Automatic Landmark Identification on IntraOralScans

Chapter · January 2023

DOI: 10.1007/978-3-031-23179-7_4


Najla al Turkestani
University of Michigan


Marcela Gurgel
Universidade Federal do Ceará


Jonas Bianchi
University of Michigan


Marilia Yatabe
University of Michigan

Automatic Landmark Identification on IntraOralScans

Baptiste Baquero[1,9](✉), Maxime Gillot[1,9], Lucia Cevidanes[1], Najla Al Turkestani[1,5], Marcela Gurgel[1], Mathieu Leclercq[4,9], Jonas Bianchi[1,3], Marilia Yatabe[1], Antonio Ruellas[1,2], Camila Massaro[8], Aron Aliaga[1], Maria Antonia Alvarez Castrillon[6], Diego Rey[6], Juan Fernando Aristizabal[7], and Juan Carlos Prieto[4]

[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.

Y. Chen et al. (Eds.): CLIP 2022, LNCS 13746, pp. 32–42, 2023.
https://doi.org/10.1007/978-3-031-23179-7_4

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 (ALIIOS) 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 ± 0.9 µm. 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).

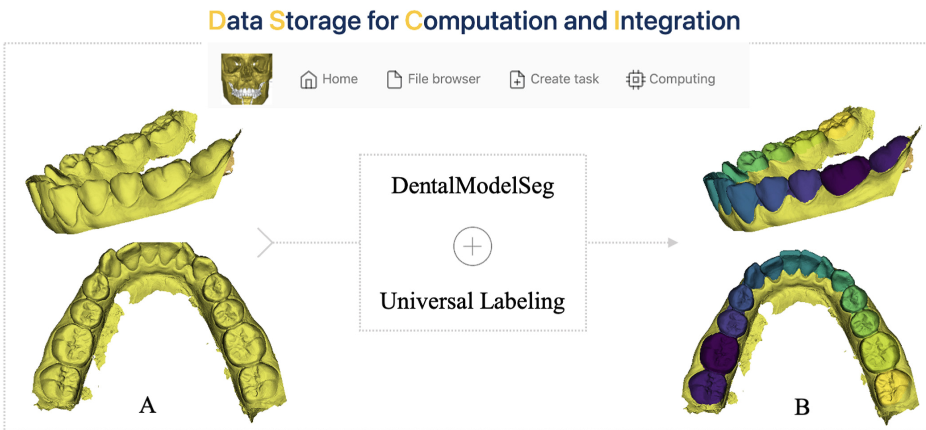

Fig. 1. IOS pre-processing A. IOS acquired using iTero® or 3Shape scanners B. Segmentation of the dental crowns with the two open-source tools: DentalModelSeg and Universal Labelling [4] to segment and assign the universal numbering.

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 ALIIOS 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^o 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 "HardPhongShading" 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 ALIIOS 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 ALIIOS 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 × 12 processors, and 2 TB disk capacity.

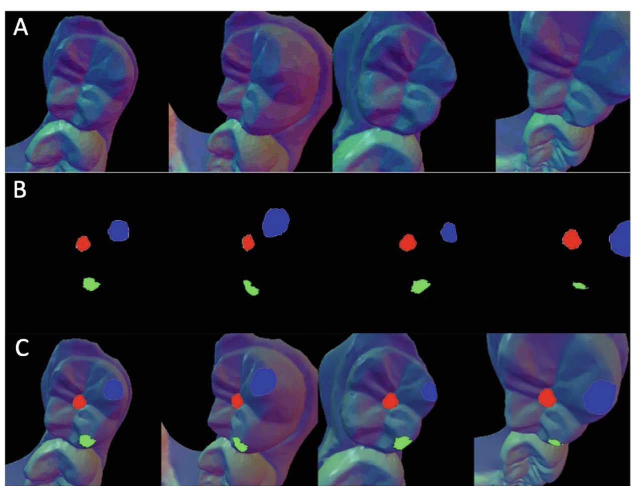

Fig. 2. A.U-Net input images. B.U-Net output patches. C.Identification of the surface meshes vertices using the U-Net output.

