Gonca et al. BMC Medical Imaging (2024) 24:294
https://doi.org/10.1186/s12880-024-01478-z

BMC Medical Imaging

RESEARCH Open Access

Does the FARNet neural network algorithm accurately identify Posteroanterior cephalometric landmarks?

Merve Gonca[1,5*], İbrahim Şevki Bayrakdar[2,3] and Özer Çelik[3,4]

Abstract

Background We explored whether the feature aggregation and refinement network (FARNet) algorithm accurately identified posteroanterior (PA) cephalometric landmarks.

Methods We identified 47 landmarks on 1,431 PA cephalograms of which 1,177 were used for training, 117 for validation, and 137 for testing. A FARNet-based artificial intelligence (AI) algorithm automatically detected the landmarks. Model effectiveness was calculated by deriving the mean radial error (MRE) and the successful detection rates (SDRs) within 2, 2.5, 3, and 4 mm. The Mann-Whitney U test was performed on the Euclidean differences between repeated manual identifications and AI trials. The direction in differences was analyzed, and whether differences moved in the same or opposite directions relative to ground truth on both the x and y-axis.

Results The AI system (web-based CranioCatch annotation software (Eskişehir, Turkey)) identified 47 anatomical landmarks in PA cephalograms. The right gonion SDRs were the highest, thus 96.4, 97.8, 100, and 100% within 2, 2.5, 3, and 4 mm, respectively. The right gonion MRE was 0.94 ± 0.53 mm. The right condylon SDRs were the lowest, thus 32.8, 45.3, 54.0, and 67.9% within the same thresholds. The right condylon MRE was 3.31 ± 2.25 mm. The AI model's reliability and accuracy were similar to a human expert's. AI was better at four skeleton points than the expert, whereas the expert was better at one skeletal and seven dental points ($P < 0.05$). Most of the points exhibited significant deviations along the y-axis. Compared to ground truth, most of the points in AI and the second trial showed opposite movement on the x-axis and the same on the y-axis.

Conclusions The FARNet algorithm streamlined orthodontic diagnosis.

Keywords Algorithms, Machine learning, Deep learning, Orthodontics

*Correspondence:
Merve Gonca
mervegonca@gmail.com
[1]Present address: Department of Orthodontics, Faculty of Dentistry, Eskisehir Osmangazi University, Eskişehir, Turkey
[2]Department of Oral and Maxillofacial Radiology, Faculty of Dentistry, Eskisehir Osmangazi University, Eskişehir, Turkey
[3]Eskisehir Osmangazi University Center of Research and Application for Computer Aided Diagnosis and Treatment in Health, Eskişehir, Turkey
[4]Department of Mathematics-Computer, Eskisehir Osmangazi University Faculty of Science, Eskişehir, Turkey
[5]Department of Orthodontics, Faculty of Dentistry, Recep Tayyip Erdoğan University, Rize, Turkey

Background

Posteroanterior (PA) cephalometric analysis usefully assesses craniofacial and dentofacial structures, and their growth patterns, in the transverse plane [1, 2]. Such analysis, combined with lateral cephalometric data, facilitate thorough three-dimensional assessments in clinical practice [3, 4].

Cone-beam computed tomography (CBCT) might provide detailed 3D data, but it is not always necessary or feasible in routine orthodontic evaluations, particularly for transverse assessments and facial asymmetry diagnoses [5]. While advanced imaging modalities like CBCT and facial scans have gained popularity due to their three-dimensional capabilities [6], PA cephalometric radiographs remain a vital diagnostic tool in many clinical scenarios due to their lower radiation dose, accessibility, and cost- effectiveness [7]. Thus, PA cephalograms play a central role in orthodontic evaluations and treatment planning, and they are indispensable when planning comprehensive orthodontic treatments [1, 8].

Artificial intelligence (AI) is the ability of a machine to emulate intelligent human activities, thus problem-solving, object identification, and decision-making [9]. In the field of orthodontics, AI has been used to identify landmarks, and to aid diagnosis and data analysis [10]. Cephalometric analysis is required for precise orthodontic diagnosis and treatment planning; this commences with precise localization of cephalometric landmarks, and AI streamlines this process in a rapid and very consistent manner [11].

Traditional two-dimensional radiographs render three dimensions in two; the associated superimpositions may obscure crucial landmarks. These superpositioned structures affect the accuracy of landmark identification. Although PAs are an important diagnostic tool in orthodontic diagnosis, especially in evaluating the transversal direction, the user experience is less than that of lateral cephalometric radiographs. This is more evident at the PA cephalometric landmark, and its errors are greater than those of lateral cephalometric analyses. The expertise and skill levels of the examiner play a major role in the accuracy of landmark identification because it is challenging to recognize landmarks with poor reproducibility. Such errors, thus empirical variances, significantly compromise cephalometric analyses. This may be why PA cephalometric analysis is not routinely performed in orthodontic practice [2, 5, 8]. AI-based programs are applicable, improvable, and time-saving for clinicians. Although the literature contains many studies on lateral cephalometric radiographs, studies examining PA landmarks are limited.

Kim et al. reported that CNNs could effectively compare conventional and automatic cephalometric analyses, demonstrating that AI could not sufficiently reach the clinically acceptable error range of less than 2 mm [8].

Here, we present a multi-stage convolutional neural network (CNN)-based automatic landmark prediction system.

Methods

The study protocol was approved by the Clinical Research Ethics Committee of the Osmangazi University Faculty of Medicine (decision date and decision number: 04.10.2022/22); we adhered to all relevant principles of the Declaration of Helsinki. The PA radiographs were those of patients treated in the Department of Orthodontics, Faculty of Dentistry, Osmangazi University. Additional informed consent was obtained from all individual participants included in the study. The inclusion criteria were:

1) Mixed or permanent dentition,
2) Incomplete or complete facial growth; the use of orthodontic appliances, or/and dental prostheses, or/and surgical screws, or/and plates; with no regard to skeletal asymmetry status;
3) No missing upper or lower permanent incisors or permanent upper or lower first molars, no craniofacial syndrome, and no dentofacial trauma.
4) Either sex.

The exclusion criteria were dentofacial trauma and any craniofacial syndrome or systemic disease. All PA radiographs were captured using the Planmeca ProMax panoramic-cephalometric device (Planmeca Oy, Helsinki, Finland) operating at 64 kVp, 5 mA, and 18 s. One orthodontist with 10 years of experience (M.G.) labeled all 47 PA cephalometric landmarks using web-based CranioCatch annotation software (Eskişehir, Turkey) and a point identification tool (the "ground truth") (Table 1).

The deep learning architecture

We applied a deep learning model based on a specific architecture known as the Feature Aggregation and Refinement Network (FARNet) developed by Ao et al. to detect cephalometric landmarks in posteroanterior (PA) radiographs [12]. The model is built using a type of neural network called a convolutional neural network (CNN), commonly used for analyzing medical images.

Our model has three main components: a backbone, a feature aggregation module, and a refinement module. The backbone is a pre-trained model that has been trained to recognize common patterns in images using a large dataset called ImageNet. This part helps the system identify important features in PA radiographs across different scales or image sizes.

