different options/parameters. On the right, the visualisation of the segmentation applied on one small field of view scan with the selected skull structures. The mandible in red, the maxilla in yellow and the root canals in green.

<https://doi.org/10.1371/journal.pone.0275033.g006>

modalities such as Magnetic Resonance Imaging, CT, micro CT, and ultrasound, which is been shown in recent manuscripts in the medical field [29, 30].

7 Conclusion

This proposal is a step towards the implementation of dentistry decision support systems, as machine learning techniques are becoming important to automatically and efficiently analyze dental images. The MONAI framework facilitated the processing of 618 CBCTs to perform fast training and data augmentation, which led to the high accuracy and robustness of the AMASSS tool. The UNETR showed high overall performance, achieving a Dice up to 0.962 ± 0.02 on heterogeneous CBCT images.

Given its robustness and performance time, this validated free tool was implemented in 2 open-source ecosystems, a web-based clinical decision support system (the Data Storage for Computation and Integration, DSCI) [31], and a user-friendly 3D Slicer module Fig 6. These computer-aided diagnostic tools will aid in diagnosis and therapy planning, especially for patients with craniomaxillofacial anomalies and deformities.

Author Contributions

Conceptualization: Jonas Bianchi, Antonio Ruellas, Lucia Cevidanes.

Data curation: Maxime Gillot, Celia Le, Romain Deleat-Besson, Jonas Bianchi, Marcela Gurgel, Marilia Yatabe, Najla Al Turkestani, Jonathan Gryak, Marcos Ioshida, Camila Massaro, Liliane Gomes, Heesoo Oh, Karine Evangelista, Cauby Maia Chaves Junior, Daniela Garib, Fábio Costa, Erika Benavides, Fabiana Soki.

Formal analysis: Maxime Gillot, Antonio Ruellas, Marcela Gurgel, Marilia Yatabe, Najla Al Turkestani, Jonathan Gryak, Marcos Ioshida, Camila Massaro, Liliane Gomes, Heesoo Oh, Karine Evangelista, Cauby Maia Chaves Junior, Daniela Garib, Fábio Costa, Erika Benavides, Fabiana Soki, Lucia Cevidanes.

Funding acquisition: Lucia Cevidanes.

Investigation: Maxime Gillot, Antonio Ruellas, Marcela Gurgel, Marilia Yatabe, Najla Al Turkestani, Jonathan Gryak, Marcos Ioshida, Camila Massaro, Liliane Gomes, Heesoo Oh, Karine Evangelista, Cauby Maia Chaves Junior, Daniela Garib, Fábio Costa, Erika Benavides, Fabiana Soki, Lucia Cevidanes.

Methodology: Maxime Gillot, Antonio Ruellas, Lucia Cevidanes, Juan Carlos Prieto.

Project administration: Maxime Gillot, Antonio Ruellas, Lucia Cevidanes.

Software: Maxime Gillot, Baptiste Baquero, Steve Pieper, Ron Kikinis, Hina Joshi, Juan Carlos Prieto.

Visualization: Maxime Gillot.

Writing – original draft: Maxime Gillot.

Writing – review & editing: Baptiste Baquero, Celia Le, Romain Deleat-Besson, Jonas Bianchi, Antonio Ruellas, Marcela Gurgel, Marilia Yatabe, Najla Al Turkestani, Kayvan Najarian, Reza Soroushmehr, Steve Pieper, Ron Kikinis, Beatriz Paniagua, Jonathan Gryak, Marcos Ioshida, Camila Massaro, Liliane Gomes, Heesoo Oh, Karine Evangelista, Cauby Maia Chaves Junior, Daniela Garib, Fábio Costa, Erika Benavides, Fabiana Soki, Jean-Christophe Fillion-Robin, Hina Joshi, Lucia Cevidanes, Juan Carlos Prieto.

References

1. Way TW, Hadjiiski LM, Sahiner B, Chan HP, Cascade PN, Kazerooni EA, et al. Computer-aided diagnosis of pulmonary nodules on CT scans: Segmentation and classification using 3D active contours. *Medical physics*. 2006; 33(7Part1):2323–2337. <https://doi.org/10.1118/1.2207129> PMID: 16898434
2. D D'Haese PF, Duay V, Li R, du Bois d'Aische A, Merchant TE, Cmelak AJ, et al. Automatic segmentation of brain structures for radiation therapy planning. In: *Medical Imaging 2003: Image Processing*. vol. 5032. International Society for Optics and Photonics; 2003. p. 517–526.
3. Fang X, Xu S, Wood BJ, Yan P. Deep learning-based liver segmentation for fusion-guided intervention. *International journal of computer assisted radiology and surgery*. 2020; 15(6):963–972. <https://doi.org/10.1007/s11548-020-02147-6> PMID: 32314228
4. Gaviño Orduña J, García García M, Dominguez P, Caviedes Bucheli J, Martin Biedma B, Abella Sans F, et al. Successful pulp revascularization of an autotransplanted mature premolar with fragile fracture apicoectomy and plasma rich in growth factors: a 3-year follow-up. *International Endodontic Journal*. 2020; 53(3):421–433. <https://doi.org/10.1111/iej.13230> PMID: 31587320
5. Yushkevich PA, Piven J, Hazlett HC, Smith RG, Ho S, Gee JC, et al. User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability. *Neuroimage*. 2006; 31(3):1116–1128. <https://doi.org/10.1016/j.neuroimage.2006.01.015> PMID: 16545965
6. Fedorov A, Beichel R, Kalpathy-Cramer J, Finet J, Fillion-Robin JC, Pujol S, et al. 3D Slicer as an image computing platform for the Quantitative Imaging Network. *Magnetic resonance imaging*. 2012; 30(9):1323–1341. <https://doi.org/10.1016/j.mri.2012.05.001> PMID: 22770690
7. Mancas M, Gosselin B, Macq B. Segmentation using a region-growing thresholding. In: *Image Processing: Algorithms and Systems IV*. vol. 5672. SPIE; 2005. p. 388–398.
8. Wang L, Gao Y, Shi F, Li G, Chen KC, Tang Z, et al. Automated segmentation of dental CBCT image with prior-guided sequential random forests. *Medical physics*. 2016; 43(1):336–346. <https://doi.org/10.1118/1.4938267> PMID: 26745927
9. Cabezas M, Oliver A, Lladó X, Freixenet J, Cuadra MB. A review of atlas-based segmentation for magnetic resonance brain images. *Computer methods and programs in biomedicine*. 2011; 104(3):e158–e177. <https://doi.org/10.1016/j.cmpb.2011.07.015> PMID: 21871688