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

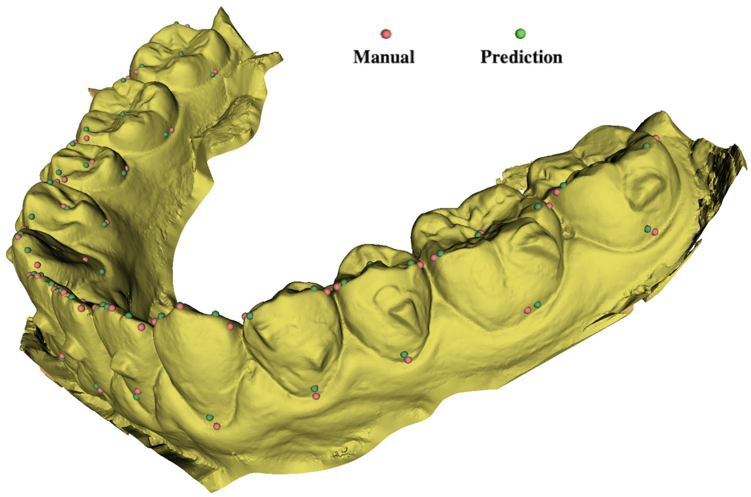

Fig. 3. Comparison between manual landmarks and predicted on the lower arch. The red spheres represent the clinician's landmarks (manual) and the green spheres are the ones predicted by ALIIOS. The diagram displays the average error (mm) for each landmark. (Color figure online)

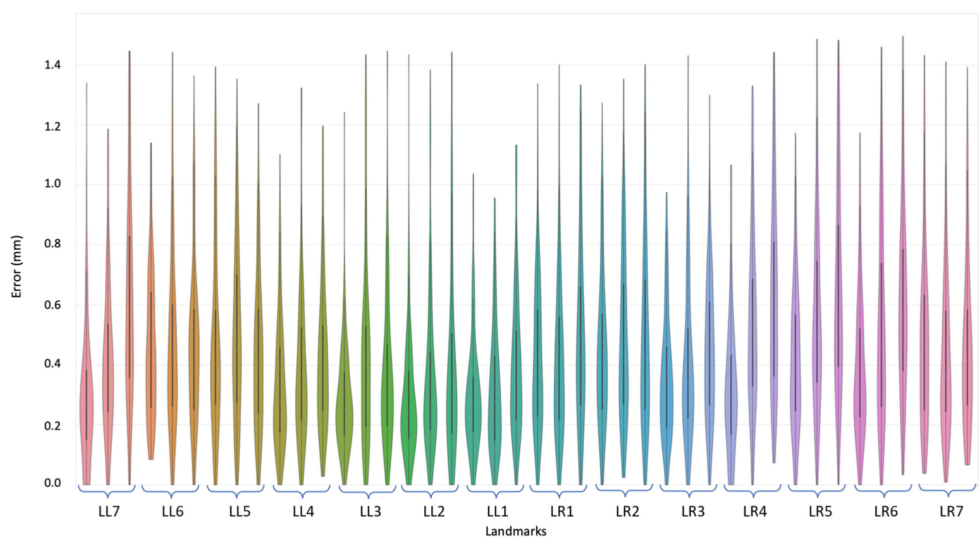

Fig. 4. Accuracy for the lower occlusal landmarks.