Table 1 Landmark definitions

Landmarks	Definitions
Bilateral skeletal landmarks	
Supraorbitale (So)	The most superior (sagittal perspective) and the middle (frontal perspective) point on the contour of the orbit
Lateroorbitale (Lo)	the intersection of the lateral wall of the orbit and the greater wing of the sphenoid
Frontozygomatic suture (FZ)	The intersection of the frontozygomatic suture and the inner rim of the orbit
Zygomatic arch (ZA)	the most lateral point of the zygomatic arch
Lateral wall of nasal cavity (LWNC)	the most lateral end of the nasal cavity
Jugal process (JG)	The intersection of the tuberosity of the maxilla and zygomatic buttress
Condyle (Co)	The most superior (sagittal perspective) and the middle (frontal perspective) point on the contour of the condyle head
Gonial (Go)	gonion (the most inferior posterior point at the angle of the mandible)
Antegonial notch (AG)	The antegonial notch at the lateral inferior margin of the antegonial protuberances
Midline skeletal landmarks	
Crista Galli (CG)	The middle point of crista galli
Nasal Septum (NS)	The approximated midpoint on the nasal septum between crista galli and anterior nasal spine
Anterior Nasal Spine (ANS)	Located at the anterior nasal spine, between the middle part of the bony portion of the nasal septum and the hard palate
A point (A)	Point A
Menton (Me)	The most inferior point of symphysis of the mandible
Bilateral dentoalveolar landmarks	
Upper central incisor tip (UC1T)	midpoint at the maxillary central incisor at the level of the incisal edges
Upper central incisor root (UC1R)	Apex of maxillary central incisor
Upper central incisor mesial (UC1M)	The most mesial point at the maxillary central incisor at the level of the incisal edges
Lower central incisor tip (LC1T)	midpoint at the mandibular central incisor at the level of the incisal edges
Lower central incisor root (LC1R)	Apex of mandibular central incisor
Lower central incisor mesial (LC1M)	The most mesial point at the mandibular central incisor at the level of the incisal edges
Upper cuspid tip (UC3T)	The cusp tip of the maxillary canine
Lower cuspid tip (LC3T)	The cusp tip of the mandibular canine
Upper molar buccal cusp (UM6BC)	The upper first molar mesiobuccal cusp tip
Upper molar buccal root (UM6BR)	The point of the maxillary first molar root apex
Lower molar buccal cusp (LM6BC)	The lower first molar mesiobuccal cusp tip
Lower molar buccal root (LM6BR)	The point of the mandibular first molar root apex

Next, the feature aggregation module works by gathering details from different parts of the image, both small and large areas, and combines them. This helps the model understand the relationships between various features across different levels of detail. In this process, the system gives preference to clearer, high-resolution parts of the image, which helps in making more accurate predictions. The aggregated image details are then combined into a slightly smaller version of the original image.

Finally, to improve the accuracy of landmark detection, the feature refinement module adjusts the image features and brings them back to the original resolution, ensuring that the model can predict landmarks with better precision (Fig. 1).

Model development

The model was developed in the Dental-AI Laboratory of the Faculty of Dentistry, Osmangazi University, using a Precision 3640 Tower CTO BASE workstation featuring an Intel Xeon W-1250P processor (six cores, 12 Mb cache, core processor frequency 4.1 GHz, max turbo frequency 4.8 GHz, 64 GB DDR4 [4×16 GB] operating at 2,666 MHz of UDIMM ECC memory, a 256-GB SSD SATA, a 2-GB Nvidia Quadro P620, a NVIDIA Tesla V100 graphics card, and a 27-inch IPS LCD monitor of resolution 1,920×1,080 pixels [Dell, TX, USA]). Model development employed the Python open-source programming language (ver. 3.6.1; Python Software Foundation, DE, USA) and the PyTorch library. A total of 1,431 PA radiographs of various sizes, all labeled at 47 points, was employed. Each point corresponded to a row and all labels were saved in a specific order as 47 points in txt format; all images and labels were resized to 640×800 pixels. The dataset was partitioned into three subsets: Training, validation, and testing (1,177, 117, and 137 radiographs with 47-point labels). No testing data were used for training or validation. The Adam Optimizer was employed for training that ran over 300 epochs using a

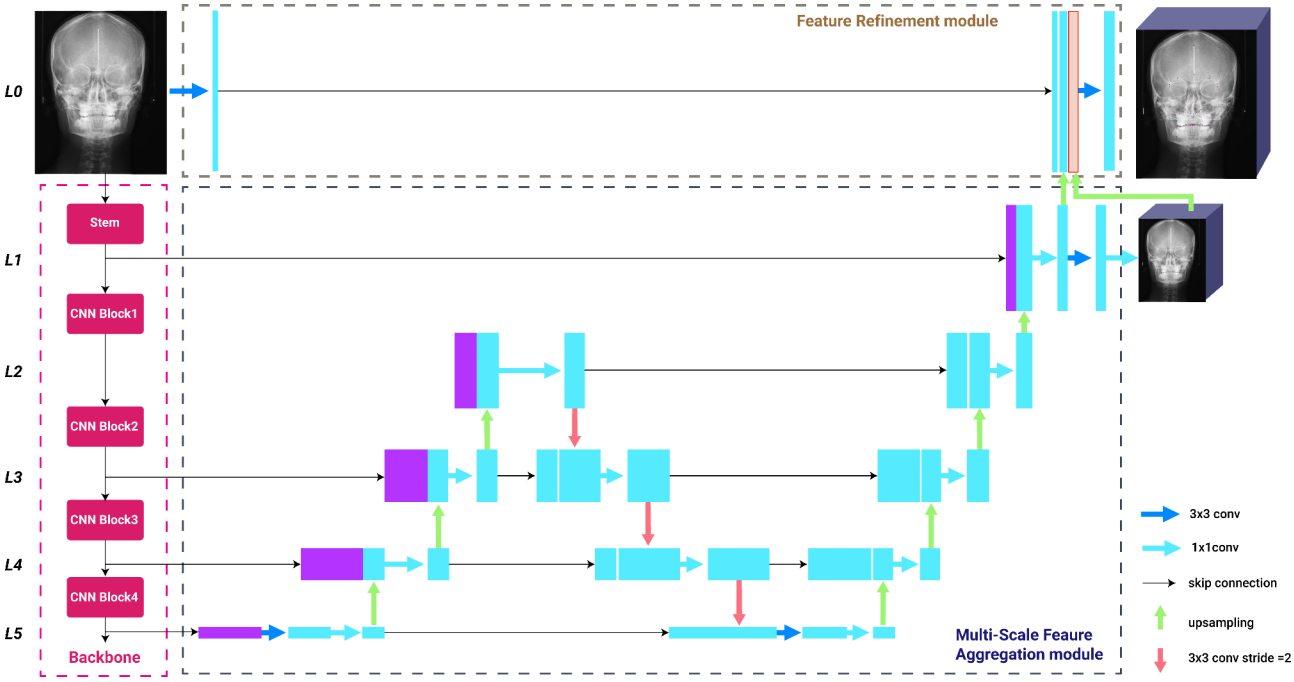

Fig. 1 The architecture of the feature aggregation and refinement network (FARNet). This includes a backbone (pink dashed box), a multi-scale feature aggregation (MSFA) module (black dashed box), and a feature refinement (FR) module (brown dashed box). The feature level labels (*L0*, *L1*, *L2*, *L3*, *L4*, *L5*) are shown on the left. The spatial resolutions of all feature maps at the same horizontal levels are identical

CNN-based deep learning approach implemented in PyTorch. The learning rate was 0.0001 (Fig. 2).

Performance evaluation

In this study, the performance of our AI model was evaluated by comparing its predicted landmark positions (AI detection) with those manually identified by a human expert (manual detection), which serve as the "gold standard" for cephalometric analysis. The primary metric used for comparison was the point-to-point error, which is the average absolute distance between the AI-predicted landmarks and the manually determined landmarks across the entire test dataset. We calculated the mean radial error (MRE) and standard deviation (SD) for each landmark to quantify these errors. The radial error (R) for any given landmark was defined as the distance between the AI-predicted position and the corresponding manual position using the formula:

$$\mathrm{R} = \sqrt{\Delta \mathrm{x}^2 + \Delta \mathrm{y}^2}$$

In this formula, Δx represents the horizontal distance, and Δy represents the vertical distance between the AI-predicted and manually identified landmark positions.

The MRE was calculated as the average of these radial errors across all test cases:

$$\mathrm{MRE} = \sum \left(\sqrt{\left(\Delta \mathrm{x}^2 + \Delta \mathrm{y}^2\right)}\right) / \mathrm{n}$$

Where n is the total number of test samples, this metric measures the overall accuracy of the AI model compared to manual detection.