10. Ahmad P, Qamar S, Shen L, Saeed A. Context aware 3D UNet for brain tumor segmentation. In: International MICCAI Brainlesion Workshop. Springer; 2020. p. 207–218.
11. Hatamizadeh A, Tang Y, Nath V, Yang D, Myronenko A, Landman B, et al. Unetr: Transformers for 3d medical image segmentation. In: Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision; 2022. p. 574–584.
12. Xie L, Udupa JK, Tong Y, Torigian DA, Huang Z, Kogan RM, et al. Automatic upper airway segmentation in static and dynamic MRI via deep convolutional neural networks. In: Medical Imaging 2021: Biomedical Applications in Molecular, Structural, and Functional Imaging. vol. 11600. SPIE; 2021. p. 131–136.
13. Han L, Chen Y, Li J, Zhong B, Lei Y, Sun M. Liver segmentation with 2.5 D perpendicular UNets. Computers & Electrical Engineering. 2021; 91:107118. <https://doi.org/10.1016/j.compeleceng.2021.107118>
14. Chen J, Lu Y, Yu Q, Luo X, Adeli E, Wang Y, et al. Transunet: Transformers make strong encoders for medical image segmentation. arXiv 2021. arXiv preprint arXiv:210204306;.
15. Isensee F, Petersen J, Klein A, Zimmerer D, Jaeger PF, Kohl S, et al. nnu-net: Self-adapting framework for u-net-based medical image segmentation. arXiv preprint arXiv:180910486. 2018;.
16. Silversmith W. connected-components-3d 3.8. 0;.
17. Chen S, Wang L, Li G, Wu TH, Diachina S, Tejera B, et al. Machine learning in orthodontics: Introducing a 3D auto-segmentation and auto-landmark finder of CBCT images to assess maxillary constriction in unilateral impacted canine patients. The Angle Orthodontist. 2020; 90(1):77–84. <https://doi.org/10.2319/012919-59.1> PMID: 31403836
18. Verhelst PJ, Smolders A, Beznik T, Meewis J, Vandemeulebroucke A, Shaheen E, et al. Layered deep learning for automatic mandibular segmentation in cone-beam computed tomography. Journal of Dentistry. 2021; 114:103786. <https://doi.org/10.1016/j.jdent.2021.103786> PMID: 34425172
19. Cuadros Linares O, Bianchi J, Raveli D, Batista Neto J, Hamann B. Mandible and skull segmentation in cone beam computed tomography using super-voxels and graph clustering. The Visual Computer. 2019; 35(10):1461–1474. <https://doi.org/10.1007/s00371-018-1511-0>
20. Zikic D, Glocker B, Criminisi A. Encoding atlases by randomized classification forests for efficient multi-atlas label propagation. Medical image analysis. 2014; 18(8):1262–1273. <https://doi.org/10.1016/j.media.2014.06.010> PMID: 25042602
21. Cha JY, Yoon HI, Yeo IS, Huh KH, Han JS. Panoptic segmentation on panoramic radiographs: Deep learning-based segmentation of various structures including maxillary sinus and mandibular canal. Journal of Clinical Medicine. 2021; 10(12):2577. <https://doi.org/10.3390/jcm10122577> PMID: 34208024
22. Wang H, Minnema J, Batenburg K, Forouzanfar T, Hu F, Wu G. Multiclass CBCT image segmentation for orthodontics with deep learning. Journal of dental research. 2021; 100(9):943–949. <https://doi.org/10.1177/00220345211005338> PMID: 33783247
23. Mimics. available at: <https://www.materialise.com/en/medical/mimics-innovation-suite>. 2021;.
24. Relu. available at: <https://relueu>. 2021;.
25. Ezhov M, Gusarev M, Golitsyna M, Yates JM, Kushnerev E, Tamimi D, et al. Clinically applicable artificial intelligence system for dental diagnosis with CBCT. Scientific reports. 2021; 11(1):1–16. <https://doi.org/10.1038/s41598-021-01678-5>
26. Schwendicke Fa, Samek W, Krois J. Artificial intelligence in dentistry: chances and challenges. Journal of dental research. 2020; 99(7):769–774. <https://doi.org/10.1177/0022034520915714> PMID: 32315260
27. Schwendicke Fa, Samek W, Krois J. Automated dental tools: <https://github.com/DCBIA-OrthoLab/SlicerAutomatedDentalTools>;
28. Tcheng JE. Optimizing strategies for clinical decision support: summary of a meeting series. National Academy of Medicine; 2017.
29. Chen C, Qin C, Qiu H, Tarroni G, Duan J, Bai W, et al. Deep learning for cardiac image segmentation: a review. Frontiers in Cardiovascular Medicine. 2020; 7:25. <https://doi.org/10.3389/fcvm.2020.00025> PMID: 32195270
30. Xu Y, Wang Y, Yuan J, Cheng Q, Wang X, Carson PL. Medical breast ultrasound image segmentation by machine learning. Ultrasonics. 2019; 91:1–9. <https://doi.org/10.1016/j.ultras.2018.07.006> PMID: 30029074
31. Brosset S, Dumont M, Cevidanes L, Soroushmehr R, Bianchi J, Gurgel ML, et al. Web infrastructure for data management, storage and computation. In: Medical Imaging 2021: Biomedical Applications in Molecular, Structural, and Functional Imaging. vol. 11600. International Society for Optics and Photonics; 2021. p. 116001N.



Published in final edited form as:

J World Fed Orthod. 2022 December ; 11(6): 207–215. doi:10.1016/j.ejwf.2022.10.006.

Three-dimensional digital applications for implant space planning in orthodontics: A narrative review

Jonas Bianchi^{a,b,*}, Gustavo Mendonca^e, Maxime Gillot^c, Heesoo Oh^a, Joorok Park^a, Najla Al Turkestani^{c,d}, Marcela Gurgel^c, Lucia Cevidanes^c

^aDepartment of Orthodontics, Arthur Dugoni School of Dentistry, University of the Pacific, San Francisco, California

^bDepartment of Morphology, Genetics, Orthodontics and Pediatric Dentistry, School of Dentistry University of the State of Sao Paulo, São Paulo State University (Unesp), São Paulo, Brazil

^cDepartment of Orthodontics and Pediatric Dentistry, School of Dentistry, University of Michigan, Ann Arbor, Michigan

^dDepartment of Restorative and Aesthetic Dentistry, Faculty of Dentistry, King Abdulaziz University, Jeddah, Saudi Arabia

^eDepartment of Biologic and Materials Sciences & Prosthodontics, University of Michigan School of Dentistry, Ann Arbor, Michigan

Abstract

In the digital dentistry era, new tools, algorithms, data science approaches, and computer applications are available to researchers and clinicians. However, there is also a strong need for better knowledge and understanding of multisource data applications, including three-dimensional imaging information such as cone-beam computed tomography images and digital dental models for multidisciplinary cases. In addition, artificial intelligence models and automated clinical decision systems are rising. The clinician needs to plan the treatment based on state-of-the-art diagnosis for better and more personalized treatment. This article aimed to review basic concepts and the current panorama of digital implant planning in orthodontics, with open-source and closed-source tools for assessing cone-beam computed images and digital dental models. The visualization and processing of the three-dimensional data allow better implant planning based on bone conditions, adjacent teeth and root positions, and the prognosis of the case. We showed that many tools for assessment, segmentation, and visualization of cone-beam computed tomographic images and digital dental models could facilitate the treatment planning of patients needing implants or space closure. The tools and approaches presented are toward personalized treatment and better prognosis, following the path to a more automated clinical decision system based on multisource three-dimensional data, artificial intelligence models, and digital planning. In summary, the orthodontist needs to analyze each patient individually and use different software or

*Corresponding author: Department of Orthodontics, Arthur Dugoni School of Dentistry, University of the Pacific, 155 5th Street, San Francisco, California, jbianchi@pacific.edu (J. Bianchi).

Competing interests: Authors have completed and submitted the ICMJE Form for Disclosure of potential conflicts of interest. None declared.

Provenance and peer review: Commissioned and internally peer reviewed.

tools that better fit their practice, allowing efficient treatment planning and satisfactory results with an adequate prognosis.

Keywords

Digital dentistry; CBCT; Digital dental models; Implant planning

1. Introduction

The combined orthodontic and prosthodontic therapy is a need that has always existed in clinical practice. The increasing number of adults seeking orthodontic treatment with periodontally compromised and missing tooth makes the combined treatment and synergism between the professionals a real need [1]. Therefore, it is important to not only diagnose the edentulous space but also to plan treatment based on the best evidence available and patients' conditions. Two major approaches can be considered when planning for an edentulous space: 1) closing the space and 2) opening space for the implant and crown placement.

Treatment choice depends on the malocclusion, skeletal facial type, bone conditions, adjacent teeth conditions, and the patients' choice. For implant restoration, clinicians should consider stress-reducing options, including a shorter cantilever, fewer offset loads to the buccal or lingual, number of implants, increased diameter of the implants, splint implant together, and optimal bone quality and quantity [2]. Ideally, all cases require a clear visualization of the results before the surgery is performed because failure of implants occurs in the range of 0% to 20% and is related to bone volume, density, and loading distribution [3].

