

# Automatic landmark identification in cone-beam computed tomography

Maxime Gillot<sup>1,2</sup> | Felicia Miranda<sup>1,3</sup>  | Baptiste Baquero<sup>1,2</sup> | Antonio Ruellas<sup>4</sup> |  
 Marcela Gurgel<sup>1</sup>  | Najla Al Turkestani<sup>1,5</sup>  | Luc Anchling<sup>1,2</sup> | Nathan Hutin<sup>1,2</sup> |  
 Elizabeth Biggs<sup>1</sup> | Marilia Yatabe<sup>1</sup>  | Beatriz Paniagua<sup>6</sup> |  
 Jean-Christophe Fillion-Robin<sup>6</sup> | David Allemang<sup>6</sup> | Jonas Bianchi<sup>7</sup>  |  
 Lucia Cevidanes<sup>1</sup>  | Juan Carlos Prieto<sup>8</sup> 

<sup>1</sup>Department of Orthodontics and Pediatric Dentistry, University of Michigan School of Dentistry, MI, Ann Arbor, USA

<sup>2</sup>CPE Lyon, Lyon, France

<sup>3</sup>Department of Orthodontics, Bauru Dental School, University of São Paulo, Bauru, Brazil

<sup>4</sup>Department of Orthodontics and Pediatric Dentistry, School of Dentistry, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil

<sup>5</sup>Department of Restorative and Aesthetic Dentistry, Faculty of Dentistry, King Abdulaziz University, Jeddah, Saudi Arabia

<sup>6</sup>Kitware Inc., Chapel Hill, NC, USA

<sup>7</sup>Department of Orthodontics, University of the Pacific, San Francisco, CA, USA

<sup>8</sup>Department 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](mailto: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  $1.54 \pm 0.87$  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.<sup>1,2</sup>

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,<sup>3</sup> treatment planning,<sup>4</sup> intervention, follow-up of anatomical changes, or disease conditions,<sup>5</sup> and simulations.<sup>6</sup> Landmark identification may serve as initialization to other algorithms such as segmentation algorithms,<sup>7</sup> or image-to-image registration.<sup>8,9</sup> Most of the available solutions for landmark detection rely on machine learning,<sup>10-12</sup> 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,<sup>11,12</sup> one-shot displacement estimation,<sup>13,14</sup> or end-to-end image mapping techniques.<sup>15,16</sup> 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.<sup>2</sup> 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](http://www.itksnap.org))<sup>17</sup> and 3D Slicer, version 4.11 ([www.slicer.org](http://www.slicer.org))<sup>18</sup> were used by clinician experts to orient the scans and place the landmarks. Head orientation was performed accordingly with a previous study.<sup>19</sup> 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.<sup>20</sup>

### 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.

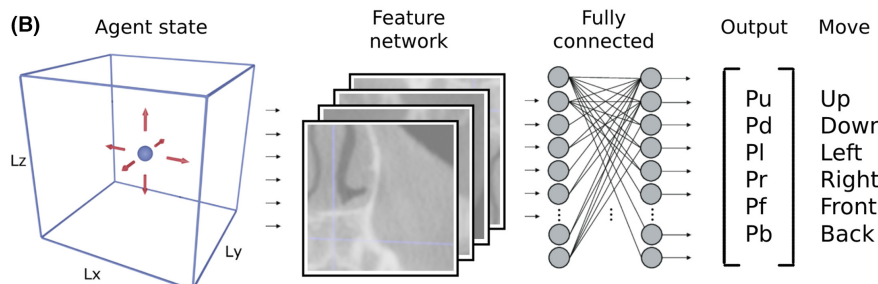
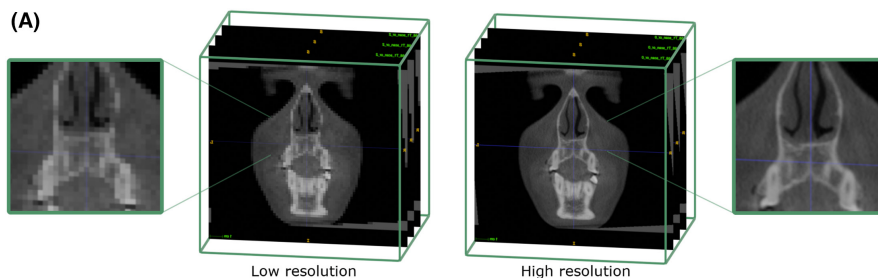
Description of the landmarks	
<b>Cranial base</b>	
Ba	Placed at the most posteroinferior point of the anterior margin of the foramen magnum in the midsagittal plane
S	Placed on the most central point of sella turcica from supero-inferior, antero-posterior, and transversal aspects
N	Placed at the most anterosuperior junction of the nasofrontal suture
<b>Maxilla</b>	
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
<b>Mandible</b>	
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.<sup>21</sup> 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.