We also calculated the standard deviation (SD) of the errors to assess the variability in the model's performance. The SD measures the spread of individual landmark detection errors around the mean radial error (MRE). It was computed as follows:

$$\mathrm{SD} = \sqrt{\left[\left(\sum \left(\sqrt{\left[\Delta \mathrm{x}^2 + \Delta \mathrm{y}^2\right]}\right)/\mathrm{n}\right)^2 - \mathrm{MRE}^2\right]/(\mathrm{n}-1)}$$

We derived successful detection rates (SDRs), thus the percentages of estimated points within 2, 2.5, 3, and 4 mm of the actual locations. For each PA cephalometric landmark, the AI-determined position was compared to the ground truth. If the difference did not exceed a defined value *d*, AI localization was viewed as successful, and the SDR was calculated.

The direction in which the deviations of the points in manual identification 1 (M1) -AI and manual identification 1-manual identification 2 (M2) were evident (x or y) was examined. It was evaluated whether the movements in M2 and AI were in the same/opposite direction on the x and y axis relative to M1.

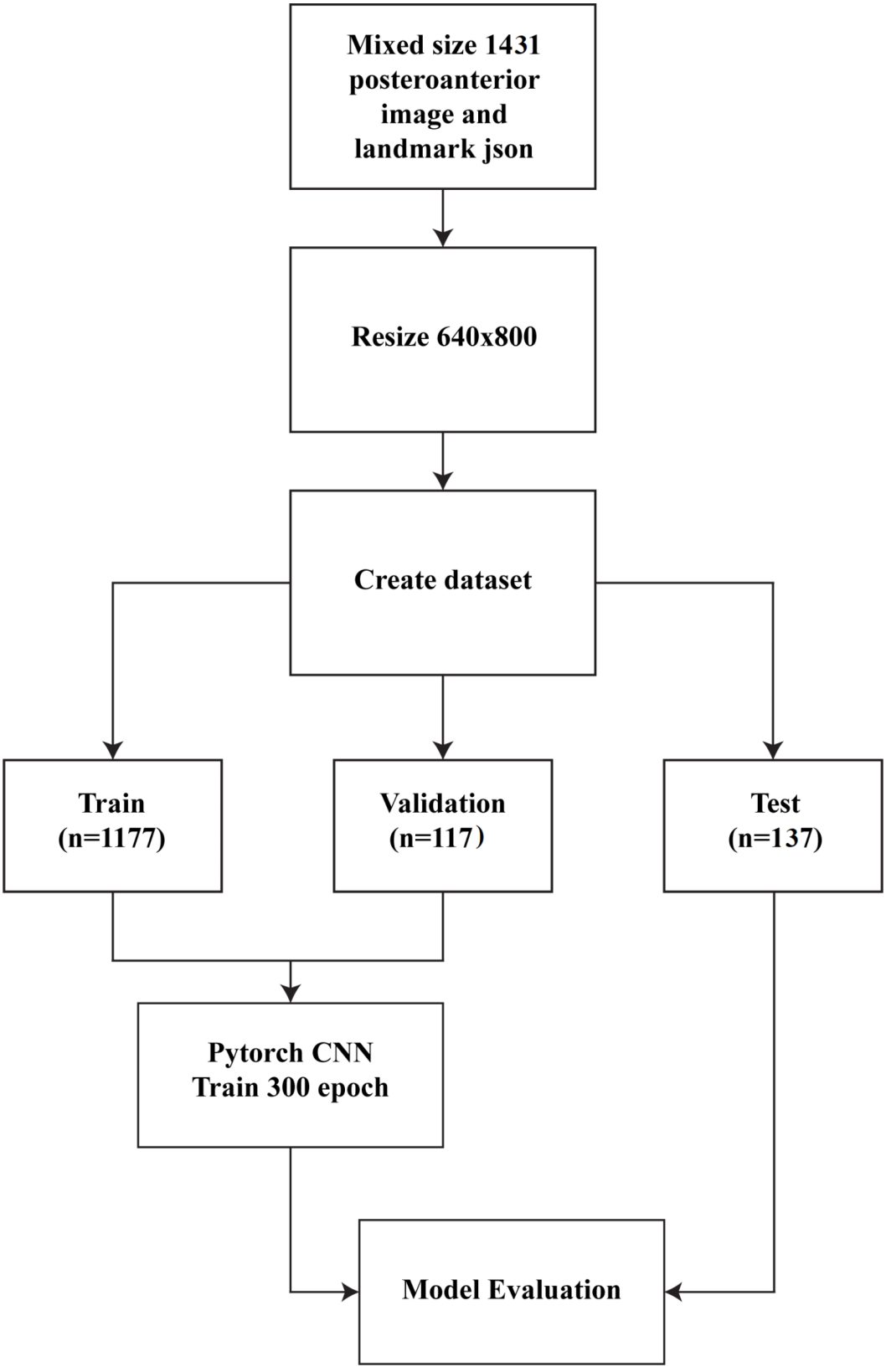

Fig. 2 The AI model pipeline for automatic landmark detection (JSON: Java Script Object Notation)

Measurement error

The same orthodontist with ten years of experience repeated the landmark localizations in the all test images one month after the initial evaluation. The distances to the right and upper edges of the radiographs were calculated in pixels to determine the coordinates of the distances in the x and y axes. The distances in pixels of landmarks along the x-y axes were used to calculate intraclass correlation coefficients (ICC) for intra-rater reliability.

Statistical analysis

The normality of data distribution was evaluated using the Kolmogorov-Smirnov test. The Mann Whitney U test and intraclass reliability tests were performed using IBM SPSS (version 21.0) software (SPSS, Chicago, IL, USA).

The Mann Whitney U test was used to compare the Euclidean differences between manual identification 1-auto-identification, and Euclidean differences between manual identification 1–2. The P value <0.05 was considered to be statistically significant.

Results

Table 2 presents the intra-class correlation coefficient (ICC) values for the 47 landmarks detected by the first and second manual identifications. ICC values less than 0.5 are indicative of poor reliability, values between 0.5 and 0.75 indicate moderate reliability, values between 0.75 and 0.9 indicate good reliability, and values greater than 0.9 indicate excellent reliability [13] The ICC values for all landmarks were above 0.75, except for the x-coordinate of the lower cuspid tip (LC3T) L-R. These points had moderate reliability.

The AI system (web-based CranioCatch annotation software (Eskişehir, Turkey)) identified 47 anatomical landmarks in PA cephalometric radiographs. The AI- (blue), human-located 1 (gold standard; red), and human-located 2 (green) locations were superimposed on all images. (Figures 3 and 4). Tables 3 and 4 summarize the SDR and MRE values of all test landmarks and show them in the supplementary content.

Between manual identifications 1 and 2, the right gonion point exhibited the highest 2-, 2.5-, 3-, and 4-mm SDRs of 94.2, 97.8, 99.3, and 100%, respectively. The average MRE (±SD) was 1.81±1.61 mm. The lowest MREs (1.01±0.71) were that for the both right upper central incisor tip, and the right lower central incisor tip. The left condyle point exhibited the lowest SDRs (36.5, 48.2, 53.3, and 67.2%) at 2, 2.5, 3, and 4 mm; and the highest MRE±SD was 3.37±2.34.

Between manual identification 1 and AI, the right gonion point exhibited the highest 2-, 2.5-, 3-, and 4-mm SDRs of 96.4, 97.8, 100.0, and 100.0%, respectively. The average MRE (±SD) was 1.84±1.60 mm. The lowest AI MRE (0.94±0.53) was that for the right gonion. The right condyle point exhibited the lowest SDRs (32.8, 45.3, 54.0, and 67.9%) at 2, 2.5, 3, and 4 mm; and the highest MRE±SD was 3.31±2.25.

Artificial intelligence showed significantly higher accuracy than manual identification of the left and right lateroorbitale points, left frontozygomatic suture, and left antegonial notch points ($P=0.001$, $P=0.005$, $P=0.015$, and $P=0.011$, respectively). Although AI showed low accuracy in the identification of left jugal process, right upper central incisor tip, right lower central incisor tip, right upper cuspid tip, left and right upper molar buccal cusp, left and right lower molar buccal cusp points ($P=0.035$, $P=0.002$, $P=0.002$, $P=0.007$, $P=0.007$, $P<0.001$, $P=0.006$ and $P=<0.001$, respectively) (Table 4). However, there was no difference in accuracy except for the measurements above between AI and manual identification ($P>0.05$).