For the reasons mentioned above, a complete assessment of the patient's condition is necessary. Before, only two-dimensional examinations such as panoramic images (Fig. 1A) were available to clinicians, limiting the capacity of the evaluation to only vertical and horizontal dimensions of the space while evaluating the edentulous area and planning implants [3]. However, with the advances in the engineering field and computational analysis, cone-beam computed tomography (Fig. 1B) and digital dental models (Fig. 1C) are the gold-standard imaging examinations for most patients undergoing orthodontics treatment combined with restorative approaches [4]. We are approaching big data era in dental medicine with the advances in mathematics, storage capacity, and data science fields [5,6]. The amount of information available has increased significantly in the last years, requiring powerful algorithms to process data and predict treatment, diagnosis, or prognosis. In the orthodontics and restorative field, researchers have been testing different machine learning models to segment cone beam computed tomography (CBCT) images for anatomical structures assessment, helping in the treatment decision making, based on previously treated patients and data [7–10].

This article aimed to provide insights into three-dimensional (3D) digital planning of implant placements, anatomical image segmentation, and visualization of CBCT images and digital dental models based on a narrative review of the current literature. We also

highlighted data science and artificial intelligence (AI) models to help clinicians and researchers better make decisions.

2. Image modalities and steps for processing

The images of choice for precise orthodontic and prosthodontic therapy are CBCT images and 3D digital dental models (intraoral scan). The first allows the clinicians to assess the 3D spatial position of adjacent dentition, including roots positions, bone quantity and quality, management of the space available, and creation of surgical guides for implant placement. In contrast, a digital dental model is a precise tool to assess the teeth' surface with accuracy because x-rays produce image noise on the enamel region. For didactic purposes, the basic steps of a fully digital management of orthodontic and implant treatments can be divided into five, as follows:

1. **Data acquisition:** the digital dental model of the upper and lower dentition needs to be acquired either with an intraoral scanner or by digitalizing the plaster model; after the digitalization, the file in the format.stl (stereolithography) needs to be stored. The CBCT scan must also be acquired for the region of interest and the Dicom files stored.
2. **Data and software processing:** the .stl models and Dicom files must be imported into the software of choice, such as the 3Shape Implant Studio, Blue Sky Plan, 3D Slicer, and others.
3. **Data integration and planning:** the two different imaging modalities need to be registered on each other using semiautomated or automated tools provided by the software; also, the bone can be segmented at this stage for visualization of the anatomy.
4. **Selection of implant:** at this stage, based on the integrated 3D image, bone available, anatomy, and orthodontic planning, the more suitable implant size and brand can be selected.
5. **Surgical guide:** at this stage, the surgical guide can be created and exported as .stl for future impressions into a 3D printer, which can be used during implant surgery.

2.1. CBCT image analysis: Visualization and planning of the edentulous space

For optimal use and measurement of x-ray-based imaging examinations, such as the CBCT, the inspection and assessment of each image slice (axial, coronal, and sagittal) are important, as seen in Fig. 2A. The 3D visualization clarifies the anatomical site conditions and adjacent teeth (Fig. 2B). Fig. 2 also shows the importance of visualization in cross-sectional views and 3D rendering. The cross-sectional inspection allows us to obtain precise information on the height and width of the future implant space and the proximity to the maxillary sinus. In contrast, the direct 3D rendering gives a general perspective of the space available. But because of the technical limitation in the automatic rendering based on picking an arbitrary threshold of voxel values, below or above which all gray values are excluded. Therefore, it is commonly used for images with the highest density values

within a particular thickness, such as impacted teeth [11]. This is also illustrated in Fig. 1B, where the imaging methods such as rapid direct 3D rendering allow the assessment of the location of the impacted teeth and mandibular dentition evaluation more accurately than the panoramic x-ray (Fig. 1A).

2.2. Digital dental models and CBCT for assessment of roots and crown position—3D Digital Integration for implant planning

The use of intraoral and dental cast scanners is a routine in most practices for patients undergoing multidisciplinary treatment. The accuracy and reliability of using digital models are comparable to plaster models, with higher precision for measuring dental changes, arch changes, and available space [12,13]. Besides the bone conditions that need to be assessed, the position of the adjacent teeth roots for proper space management and implant planning are essential. Unfortunately, digital dental models cannot accurately provide this information because it only shows the crown position and gingival margins. However, our group has demonstrated that it is possible to calculate tooth crown/root movements with the 3D dental model [14]. Still, the implant surgery would need the exact position of the long axis to assess which type of implant should be chosen.

Fig. 3 shows an automated algorithm created using tools under development by the Dental and Craniofacial Bionetwork for Image Analysis - DCBIA laboratory [15,16]. The figure shows the segmentation of the roots based on AI approaches and crown segmentation from digital dental models. Still, the proper implant surgery would require integrating root information, crown, bone, and soft tissue conditions. Fig. 4 shows the 3D reconstructions of the bone, gum information, and teeth (crown and roots) in a fully integrated approach; Fig. 4A shows the assessment of the available space, and Fig. 4B shows the available area for the implant placement simulation.

The advantages of using this fusion technique are to have a precise 3D model of the teeth crowns and space available in the arch, because with a CBCT image, because of the high density of the enamel and presence of artifacts such as metal restorations, the visualization, and rendering of the dentition is limited, not allowing an accurate implant planning of tooth movement. Current AI methods have tried to improve this limitation; however, there is still a need for improvement [17]. The integration and merging of CBCT with digital dental models is an approach that has been investigated and proven to be successful [18,19]. In addition, this approach allows for overcoming the challenges of proper segmentation of the crowns using information from the 3D dental models and not from CBCT images. In summary, Fig. 5 shows what would be an ideal assessment and workflow for implant surgery and orthodontic planning.

2.3. AI approaches and clinical decision systems

AI has become an important tool in dentistry. Because of the amount of clinical and imaging data that clinicians and researchers have available and the implementation of better data science approaches, the decision-making process can be facilitated by trained algorithms. In addition, clinical decision support systems have been incorporating knowledge with patient-specific data to serve clinicians with tools that enhance this process [7,20]. Pareek

et al. [21] reported that almost 4000 dental implants are marketed worldwide, with varying treatment techniques and structures. Therefore, knowing which one is better applicable to a specific patient, based on their condition, is primordial for success. In this field, AI can help the computer-aided design/computer-aided manufacturing, and panoramic radiographs classify the implant structure, and the AI approaches can help the dentist recognize and rank the implants, avoiding complications.

Researchers have also focused on detecting dental implant failures and fractures using AI methods. In 2020, Lee et al. [22] evaluated the reliability of three deep convolutional neural networks (VGGNet-19, GoogLeNet Inception-v3, and automated DCNN) for the classification of fractured dental implants in panoramic and periapical rays. They used a database with 21,398 fractured implants and found that the AI approaches had acceptable accuracy in the detection and classification of fractured implants, with the best performance obtained while using periapical radiographic images alone (automated DCNN with an area under the curve of 0.984).

AI can also assess the shape of the definitive monolithic zirconia restorations because they cannot be retouched in the mouth. Therefore, Lerner et al. [10] have demonstrated digital dental models and machine learning in fixed implant prosthodontics. The author used a fully digital protocol using AI, which allowed the successful restoration of single locking-taper implants with monolithic zirconia restorations, and they stated that the marginal adaptation, quality of interproximal and occlusal contacts, and aesthetic integration were excellent.