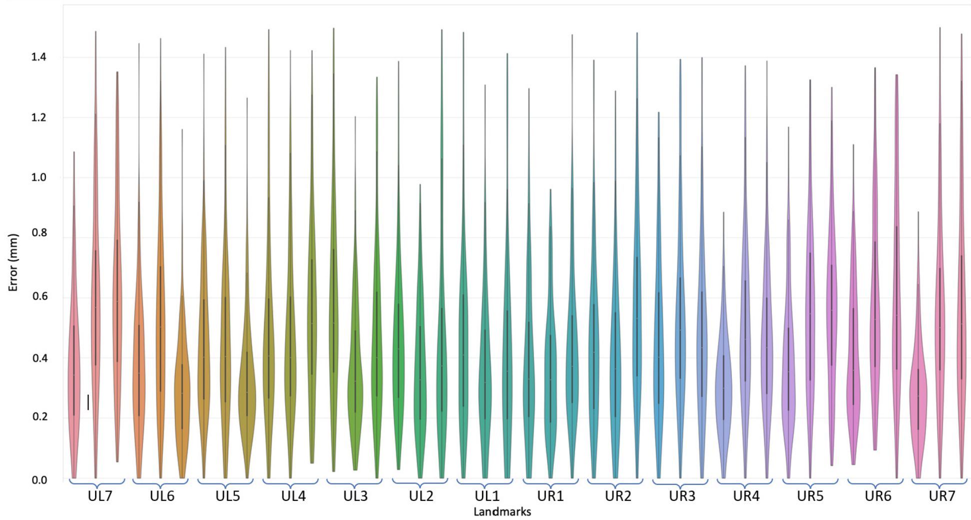

Fig. 5. Accuracy for the upper occlusal landmarks.

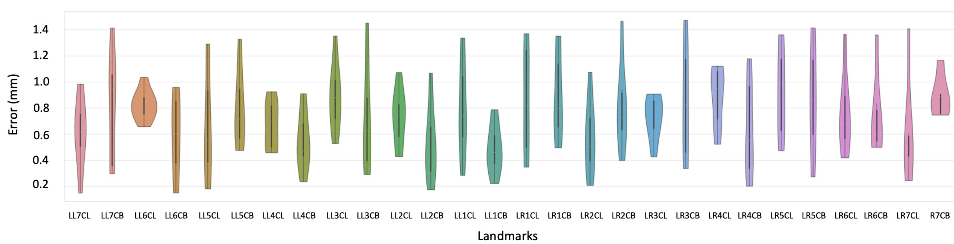

Fig. 6. Accuracy for the lower cervical landmarks.

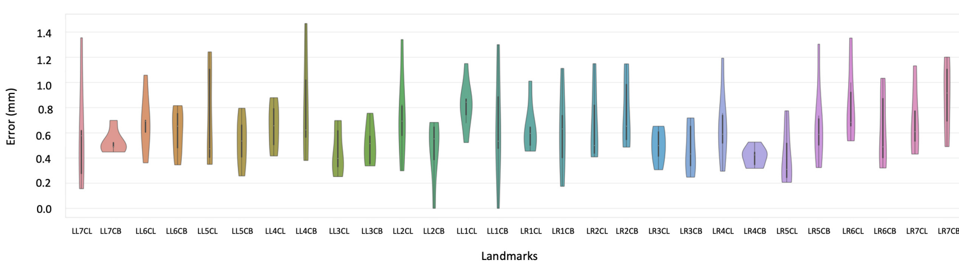

Fig. 7. Accuracy for the upper cervical landmarks.

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 langual and “CB” stands for cervical buccal.

5 Discussion

This paper presents the ALIIOS algorithm, a novel method for robust and accurate automatic landmark identification on IOS. The ALIIOS 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 ALIIOS 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 ALIIOS 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 ALIIOS 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 ALIIOS intuitive interface allows users to predict occlusal or cervical landmarks on the selected tooth. Additionally, to allow users to automatically compute measurements between the ALIIOS 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 ALIIOS 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.

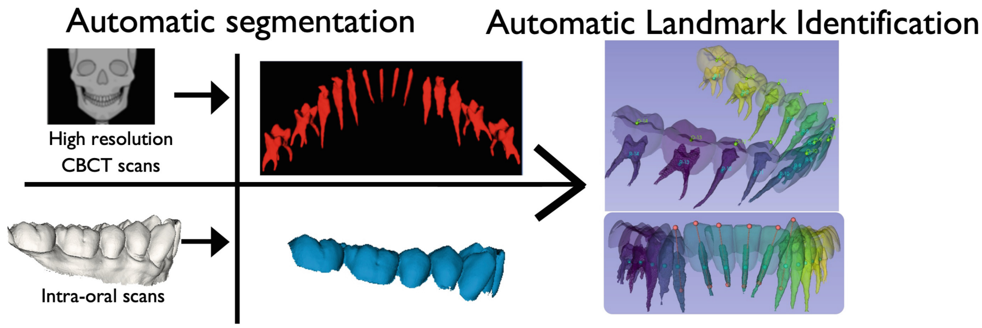

Fig. 8. Proposed future work: Segmentation of root canal and prediction of the tooth long axis.

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 (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)