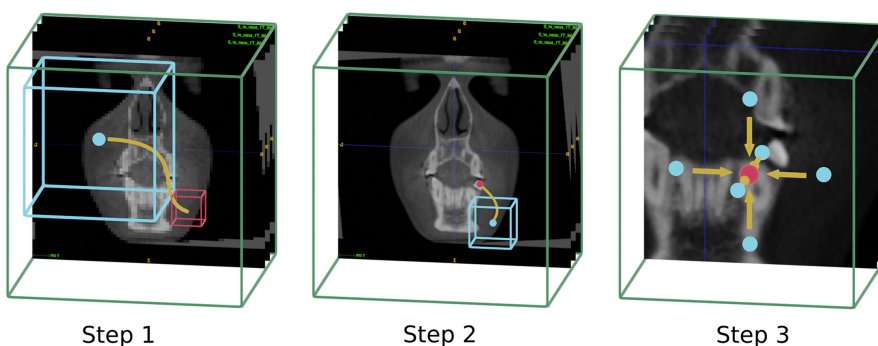
## 2.3 | Training the agents

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

**FIGURE 1** A, Visualization of an environment. On the left, the low 1 mm resolution was re-scaled from the high-resolution 0.3mm 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).



trained using 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 200000000 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.<sup>22,23</sup> 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.<sup>24</sup> 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.<sup>25,26</sup> Landmark placement using both surface models and MPR images took an average of 10:41 ± 4:01 minutes to trace each patient.<sup>25</sup> 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.<sup>27-29</sup> 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.<sup>2</sup> 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.<sup>29</sup> 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.<sup>29</sup> 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