3. Clinical cases and clinical applications

3.1. Case 1—Guided implant surgery in a prosthodontics treatment planning

Fig. 6 shows a prosthodontics case and treatment planning with a digital setup (Integrating Digital Surface Scanners and CBCT Images) for a patient where the inferior lower right incisor is compromised because of periodontal health. In this report, the software 3Shape Implant Studio software (3Shape) was used to integrate CBCT images and digital dental models to create the implant surgical guide. This software allows the clinician to incorporate the CBCT image and intra oral-scan for decision making based on the quantity of bone available and prosthetic space available for the future crown. The first step is to perform an intraoral scanner of the patient to obtain the digital dental model and take a CBCT examination, having access to the Dicom files. Then, the 3D.stls (stereolithography) models of the upper and lower dentition and the Dicom files (CBCT) are imported into the software. A semi-automatic approach allows the registration of both image modalities, and a library containing prosthetic component information can be used to select the more accurate implant. Ultimately, the surgical guide can be virtually fabricated and sent to 3D impression. Fig. 7 shows the final implant and components using the patient dentition data (.stl models) and bone anatomy (CBCT image).

3.2. Case 2—Orthodontic and restorative planning: minimizing the number of implants using optimal biomechanics and digital diagnosis

Figs. 8 and 9 illustrate a patient that had orthodontic treatment and implant surgery at the University of the Pacific—Orthodontics Department. The 38-year-old female patient was diagnosed with canine Class I on the right side, full cusp Class II on the left, lower midline deviation to the left, moderate to severe crowding in the upper and lower arches, and missing lower first molars bilateral, upper right first and second premolar, and upper left first premolar. Her chief complaint was the edentulous spaces, and her general dentist recommended an orthodontic consultation before her restorative treatment planning. The initial lateral cephalogram and panoramic image were generated from the CBCT and are shown in Fig. 10. The treatment proposed was to reduce the number of implants by retracting the laterals and canines into the first premolar space with moderate anchorage, maintaining the right maxillary second premolar, upright the mandibular second and third molars, close the first molar spaces, and achieving a full cusp CI II molar and CI I canine relationship bilaterally. This proposed plan allows the patient to have only a single implant in the first right upper premolar region. The pre-adjusted 0.022×0.028 " MBT prescription fixed appliance (Victory, 3M Unitek) was used. Fig. 9A shows the progress in the orthodontic treatment. TADs (1.4 mm diameter, 6 mm length, Vector, Ormco) were used to minimize the side effects of the continuous archwire in the lower dentition that was used to upright the lower molars and to close the space without losing anterior anchorage, and gable bends were added to promote mesial root tip of the second molars in the 0.017×0.025 " SS archwire. In the upper arch, the sequence progressed from 0.016×0.022 " NiTi to 0.019×0.025 " NiTi to 0.019×0.025 " TMA. The spaces were closed by sliding mechanics except for the upper right second premolar, which was maintained, and the space was adequate for an implant. Fig. 9B shows the final results after single implant placement, and Fig. 11 shows the lateral cephalogram, panoramic image, and cephalometric tracing immediately after treatment.

3.3. Case 3—Orthodontic planning using digital tools for orthodontics and implant space treatment simulation

A 17-year-old male patient presented to the University of the Pacific - Orthodontics clinic with the chief complaint of retained deciduous mandibular and maxillary molars and impacted upper second premolars. In addition, he had the absence of the lower second premolar, requiring detailed treatment planning to address his condition (Fig. 12A). Therefore, a virtual setup was performed using the software Archform and ClinCheck. Fig. 12B shows his initial digital models, and Fig. 12C shows his CBCT images for assessing the conditions of the lower deciduous molar. Because the patient's guardian refuses to extract the second upper premolar, the two main treatments proposed were as follows: 1) orthodontic traction of the impacted upper second premolars and maintenance of the lower space for future implant in the deciduous molar region or 2) orthodontic premolar traction of the impacted upper second premolars and maintenance of the space for future implant in the lower deciduous molar region. Next, a virtual simulation was performed with two different software: Archform (Fig. 12E) and ClinCheck (Fig. 12D). This setup allowed us to see that the maintenance of space was an appropriate choice because of the occlusion of the second upper molar.

4. Discussion

The digitalization of treatments is rising in orthodontics and prosthodontics with AI approaches. Especially with the use of aligners, digital dental models, and more access to CBCT images, better and more personalized treatment planning is possible. In a recent publication, Shroff et al. [1] showed two case reports using a digital workflow with virtual orthodontic planning and the design of surgical guides for implant placement. As the main advantages of using digital aligner therapy, they cited the possibility of using a 3D setup and accurate planning of the final tooth position, simulation of orthodontic movement, and space creation for the implant; but the disadvantages are the relatively high cost compared with fixed appliances and the need for patient cooperation while using the aligners. In comparison, we also showed in the current study that the simulation of tooth movement could be done without the need for aligner therapy. Fig. 12D and 12E show a virtual setup with the purpose of treatment planning only, where the patient received fixed appliances but had the digital setup for a better prediction of the occlusion after closing the implant space (Fig. 12D) or implant space maintenance (Fig. 12E).

It is important to highlight the multidisciplinary aspects of orthodontics and implant treatment involving specialties such as periodontics, dental implant, orthodontics, and prosthodontics. Blasi et al. [23] presented the digital preorthodontic implant placement, showing the steps for a fully digitized treatment using digital bracket placement and guided surgeries for periodontics and implant purposes. In addition, Tarraf et al. [24], in 2018, pointed out that digital technology has a great impact in the medical and dental field, allowing personalized and better treatment options with computer-aided design/computer-aided manufacturing techniques, indirect bonding trays, customized wires, and even remote monitoring of treatment. The authors also showed the integration of 3D photos and facial scans in treatment planning.

Overall, the digital combined orthodontics and implant surgery is an emerging and needed tool. This approach allows for better and more robust treatment planning, with fewer variables and more predictability. This was also demonstrated by Spalthoff et al. [25], who evaluated the efficiency of a digital workflow of prosthetic teeth positioning between virtual standard-sized digitally constructed and conventional dental laboratory-fabricated prostheses. They found that the digital workflow provided accurate final results.

5. Conclusion

Combining multisource images such as CBCT and 3D digital dental models is essential for proper planning and managing implant surgery combined with orthodontic treatment. In addition, the use of data science approaches, advances in the image analysis field, and new AI approaches are becoming more popular among clinicians because of translational research and software availability. Therefore, better and more personalized treatment can be available, helping the clinical decision making and the prognosis.

Acknowledgments

American Association of Orthodontists Foundation; National Institute of Dental and Craniofacial Research, Grant/Award Number: R01DE024450 and the Research Enhancement Award Activity 141 from the University of the Pacific, Arthur A. Dugoni School of Dentistry.