Most of the points in both M1-AI and M1-M2 exhibited significant deviations along the y-axis. When the differences M1-AI and M1-M2 were compared, the directions of seven points ((R) Frontozygomatic suture, (R) Zygomatic arch, (L, R) Gonial, (R) Upper central incisor tip, and (L, R) Lower molar buccal cusp) were found to differ. Compared to M1, most of the points in AI and M2 showed opposite movement on the x-axis and the same on the y-axis.

Discussion

Although PA cephalometric analysis allows comprehensive evaluation of cranial-dentofacial features, more anatomical structures are superimposed and layered than in lateral cephalograms, rendering precise landmark identification difficult. Accurate human identification requires high-level expertise, perhaps explaining why PA cephalometric analysis is not common in orthodontic settings [1, 14].

We present several novel findings: (1) We assessed the automated identification accuracy of 47 PA cephalometric landmarks. (2) We built a comprehensive algorithm featuring a multi-stage CNN. (3) The average SDRs for AI were 68.2, 78.4, 85.3, and 92.5% within 2.0, 2.5, 3.0, and 4.0 mm, respectively (4). The mean point-to-point error was 1.84 mm, which is clinically acceptable. Several studies have found that deep learning algorithms accurately and rapidly detect landmarks with a precision that attained 2.0 mm. Our SDRs for the 2- and 4-mm thresholds exceeded 70 and 90% respectively [15–17]. However, most prior studies used lateral cephalograms. We calculated the 2-, 2.5-, 3-, and 4-mm SDRs of 47 landmarks in PA cephalograms [15, 18–21].

To the best of our knowledge, few studies have used AI to determine automatically landmarks in PA radiographs [8, 22, 23]. The multi-center study of Gil et al.

Table 2 Intraobserver agreement of the first and second trial

	X-axis			Y-axis		
Anatomik Landmarks	**ICC**	**Lower 95% CI**	**Upper 95% CI**	**ICC**	**Lower 95% CI**	**Upper 95% CI**
Bilateral skeletal landmarks						
Supraorbitale (So), L	0.885	0.843	0.917	0.994	0.992	0.996
Supraorbitale (So), R	0.895	0.850	0.926	0.994	0.992	0.996
Lateroorbitale (Lo), L	0.974	0.964	0.982	0.951	0.887	0.974
Lateroorbitale (Lo), R	0.984	0.977	0.989	0.959	0.931	0.974
Frontozygomatic suture (FZ), L	0.946	0.923	0.962	0.957	0.934	0.971
Frontozygomatic suture (FZ), R	0.957	0.937	0.970	0.953	0.922	0.970
Zygomatic arch (ZA), L	0.957	0.933	0.972	0.918	0.888	0.941
Zygomatic arch (ZA), R	0.957	0.939	0.969	0.910	0.877	0.935
Lateral wall of nasal cavity (LWNC), L	0.902	0.770	0.949	0.949	0.928	0.964
Lateral wall of nasal cavity (LWNC), R	0.953	0.934	0.967	0.961	0.942	0.973
Jugal process (JG), L	0.933	0.908	0.952	0.946	0.926	0.961
Jugal process (JG), R	0.949	0.930	0.964	0.950	0.931	0.964
Condyle (Co), L	0.924	0.893	0.946	0.812	0.746	0.862
Condyle (Co), R	0.931	0.900	0.952	0.841	0.758	0.893
Gonial (Go), L	0.991	0.987	0.994	0.99	0.987	0.993
Gonial (Go), R	0.991	0.984	0.995	0.992	0.989	0.994
Antegonial notch (AGL), L	0.972	0.946	0.984	0.991	0.983	0.995
Antegonial notch (AGL), R	0.980	0.972	0.986	0.992	0.989	0.994
Midline skeletal landmarks						
Crista Galli (CG)	0.938	0.913	0.956	0.924	0.895	0.945
Nasal Septum (NS)	0.953	0.935	0.966	0.933	0.908	0.952
Anterior Nasal Spine (ANS)	0.957	0.934	0.971	0.970	0.958	0.979
A point (A)	0.958	0.936	0.972	0.946	0.925	0.961
Menton (Me)	0.902	0.861	0.930	0.994	0.992	0.996
Bilateral dentoalveolar landmarks						
Upper central incisor tip (UC1T), L	0.969	0.930	0.983	0.991	0.987	0.994
Upper central incisor tip (UC1T), R	0.973	0.962	0.981	0.993	0.990	0.995
Upper central incisor root (UC1R), L	0.948	0.927	0.962	0.973	0.962	0.980
Upper central incisor root (UC1R), R	0.953	0.934	0.967	0.974	0.962	0.981
Upper central incisor mesial (UC1M), L	0.976	0.965	0.983	0.989	0.980	0.994
Upper central incisor mesial (UC1M), R	0.979	0.961	0.988	0.988	0.977	0.993
Lower central incisor tip (LC1T), L	0.943	0.908	0.963	0.972	0.961	0.980
Lower central incisor tip (LC1T), R	0.941	0.919	0.958	0.956	0.936	0.969
Lower central incisor root (LC1R), L	0.971	0.959	0.979	0.973	0.960	0.982
Lower central incisor root (LC1R), R	0.974	0.963	0.981	0.977	0.967	0.984
Lower central incisor mesial (LC1M), L	0.952	0.932	0.966	0.963	0.930	0.978
Lower central incisor mesial (LC1M), R	0.942	0.919	0.958	0.964	0.926	0.980
Upper cuspid tip (UC3T), L	0.874	0.827	0.909	0.954	0.933	0.969
Upper cuspid tip (UC3T), R	0.841	0.784	0.884	0.957	0.941	0.970
Lower cuspid tip (LC3T), L	0.704	0.609	0.780	0.964	0.951	0.974
Lower cuspid tip (LC3T), R	0.715	0.623	0.788	0.968	0.955	0.977
Upper molar buccal cusp (UM6BC), L	0.914	0.861	0.945	0.967	0.954	0.977
Upper molar buccal cusp (UM6BC), R	0.916	0.869	0.945	0.975	0.966	0.982
Upper molar buccal root (UM6BR), L	0.788	0.617	0.873	0.943	0.920	0.959
Upper molar buccal root (UM6BR), R	0.880	0.836	0.913	0.955	0.938	0.968
Lower molar buccal cusp (LM6BC), L	0.877	0.832	0.911	0.959	0.943	0.971
Lower molar buccal cusp (LM6BC), R	0.921	0.889	0.944	0.979	0.971	0.985
Lower molar buccal root (LM6BR), L	0.862	0.757	0.916	0.938	0.914	0.956
Lower molar buccal root (LM6BR), R	0.922	0.892	0.944	0.962	0.944	0.973