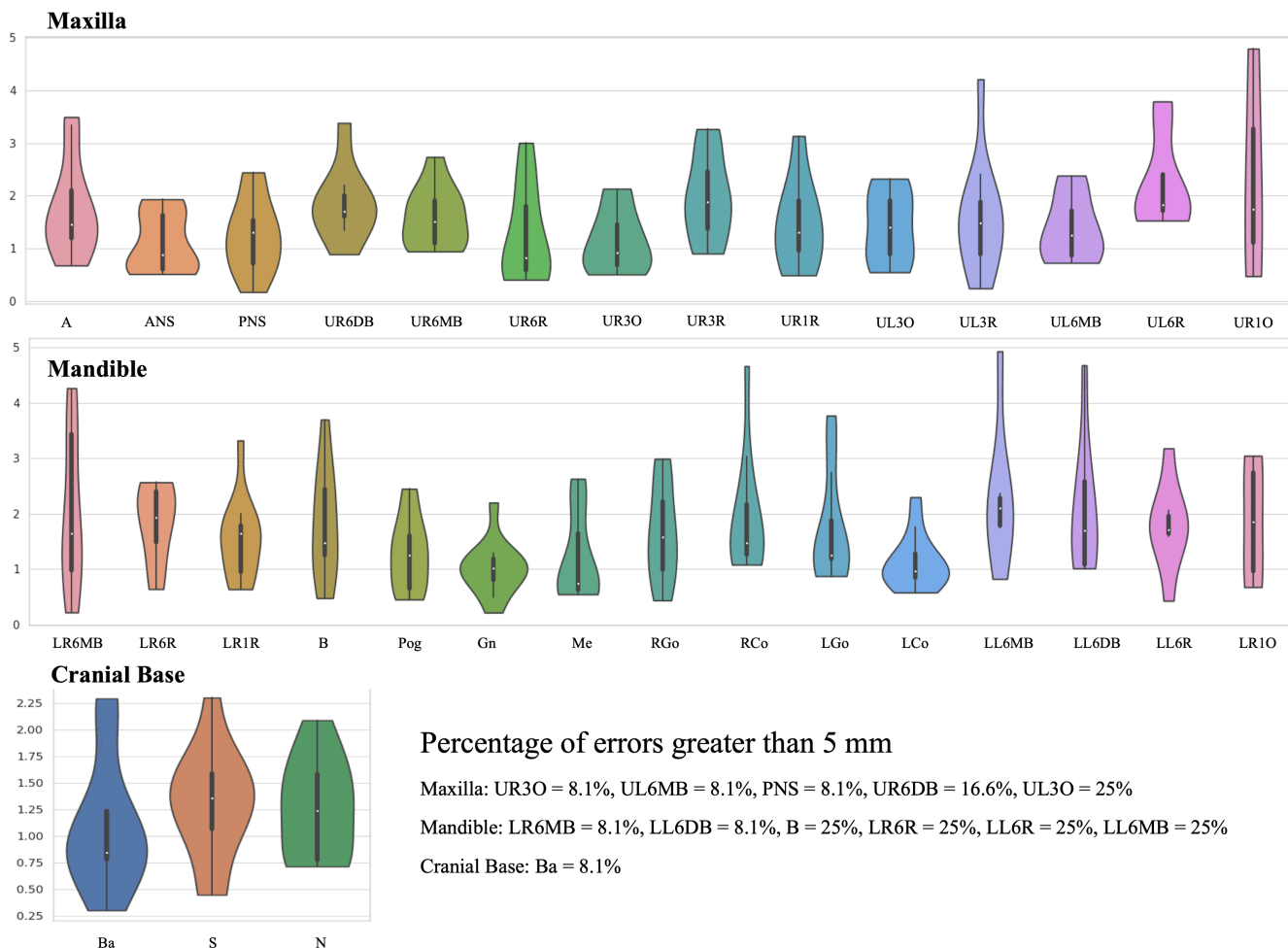


FIGURE 3 Violin plot of the cross-validation result on the cranial base (top) maxilla (middle) and mandible (bottom) and a summary of the fails (top right). Each landmark is represented with its error distribution in mm. The white dot and the black strip are respectively the mean error and the standard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

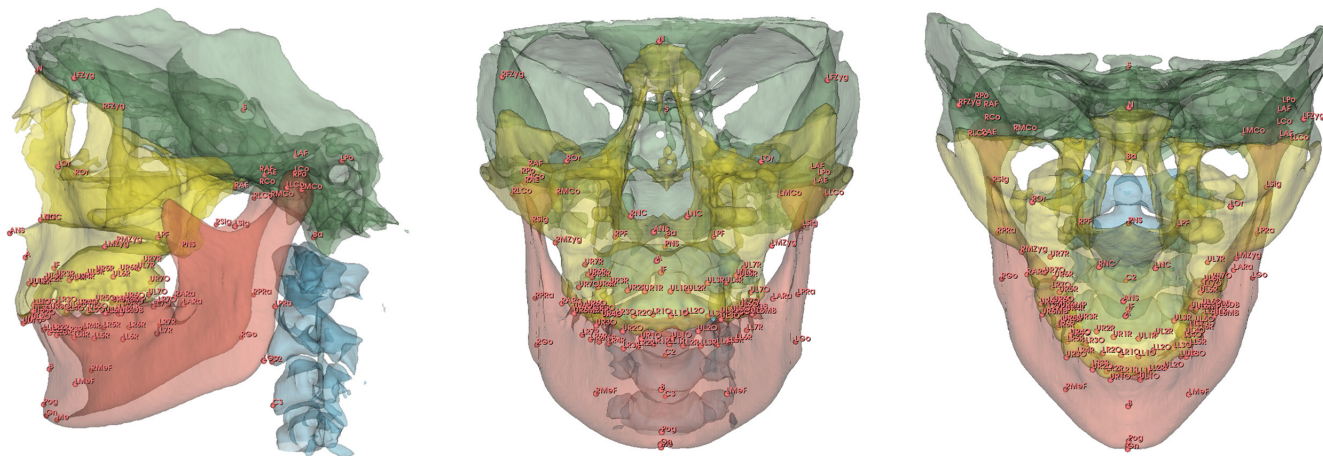


FIGURE 4 All 119 landmark that can be identify by the latest trained agents.

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.<sup>30</sup>

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).<sup>31,32</sup> 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.<sup>31</sup> 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.<sup>32</sup> 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.<sup>22,23</sup> 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. Cebal 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. Kragsskov 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;26:560-567. doi:[10.1111/ocr.12642](https://doi.org/10.1111/ocr.12642)