References

- [1]. Shroff B, Siegel SM, Feldman S, Siegel SC. Combined orthodontic and prosthetic therapy. Special considerations. *Dent Clin North Am* 1996;40:911–43 [PubMed: 8886545]
- [2]. Misch CE, Goodacre CJ, Finley JM, et al. Consensus Conference Panel Report: crown-height space guidelines for implant dentistry - part 1. *Implant Dent* 2005;14:312–18 [PubMed: 16361879]
- [3]. Goodacre CJ, Bernal G, Rungcharassaeng K, Kan JYK. Clinical complications with implants and implant prostheses. *J Prosthet Dent* 2003;90:121–32. [PubMed: 12886205]
- [4]. Howerton WB, Mora MA. Advancements in digital imaging: what is new and on the horizon? *J Am Dent Assoc* 2008;139 (Suppl.):20S–4S [PubMed: 18539868]
- [5]. Raghupathi W, Raghupathi V. Big data analytics in healthcare: promise and potential. *Health Inf Sci Syst* 2014;2:3 [PubMed: 25825667]
- [6]. Alyass A, Turcotte M, Meyre D. From big data analysis to personalized medicine for all: challenges and opportunities. *BMC Med Genomics* 2015;8:33 [PubMed: 26112054]
- [7]. Al Turkestani N, Bianchi J, Deleat-Besson R, et al. Clinical decision support systems in orthodontics: a narrative review of data science approaches. *Orthod Craniofac Res* 2021;24(Suppl. 2):26–36. [PubMed: 33973362]
- [8]. Memon AR, Li J, Egger J, Chen X. A review on patient-specific facial and cranial implant design using Artificial Intelligence (AI) techniques. *Expert Rev Med Devices* 2021;18:985–94. [PubMed: 34404280]
- [9]. Bernauer SA, Zitzmann NU, Joda T. The use and performance of artificial intelligence in prosthodontics: a systematic review. *Sensors (Basel)* 2021;21:6628. [PubMed: 34640948]
- [10]. Lerner H, Mouhyi J, Admakin O, Mangano F. Artificial intelligence in fixed implant prosthodontics: a retrospective study of 106 implant-supported monolithic zirconia crowns inserted in the posterior jaws of 90 patients. *BMC Oral Health* 2020;20:80. [PubMed: 32188431]
- [11]. Venkatesh E, Elluru SV. CBCT: basics and applications in dentistry. *J Istanb Univ Fac Dent* 2017;51:102–21.
- [12]. De Luca Canto G, Pachêco-Pereira C, Lagravere MO, Flores-Mir C, Major PW. Intra-arch dimensional measurement validity of laser-scanned digital dental models compared with the original plaster models: a systematic review. *Orthod Craniofac Res* 2015;18:65–76.
- [13]. Leifert MF, Leifert MM, Efstratiadis SS, Cangialosi TJ. Comparison of space analysis evaluations with digital models and plaster dental casts. *Am J Orthod Dentofacial Orthop* 2009;136 16.e1–4 [PubMed: 19577140]
- [14]. Cong A, Massaro C, Ruellas ACO, et al. Dental long axes using digital dental models compared to cone-beam computed tomography. *Orthod Craniofac Res* 2022;25:64–72. [PubMed: 33966340]
- [15]. Deleat-Besson R, Le C, Al Turkestani N, et al. Automatic segmentation of dental root canal and merging with crown shape. *Annu Int Conf IEEE Eng Med Biol Soc* 2021;2021:2948–51. [PubMed: 34891863]
- [16]. Github. SlicerAutomatedDentalTools. Available from: <https://github.com/DCBIA-OrthoLab/SlicerAutomatedDentalTools> [Accessed 10/02/2022].
- [17]. Wang H, Minnema J, Batenburg KJ, Forouzanfar T, Hu FJ, Wu G. Multiclass CBCT image segmentation for orthodontics with deep learning. *J Dent Res* 2021;100:943–9. [PubMed: 33783247]
- [18]. Rangel FA, Maal TJJ, Bronkhorst EM, et al. Accuracy and reliability of a novel method for fusion of digital dental casts and cone beam computed tomography scans. *PLoS One* 2013;8:e59130. [PubMed: 23527111]

- [19]. Lin HH, Chiang WC, Lo LJ, Wang CH. A new method for the integration of digital dental models and cone-beam computed tomography images. *Annu Int Conf IEEE Eng Med Biol Soc* 2013;2013:2328–31. [PubMed: 24110191]
- [20]. Shan T, Tay FR, Gu L. Application of artificial intelligence in dentistry. *J Dent Res* 2021;100:232–44. [PubMed: 33118431]
- [21]. Pareek M, Kaushik B. Artificial intelligence in prosthodontics: a scoping review on current applications and future possibilities. *Int J Adv Med* 2022;9:367
- [22]. Lee DW, Kim SY, Jeong SN, Lee JH. Artificial intelligence in fractured dental implant detection and classification: evaluation using dataset from two dental hospitals. *Diagnostics (Basel)* 2021;11:233. [PubMed: 33546446]
- [23]. Blasi G, Blasi I, Blasi A, Elnabawi O, Murphy KG, Stappert D. The digital POIP concept: Preorthodontic implant placement. *J Esthet Restor Dent* 2020;32:545–53. [PubMed: 32613725]
- [24]. Tarraf NE, Ali DM. Present and the future of digital orthodontics ★. *Semin Orthod* 2018;24:376–85.
- [25]. Spalthoff S, Borrmann M, Jehn P, Rahlf B, Gellrich NC, Korn P. Comparison of conventional and digital workflow for dental rehabilitation with a novel patient-specific framework implant system: an experimental dataset evaluation. *Int J Implant Dent* 2022;8:4. [PubMed: 35072825]

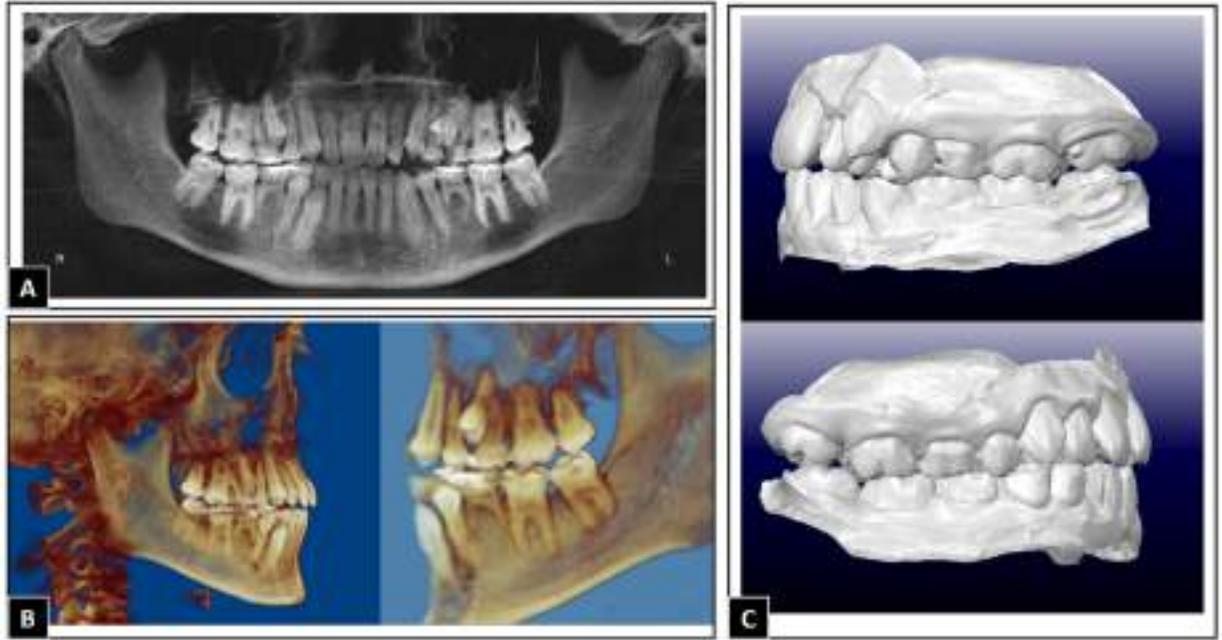


Fig. 1. Example of patient's digital records. (A) Panoramic image; (B) Cone beam computed tomography 3D rendering; (C) Digital dental model from an intraoral scanner.

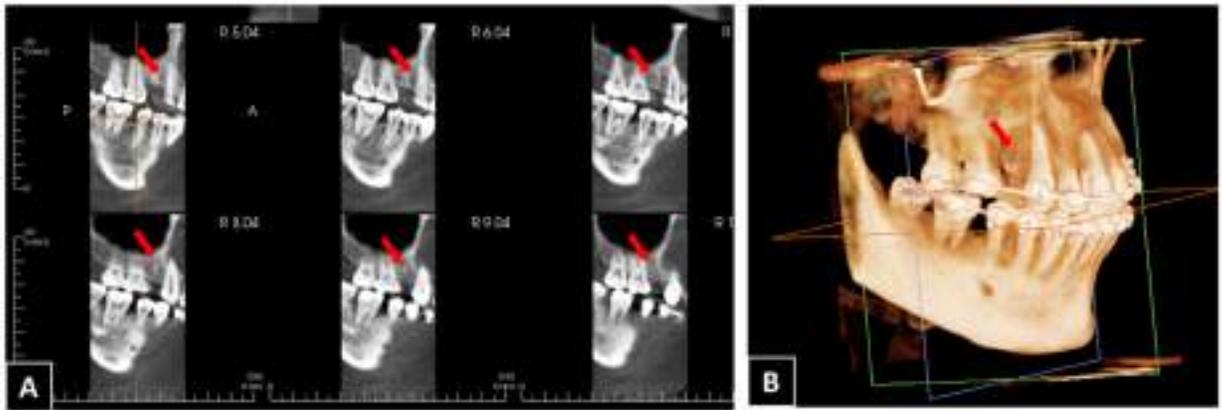


Fig. 2. CBCT examination of a patient in preparation for implant placement in the first premolar region. (A) Sagittal slices show the extension of the implant area. (B) 3D rendering showing the maxilla, mandible, and adjacent teeth using Invivo software (V 6.5.0).