ICC; Intraclass Correlation Coefficient

CI; Confidence Interval

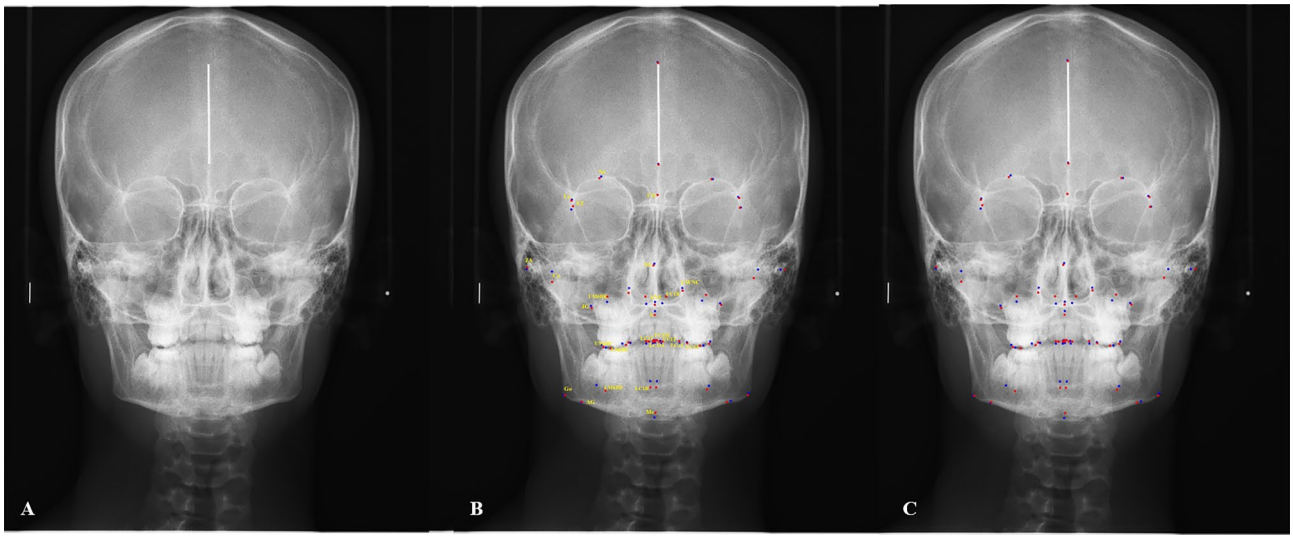

Fig. 3 Automatic detection of cephalometric points by the AI model. (**A**) Original image. (**B**) Demonstration of landmark detection. (**C**) A comparison of landmark detection by an expert in the first trial and AI. Red: The expert locations in the first trial. Blue: The AI locations

[22] examined 16 landmarks. The single-center work of Kim et al. [8] synthesized PAs from cone beam computed tomographic data and examined 22 PA cephalometric landmarks. Lee et al. [23] examined 19 PA landmarks in a single-center study on directly observed PAs. Kim et al. [8] used only the MRE to assess accuracy; the other two groups employed both the MRE and SDR [22, 23]. Our present single- center investigation of directly observed PAs derived both the MREs and SDRs for 47 PA cephalometric landmarks.

The average SDR within the 2 mm threshold was 68.2%, similar to the 67.5% of Lee et al. [23]; the figure of Gil et al. [22] came to 83.3%. As some landmarks are identified less accurately than others, the same numbers of points in particular locations must be used when comparing the average SDRs. Our landmark numbers were 2–2.5-fold more than those of the 47 landmarks of the cited Works [22, 23]. Hence, a comparison of the average SDRs may not reliably assess model efficacies.

The MREs of all PA cephalometric landmarks were under 4 mm for AI. However, the point-to-point errors of the (L) jugal process, (L, R) condyles, crista galli, (L, R) upper cuspids, and all (L, R) molar measurements exceeded 2 mm. The lowest MRE and the highest SDR was that of the right gonial point. The highest MRE and the lowest SDR was the right condyle point.

The CNN model replicates the landmark identification performed by human examiners when identifying the landmarks of test scans. Thus, difficulties encountered by humans affect model accuracy; the AI predictions mirror challenges encountered by observers [8]. Landmarks that lie on pronounced curves or where two curves converge tend to be more readily discernible than points in flat regions or on gentle curves. Points in high-contrast regions are more obvious than those in low-contrast regions. Identification is difficult when other structures are superimposed on a landmark [8, 14, 24]. Auto-identification was associated with high MREs and low SDRs for the condyle (L-R), and jugal processes (L), in contrast to the skeletal points. These results are similar to those of prior studies; auto-identification errors were higher for the the condyles because these overlapped with other anatomical structures [8, 23]. Our (L, R) condyle MREs were similar to those of Kim et al. [8] (4.05±2.44, 4.24±2.21 mm), and Lee et al. [23] (3.16±1.88, 3.47±2.09 mm). The condyle points (3.21±2.03, 3.31±2.25 mm) exhibited the highest MREs of all PA skeletal landmarks.

The landmarks that are most difficult to define lie on curved trajectories, within regions of poor contrast, or overlap with other structures [25]. Here, certain midline landmarks contrasted poorly, rendering them difficult to distinguish. These landmarks frequently overlapped or were obscured by adjacent anatomical structures. Note that the crista galli serves as the central point of the asymmetric quadrilateral construct [2]. We found that the midline landmarks exhibited mean MREs of less than 2 mm, with the exception of the crista galli (2.03±1.77 mm). Our crista galli MREs were higher than those of Kim et al. [8] (1.33±1.59 mm) and Gil et al. [22] (1.89±1.61 mm) but similar to those of Lee et al. [23] (2.57±1.63 mm). We included images of individuals who wore appliances, or who had brackets or impacted canines, perhaps explaining why the MREs of the canines and molars sometimes exceeded 2 mm. maxillary and mandibular molar root MREs over 2 mm may reflect multiroot overlaps.

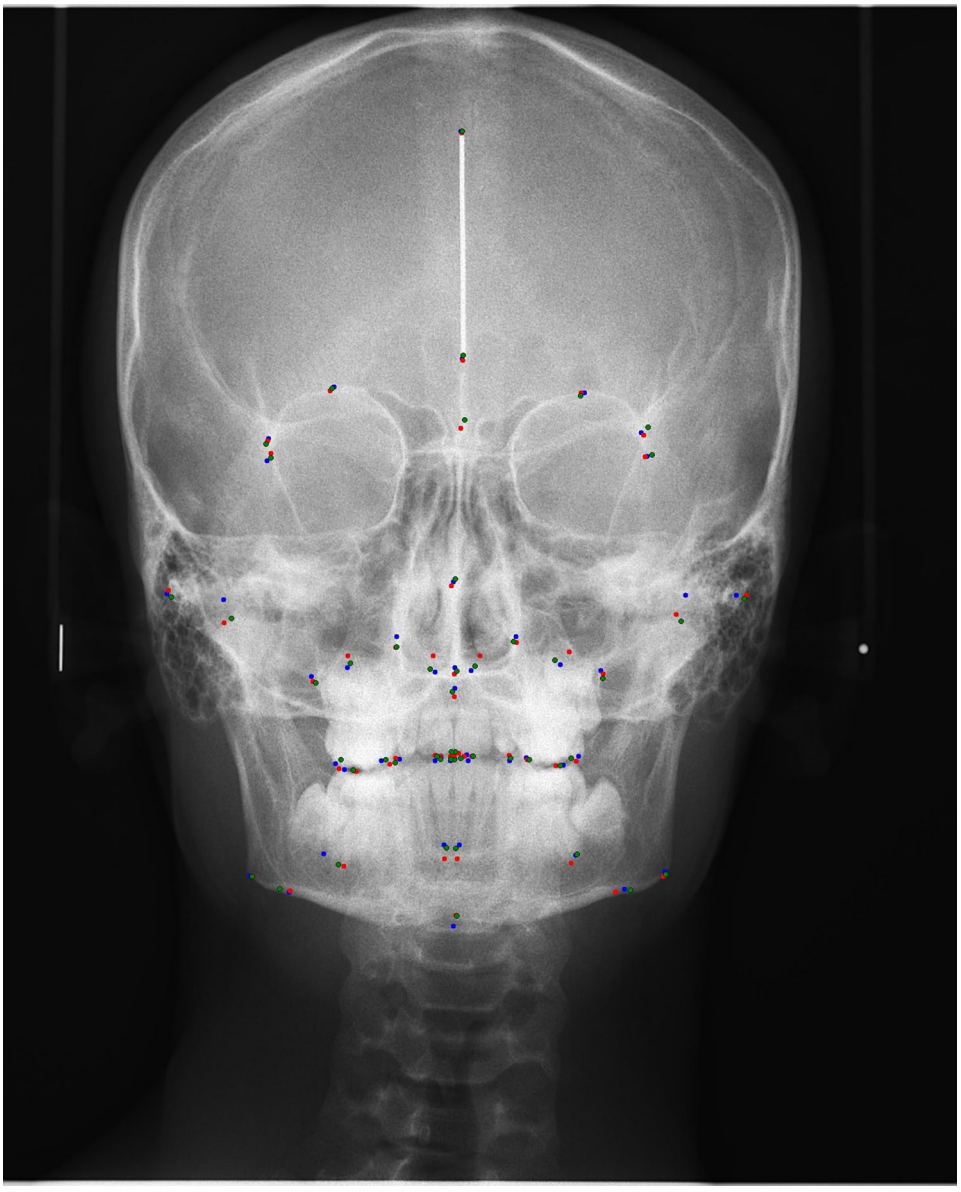

Fig. 4 Demonstration of landmark detection Red: The expert locations in the first trial. Green: The expert locations in the second trial. Blue: The AI locations

Gil et al. [22] and Lee et al. [23] evaluated the most lateral point of the crown. Kim et al. [8] focused on the molar mesiobuccal cusp tip, as did we. In this work, the molar teeth MREs exceeded 2 mm, as also reported by Kim et al. [8]. Molar roots on PA radiographs were previously assessed only by Gil et al. [22] The reason why the dental MREs in the cited works were lower than ours may be because the test sets included only pretreatment radiographs [22].