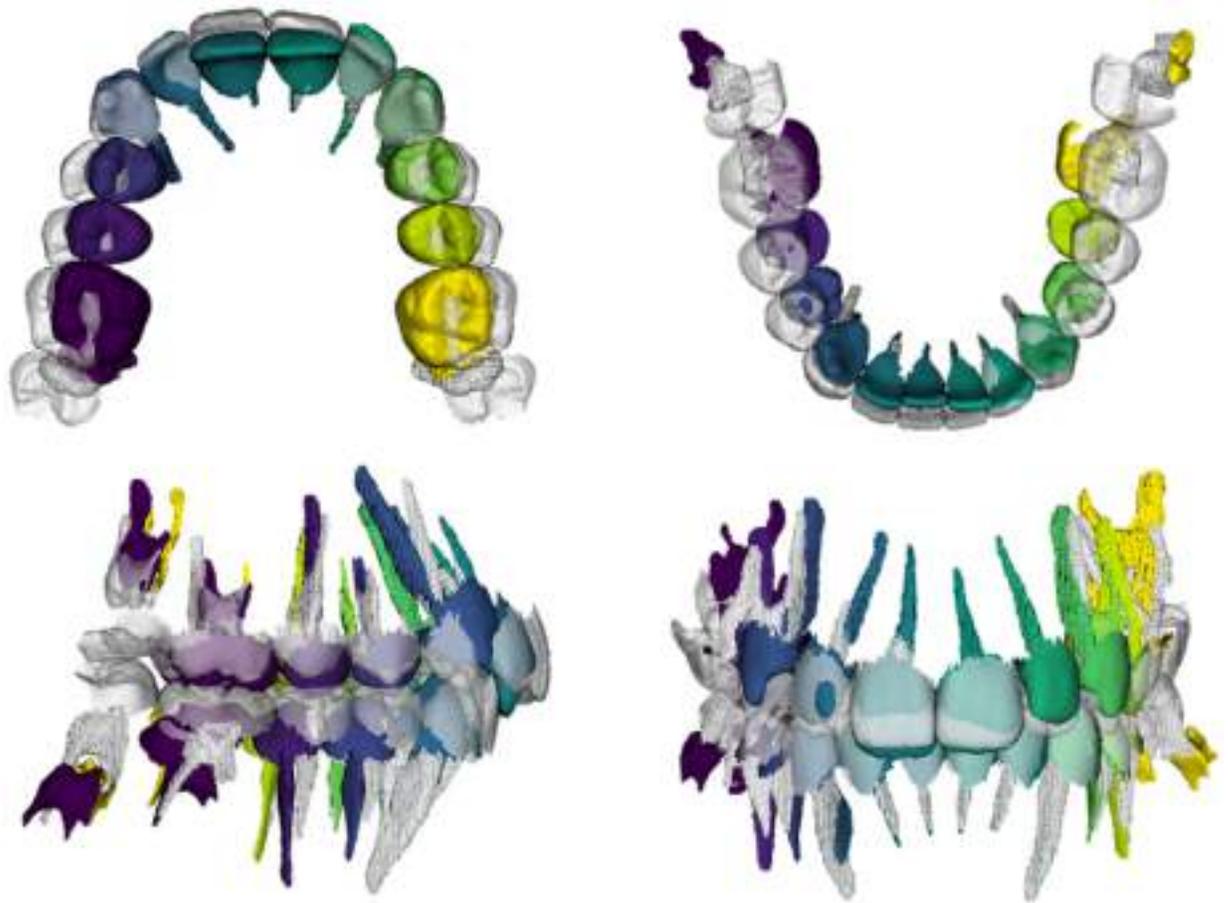


Fig. 3.
3D visualization of the crown and roots position using a digital dental model and CBCT segmentation images.

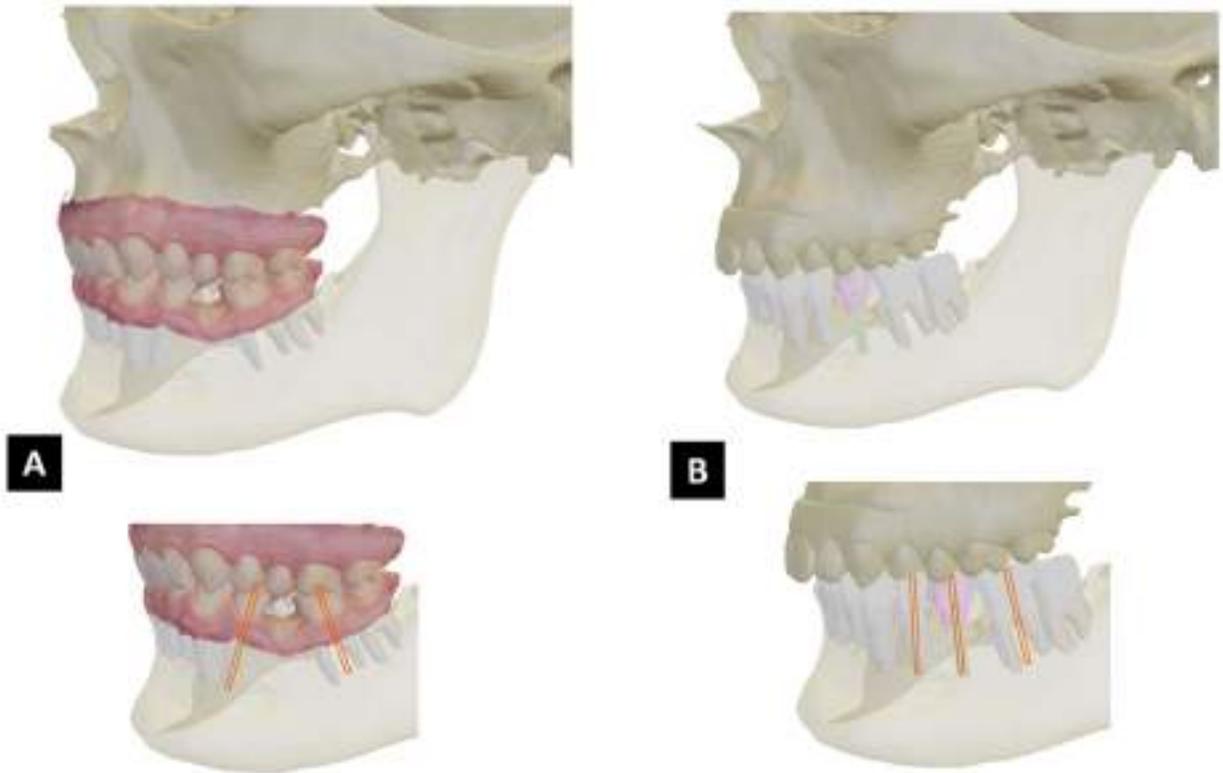


Fig. 4. Multisource images for implant planning using CBCT and digital dental models. (A) Visualization of the space available and adjacent teeth long axis. (B) Implant simulation.

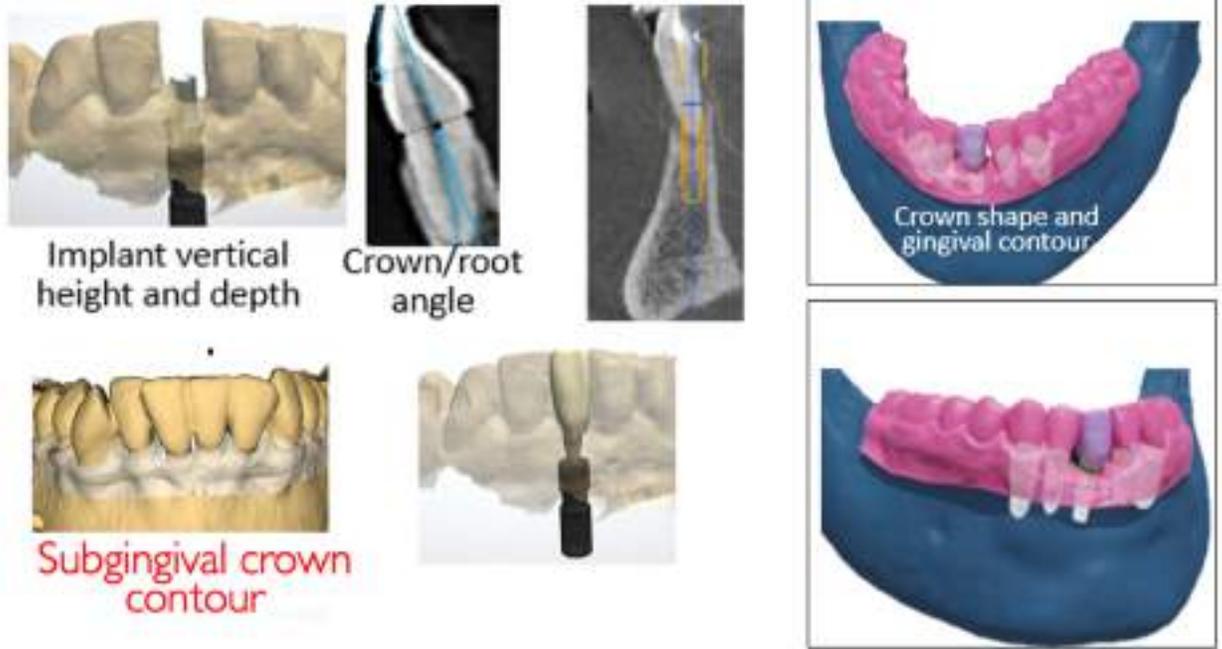


Fig. 5.
Digital 3D workflow for implant surgery simulation.



Fig. 6. Patient data imported to the software for implant planning (3shape implant studio). (A) Visualization of the CBCT reconstruction; (B) creation of the surgical guide.

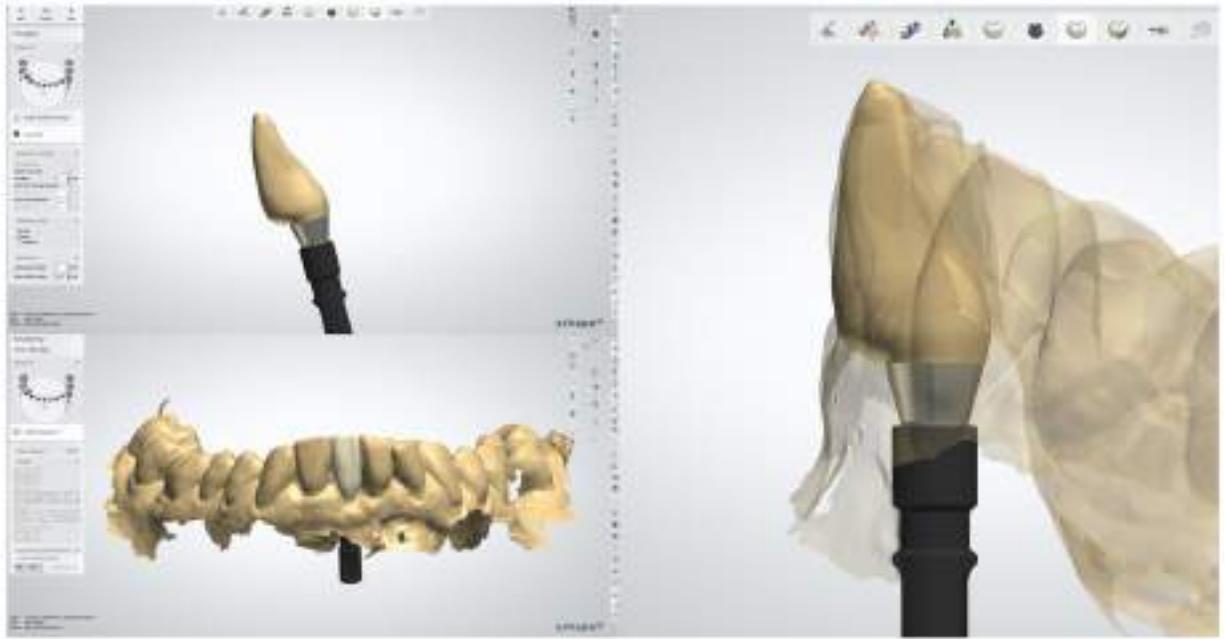


Fig. 7. Implant components were planned into the 3shape implant studio. Design of the implant based on the anatomy of the bone and dentition.



Fig. 8. CBCT 3D rendering of a patient needing an implant. The image allows the visualization of the treatment choice and planning of the biomechanics.



Fig. 9. Clinical case showing a patient that had orthodontic combined with surgery for the implant of the upper right premolar. (A) Initial, progress, and final photos of the orthodontics treatment. Most spaces were closed, and only one implant was planned. (B) Implant x-rays and final photos after the implant surgery.