In clinical practice, patients often undergo cephalometric evaluations while wearing orthodontic appliances or after undergoing surgeries involving screws and plates. Excluding these patients would not accurately represent the diversity of routine orthodontic treatment planning cases. Including these images allows to evaluate the robustness and adaptability of the AI-based program to handle complex cases, including those with increased radiopacity due to appliances or surgical hardware [26].

Research has shown that AI-based systems, particularly those using convolutional neural networks (CNNs), can identify landmarks even in obstructions or noise, such as those caused by orthodontic hardware or surgical materials [26, 27]. The ability of AI systems to adapt to such conditions is a key factor in their utility. While a more uniform sample may have improved the algorithm's baseline accuracy, it would have limited the study's generalizability. Including these patients provided a more realistic evaluation of the program's performance and ensured

Table 3 The SDR value of landmarks obtained from test data

Anatomik Landmarks	(Manual 1-AI)				(Manual 1-Manual 2)			
	<=2.0	<=2.5	<=3.0	<=4.0	<=2.0	<=2.5	<=3.0	<=4.0
Supraorbitale (So), L	80.3	92.0	96.4	97.8	79.6	88.3	94.9	99.3
Supraorbitale (So), R	75.2	88.3	92.7	97.8	65.0	78.8	87.6	95.6
Lateroorbitale (Lo), L	83.9	94.2	96.4	98.5	67.9	81.0	88.3	95.6
Lateroorbitale (Lo), R	81.0	90.5	94.9	100.0	63.5	80.3	86.9	95.6
Frontozygomatic suture (FZ), L	78.1	86.9	94.2	99.3	67.2	78.8	85.4	96.4
Frontozygomatic suture (FZ), R	68.6	80.3	89.1	95.6	59.1	76.6	83.9	93.4
Zygomatic arch (ZA), L	76.6	84.7	89.8	91.2	72.3	82.5	86.9	91.2
Zygomatic arch (ZA), R	70.1	78.8	86.1	90.5	67.2	78.1	83.9	86.9
Lateral wall of nasal cavity (LWNC), L	61.3	72.3	84.7	94.9	66.4	74.5	81.8	92.0
Lateral wall of nasal cavity (LWNC), R	68.6	73.7	83.2	94.9	73.0	81.8	88.3	94.2
Jugal process (JG), L	57.7	70.8	80.3	92.7	69.3	76.6	80.3	89.8
Jugal process (JG), R	70.8	78.8	83.2	94.9	66.4	73.0	81.8	91.2
Condyle (Co), L	32.8	45.3	52.6	68.6	36.5	48.2	53.3	67.2
Condyle (Co), R	32.8	45.3	54.0	67.9	38.7	48.9	55.5	70.1
Gonial (Go), L	92.7	97.8	98.5	99.3	90.5	97.1	97.8	100.0
Gonial (Go), R	96.4	97.8	100.0	100.0	94.2	97.8	99.3	100.0
Antegonial notch (AGL), L	95.6	98.5	98.5	100.0	81.0	90.5	94.9	99.3
Antegonial notch (AGL), R	93.4	96.4	99.3	100.0	86.9	93.4	97.1	100.0
Crista Galli (CG)	62.8	73.0	79.6	87.6	59.1	68.6	76.6	86.1
Nasal Septum (NS)	70.8	79.6	86.9	92.7	67.9	78.8	85.4	92.7
Anterior Nasal Spine (ANS)	76.6	84.7	90.5	97.8	81.8	90.5	93.4	96.4
A point (A)	69.3	81.0	85.4	91.2	65.7	73.7	83.2	92.0
Menton (Me)	82.5	92.0	98.5	98.5	80.3	92.7	96.4	99.3
Upper central incisor tip (UC1T), L	87.6	94.2	97.8	100.0	92.7	96.4	97.8	99.3
Upper central incisor tip (UC1T), R	90.5	95.6	99.3	99.3	88.3	94.9	98.5	100.0
Upper central incisor root (UC1R), L	68.6	86.9	93.4	97.8	69.3	82.5	90.5	95.6
Upper central incisor root (UC1R), R	78.8	87.6	94.9	97.8	67.9	76.6	89.1	97.8
Upper central incisor mesial (UC1M), L	84.7	94.9	97.1	100.0	85.4	91.2	95.6	100.0
Upper central incisor mesial (UC1M), R	89.8	93.4	96.4	98.5	85.4	91.2	94.2	99.3
Lower central incisor tip (LC1T), L	71.5	83.9	91.2	97.8	71.5	81.0	86.9	92.7
Lower central incisor tip (LC1T), R	67.9	81.8	89.8	96.4	71.5	78.8	86.9	93.4
Lower central incisor root (LC1R), L	70.1	81.0	90.5	97.8	67.2	78.8	88.3	94.9
Lower central incisor root (LC1R), R	71.5	83.2	90.5	97.8	75.9	83.2	89.1	95.6
Lower central incisor mesial (LC1M), L	63.5	79.6	87.6	95.6	58.4	70.8	82.5	89.1
Lower central incisor mesial (LC1M), R	67.2	81.0	88.3	96.4	58.4	73.0	80.3	88.3
Upper cuspid tip (UC3T), L	50.4	62.0	73.0	85.4	53.3	67.2	74.5	88.3
Upper cuspid tip (UC3T), R	49.6	64.2	73.0	85.4	61.3	72.3	81.0	86.9
Lower cuspid tip (LC3T), L	62.8	76.6	82.5	90.5	66.4	78.1	84.7	94.9
Lower cuspid tip (LC3T), R	67.9	75.9	82.5	89.8	69.3	78.8	84.7	92.7
Upper molar buccal cusp (UM6BC), L	61.3	73.7	82.5	89.1	68.6	82.5	87.6	92.7
Upper molar buccal cusp (UM6BC), R	52.6	62.0	73.0	89.1	70.1	80.3	86.9	92.0
Upper molar buccal root (UM6BR), L	32.8	48.9	63.5	76.6	38.7	52.6	63.5	79.6
Upper molar buccal root (UM6BR), R	38.7	51.8	67.2	86.1	48.9	59.9	67.9	84.7
Lower molar buccal cusp (LM6BC), L	55.5	72.3	80.3	88.3	68.6	75.9	83.9	91.2
Lower molar buccal cusp (LM6BC), R	59.9	68.6	81.8	89.8	70.8	81.0	89.1	92.7
Lower molar buccal root (LM6BR), L	41.6	50.4	61.3	76.6	44.5	56.9	65.0	79.6
Lower molar buccal root (LM6BR), R	44.5	53.3	58.4	74.5	48.9	59.9	68.6	83.2
Average	68.2	78.4	85.3	92.5	68.1	78.2	84.7	92.1