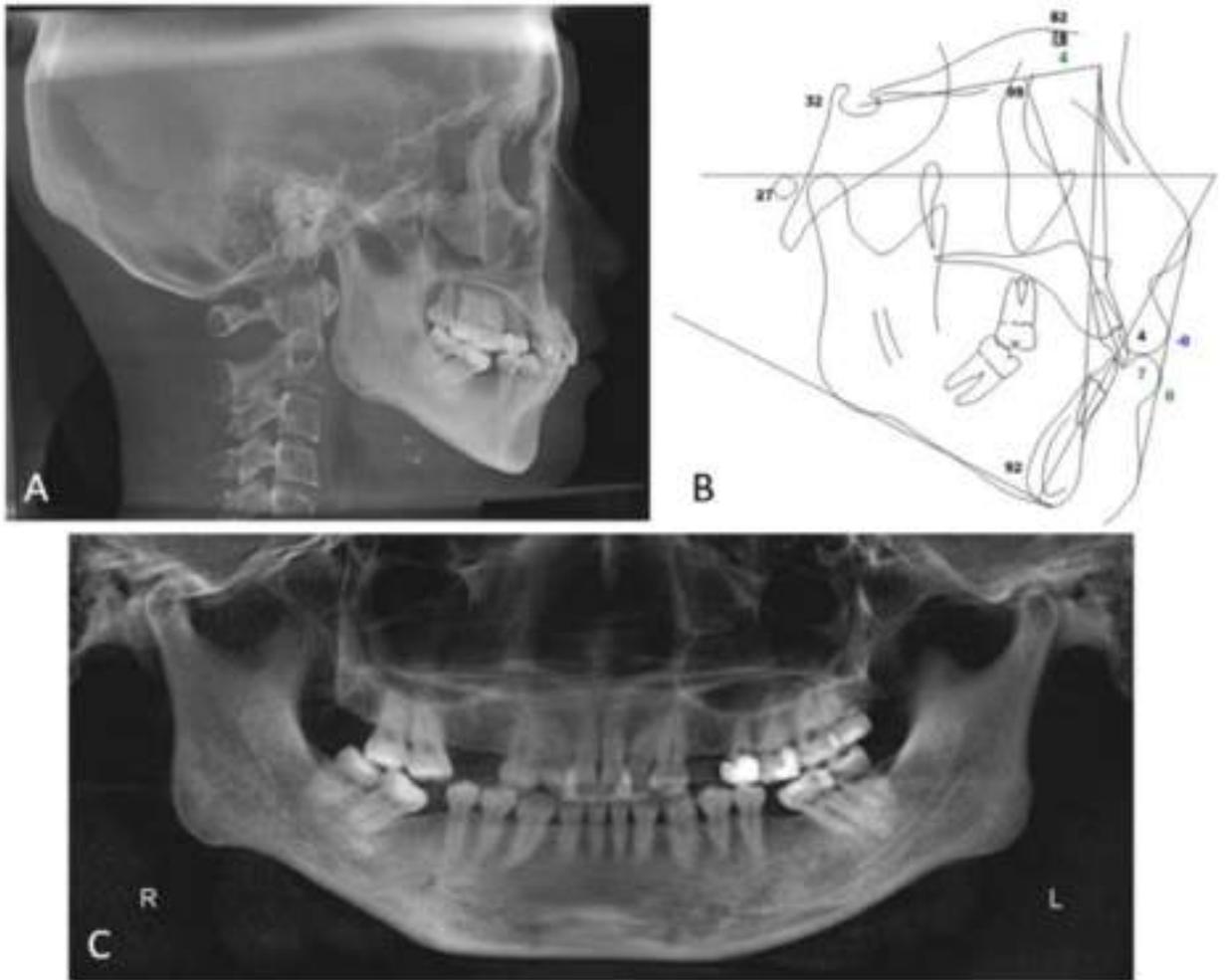


Fig. 10. Initial treatment records. (A) Pretreatment lateral cephalogram, (B) lateral cephalometric tracing, and (C) panoramic image.

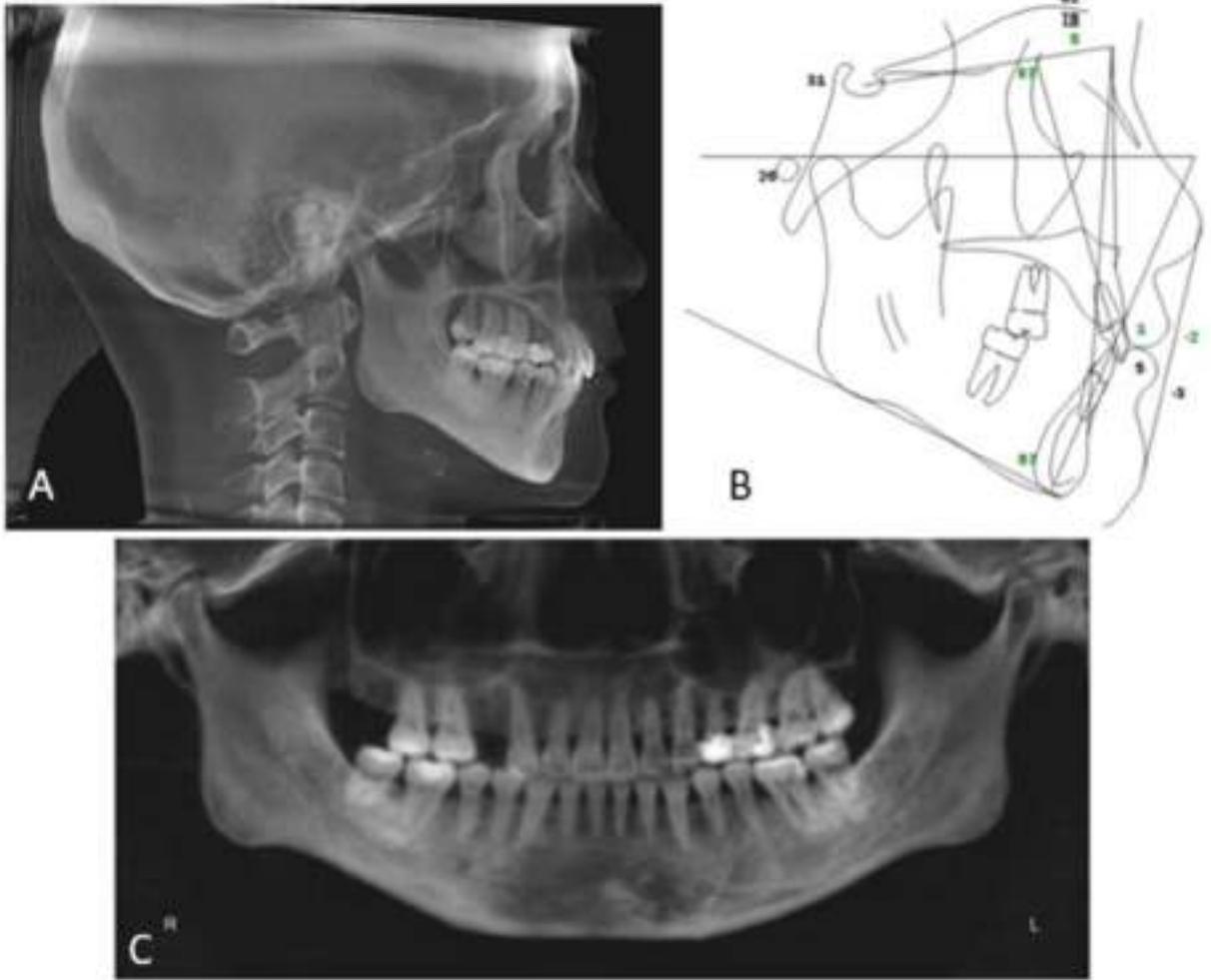


Fig. 11. Final treatment records. (A) Final treatment lateral cephalogram, (B) lateral cephalometric tracing, and (C) panoramic image.

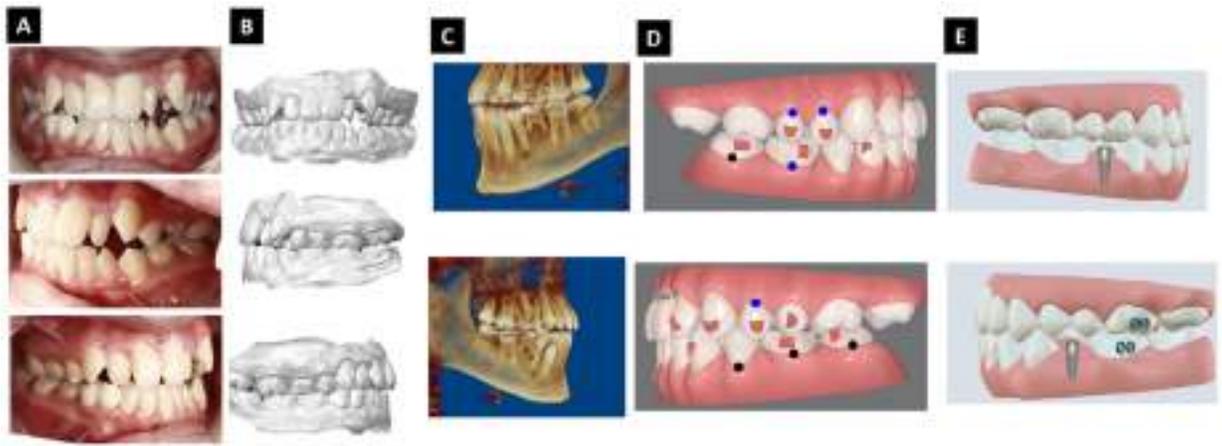


Fig. 12. Simulation of treatment in a patient with the presence of bilateral molar deciduous and congenital absence of bilateral lower second premolars. (A) Intraoral photos; (B) digital models; (C) CBCT 3D rendering; (D) simulation of treatment with the closure of the spaces; and (E) simulation of treatment with preservation of the space for implant planning.