SDR, Success Detection Rate

Table 4 The MRE values of landmarks obtained from test data

Anatomik landmarks	Diff. M1-AI			Diff. M1-M2			*P*	More accurate		(Diff. M1-M2) - (Diff. M1-AI)	Diff. direction	
	MRE	SD	Distrubition	MRE	SD	Distrubition						
Bilateral skeletal landmarks								Manual	AI		X	Y
Supraorbitale (So), L	1.31	0.84	—	1.38	0.92	—	0.674			0.06	OD	SD
Supraorbitale (So), R	1.52	0.88	—	1.75	1.18	—	0.222			0.23	OD	SD
Lateroorbitale (Lo), L	1.27	0.84	|	1.68	1.06	|	**0.001**		✓	0.41	OD	OD
Lateroorbitale (Lo), R	1.30	0.81	|	1.68	1.06	|	**0.005**		✓	0.38	OD	OD
Frontozygomatic suture (FZ), L	1.52	0.90	|	1.82	1.07	|	**0.015**		✓	0.30	OD	OD
Frontozygomatic suture (FZ), R	1.73	1.04	—	1.97	1.32	|	0.183			0.24	SD	OD
Zygomatic arch (ZA), L	1.73	1.79	O	1.86	1.90	O	0.434			0.13	OD	OD
Zygomatic arch (ZA), R	1.84	1.77	|	2.22	2.61	O	0.366			0.38	SD	OD
Lateral wall of nasal cavity (LWNC), L	1.85	1.23	|	1.80	1.41	|	0.263			-0.05	OD	SD
Lateral wall of nasal cavity (LWNC), R	1.72	1.25	|	1.62	1.19	|	0.532			-0.10	SD	OD
Jugal process (JG), L	2.17	2.01	|	1.93	2.08	|	**0.035**	✓		-0.24	OD	OD
Jugal process (JG), R	1.94	2.22	|	1.97	1.93	|	0.616			0.03	OD	OD
Condyle (Co), L	3.21	2.03	|	3.37	2.34	|	0.862			0.16	OD	SD
Condyle (Co), R	3.31	2.25	|	3.22	2.47	|	0.484			-0.09	OD	SD
Gonial (Go), L	1.06	0.70	|	1.07	0.66	O	0.928			0.01	SD	OD
Gonial (Go), R	0.94	0.53	|	1.04	0.59	O	0.173			0.10	OD	SD
Antegonial notch (AGL), L	0.97	0.58	O	1.25	0.84	O	**0.011**		✓	0.29	SD	OD
Antegonial notch (AGL), R	0.97	0.61	O	1.10	0.75	O	0.282			0.13	OD	SD
Midline skeletal landmarks												
Crista Galli (CG)	2.03	1.77	|	2.18	1.79	|	0.336			0.14	OD	OD
Nasal Septum (NS)	1.71	1.26	|	1.80	1.34	|	0.598			0.09	OD	SD
Anterior Nasal Spine (ANS)	1.41	1.09	|	1.30	1.07	|	0.383			-0.11	OD	SD
A point (A)	1.80	1.36	|	1.81	1.33	|	0.904			0.01	OD	SD
Menton (Me)	1.33	1.06	|	1.40	0.97	|	0.353			0.08	OD	SD
Bilateral dentoalveolar landmarks												
Upper central incisor tip (UC1T), L	1.16	0.69	O	1.12	0.73	|	0.603			-0.04	OD	SD
Upper central incisor tip (UC1T), R	1.19	0.66	|	1.01	0.71	|	**0.002**	✓		-0.18	SD	SD
Upper central incisor root (UC1R), L	1.62	0.90	|	1.58	1.08	|	0.332			-0.04	SD	SD
Upper central incisor root (UC1R), R	1.51	0.85	|	1.65	0.99	|	0.411			0.14	OD	SD
Upper central incisor mesial (UC1M), L	1.22	0.74	|	1.20	0.81	|	0.572			-0.02	OD	OD
Upper central incisor mesial (UC1M), R	1.18	0.79	|	1.19	0.91	|	0.611			0.01	OD	OD
Lower central incisor tip (LC1T), L	1.16	0.69	|	1.12	0.73	|	0.603			-0.04	OD	SD
Lower central incisor tip (LC1T), R	1.19	0.66	|	1.01	0.71	|	**0.002**	✓		-0.18	SD	SD
Lower central incisor root (LC1R), L	1.65	1.05	|	1.64	1.34	|	0.27			-0.01	OD	OD
Lower central incisor root (LC1R), R	1.56	1.02	|	1.53	1.21	|	0.3			-0.03	OD	OD
Lower central incisor mesial (LC1M), L	1.78	1.13	|	2.01	1.45	|	0.4			0.23	OD	SD
Lower central incisor mesial (LC1M), R	1.70	1.07	|	1.99	1.45	|	0.301			0.28	OD	SD

Table 4 (continued)

Anatomik landmarks	Diff. M1-AI			Diff. M1-M2			*P*	More accurate		(Diff. M1-M2) - (Diff. M1-AI)	Diff. direction	
	MRE	SD	Distrubition	MRE	SD	Distrubition		Manual	AI		X	Y
Bilateral skeletal landmarks												
Upper cuspid tip (UC3T), L	2.64	2.47	|	2.39	2.18	|	0.374			-0.25	SD	SD
Upper cuspid tip (UC3T), R	2.59	2.31	|	2.21	2.36	|	**0.007**	✓		-0.38	SD	OD
Lower cuspid tip (LC3T), L	1.61	0.99	—	1.71	1.35	—	0.74			0.10	OD	SD
Lower cuspid tip (LC3T), R	1.82	1.68	—	1.85	1.94	—	0.411			0.03	SD	SD
Upper molar buccal cusp (UM6BC), L	2.07	1.61	|	1.71	1.59	|	**0.007**	✓		-0.36	OD	SD
Upper molar buccal cusp (UM6BC), R	2.31	1.68	|	1.65	1.52	|	**< 0.001**	✓		-0.66	SD	SD
Upper molar buccal root (UM6BR), L	2.95	1.72	—	2.69	1.76	—	0.196			-0.26	OD	OD
Upper molar buccal root (UM6BR), R	2.62	1.71	|	2.40	1.54	|	0.311			-0.22	OD	OD
Lower molar buccal cusp (LM6BC), L	2.18	1.44	|	1.96	1.94	—	**0.006**	✓		-0.22	SD	SD
Lower molar buccal cusp (LM6BC), R	2.18	1.57	|	1.67	1.28	—	**< 0.001**	✓		-0.51	OD	SD
Lower molar buccal root (LM6BR), L	3.08	2.49	|	2.85	2.14	|	0.76			-0.23	OD	OD
Lower molar buccal root (LM6BR), R	2.82	1.96	|	2.39	1.75	|	0.069			-0.43	SD	SD
Average	1.84	1.60		1.81	1.61							

MRE, mean radial error; SD, standard deviation; M1, manual first trial; M2, manual second trial; AI, artificial intelligence, Mann-Whitney U test, Significance on $p < 0.05$ scale; O, Circular pattern; —, Primary along horizontal axis; |, Primary along vertical axis

Diff, difference; SD, expresses that the "Diff. M1-M2" and "Diff. M1-AI" are in the same direction on the axes; OD, expresses that the differences "Diff. M1-M2" and "Diff. M1-AI" are in the opposite direction on the axes

that the results apply to a broader clinical population rather than to an idealized subset [26].

This work introduced a new method for accurately and automatically detecting PA cephalometric landmarks using a deep learning system. The suggested model's accuracy and reliability were compared to an human examiner.

Our study's findings indicate that the accuracy and reliability of the developed AI model are at the similar level as that of a human expert. While AI identification was superior in four skeletal points, manual identification was superior in one skeletal point and seven dental points. Differences (M1-M2)—(M1-AI) varied between 0.18 mm and 0.66 mm. Although statistically significant differences were demonstrated, they were not clinically important. These results suggest that using AI technology significantly enhances the efficiency and precision of cephalometric analysis by automatically identifying landmarks, lowering the time and effort needed.

It is important to note that the deviation of distance errors along a certain axis holds greater significance for some points. Therefore, the distribution of errors in the horizontal and vertical planes has been addressed separately. Most of the points in both M2 and AI varied in the y-axis direction according to the M1 in the present study.

Both vertical and horizontal deviations of points on PAs are important for assessing facial skeletal asymmetry, especially in diagnosis. The amount of chin deviation was associated with the absolute differences of the left and right antegonial point to the y-axis and the zygomaticofrontal suture to the x-axis in the study by Fong et al. [28].

Turning to the limitations of our work: The gold standard PA cephalometric landmark measurements were those of a single (not several) orthodontist(s), and single-center data may not be generalizable. One limitation of this study is that the average SDRs for AI were calculated within 2.0, 2.5, 3.0, and 4.0 mm, respectively; however, they were not assessed within 1.0 mm. The localizations of the points were evaluated, but the study was not focused on clinical diagnosis.

Conclusions

1. The cascade CNN algorithm auto-identified 47 PA cephalometric landmarks. The point-to-point error (mean 1.84 mm) was clinically acceptable. The landmark average SDRs were 68.2% within 2.0 mm and 92.5% within 4.0 mm.
2. The algorithm automatically identified PA cephalometric landmarks; this is an effective alternative to manual identification.

3. Although all MREs were below 4 mm, the thresholds of only 33 landmarks were below the acceptable upper bound of 2 mm.
4. The MREs of the (L, R) condyles, the crista galli, the (L) jugal processes, the (L, R) upper cuspids, and all (L, R) molars exceeded 2 mm, and differed markedly between the AI- predicted and manual methods.

Abbreviations

FARNet	Feature aggregation and refinement network
PA	Posteroanterior
AI	Artificial intelligence
SDR	Successful detection rates
CNN	Convolutional neural network
MSFA	Multiscale feature aggregation
FR	Feature refinement
MRE	Mean radial error
SD	Standard deviations
ICC	Intraclass correlation coefficients

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12880-024-01478-z.

Supplementary Material 1

Supplementary Material 2

Acknowledgements
I thank Mehmet Ugurlu and Elif Bilgir for access to the radiological archive, and assistance in development of the study concept and the AI model.

Author contributions
MG: Conceptualization, methodology, annotation of ground truth and test, data analysis, writing original draft preparation, reviewing and editing. İŞB: Conceptualization, methodology, acquisition of the data, writing original draft preparation, reviewing and editing. ÖÇ: Conceptualization, methodology, data analysis, writing original draft preparation, reviewing and editing.

Funding
The authors did not receive any funding from any organizations.

Data availability
The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethical statement and consent to participate
The study design was approved by the Ethics Committee of the Osmangazi University, Faculty of Medicine (decision date and decision number: 04.10.2022/22). All participants gave informed consent. Additional informed consent was obtained from all individual participants included in the study. We confirmed that the study was carried out according to the relevant guidelines and regulations of the Helsinki Declaration.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 8 August 2024 / Accepted: 23 October 2024
Published online: 30 October 2024

References

1. Na ER, Aljawad H, Lee KM, Hwang HS. A comparative study of the reproducibility of landmark identification on posteroanterior and anteroposterior cephalograms generated from cone-beam computed tomography scans. Korean J Orthod. 2019;49(1):41–8.
2. Ulkur F, Ozdemir F, Germec-Cakan D, Kaspar EC. Landmark errors on posteroanterior cephalograms. Am J Orthod Dentofac Orthop. 2016;150(2):324–31.
3. Shaw K, McIntyre G, Mossey P, Menhinick A, Thomson D. Validation of conventional 2D lateral cephalometry using 3D cone beam CT. J Orthod. 2013;40(1):22–8.
4. Bajaj K, Rathee P, Jain P, Panwar VR. Comparison of the Reliability of Anatomic Landmarks based on PA Cephalometric Radiographs and 3D CT Scans in Patients with Facial Asymmetry. Int J Clin Pediatr Dent. 2011;4(3):213–23.
5. Yousefi F, Rafiei E, Mahdian M, Mollabashi V, Saboonchi SS, Hosseini SM. Comparison Efficiency of Posteroanterior Cephalometry and Cone-beam Computed Tomography in Detecting Craniofacial Asymmetry: A Systematic Review. Contemp Clin Dent. 2019;10(2):358–71.
6. Park JH, Lee G-H, Moon D-N, Yun K-D, Kim J-C, Lee KC. Creation of Digital virtual patient by integrating CBCT, Intraoral scan, 3D facial scan: an Approach to Methodology for Integration Accuracy. J Craniofac Surg. 2022;33(4):e396–8.
7. Apaydın BK, Icoz D, Uzun E, Orhan K. Investigation of the relationship between the mandibular third molar teeth and the inferior alveolar nerve using posteroanterior radiographs: a pilot study. BMC Oral Health. 2024;24(1):371.
8. Kim MJ, Liu Y, Oh SH, Ahn HW, Kim SH, Nelson G. Evaluation of a multi-stage convolutional neural network-based fully automated landmark identification system using cone-beam computed tomographysynthesized posteroanterior cephalometric images. Korean J Orthod. 2021;51(2):77–85.
9. Shan T, Tay FR, Gu L. Application of Artificial Intelligence in Dentistry. J Dent Res. 2021;100(3):232–44.
10. Bichu YM, Hansa I, Bichu AY, Premjani P, Flores-Mir C, Vaid NR. Applications of artificial intelligence and machine learning in orthodontics: a scoping review. Prog Orthod. 2021;22(1):18.
11. Subramanian AK, Chen Y, Almalki A, Sivamurthy G, Kafle D. Cephalometric Analysis in Orthodontics Using Artificial Intelligence-A Comprehensive Review. Biomed Res Int 2022, 2022:1880113.
12. Ao Y, Wu H. Feature Aggregation and Refinement Network for 2D Anatomical Landmark Detection. J Digit Imaging. 2023;36(2):547–61.
13. Koo T, Li M. Cracking the code: providing insight into the fundamentals of research and evidence-based practice a guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med. 2016;15(2):155–63.
14. Sicurezza E, Greco M, Giordano D, Maiorana F, Leonardi R. Accuracy of landmark identification on postero-anterior cephalograms. Prog Orthod. 2012;13(2):132–40.
15. Lindner C, Wang CW, Huang CT, Li CH, Chang SW, Cootes TF. Fully Automatic System for Accurate Localisation and Analysis of Cephalometric Landmarks in Lateral Cephalograms. Sci Rep. 2016;6:33581.
16. de Mesquita QTB, Vieira G, Vidigal WA, Travencolo MTC, Beaini BAN, Spin-Neto TL, Paranhos R. de Brito Junior RB: Artificial Intelligence for Detecting Cephalometric Landmarks: A Systematic Review and Meta-analysis. J Digit Imaging. 2023;36(3):1158–79.
17. Schwendicke F, Chaurasia A, Arsiwala L, Lee JH, Elhennawy K, Jost-Brinkmann PG, Demarco F, Krois J. Deep learning for cephalometric landmark detection: systematic review and meta-analysis. Clin Oral Investig. 2021;25(7):4299–309.
18. Wang CW, Huang CT, Hsieh MC, Li CH, Chang SW, Li WC, Vandaele R, Maree R, Jodogne S, Geurts P, et al. Evaluation and Comparison of Anatomical Landmark Detection Methods for Cephalometric X-Ray Images: A Grand Challenge. IEEE Trans Med Imaging. 2015;34(9):1890–900.
19. Park JH, Hwang HW, Moon JH, Yu Y, Kim H, Her SB, Srinivasan G, Aljanabi MNA, Donatelli RE, Lee SJ. Automated identification of cephalometric landmarks: Part 1-Comparisons between the latest deep-learning methods YOLOV3 and SSD. Angle Orthod. 2019;89(6):903–9.
20. Hwang HW, Moon JH, Kim MG, Donatelli RE, Lee SJ. Evaluation of automated cephalometric analysis based on the latest deep learning method. Angle Orthod. 2021;91(3):329–35.
21. Ugurlu M. Performance of a Convolutional Neural Network- Based Artificial Intelligence Algorithm for Automatic Cephalometric Landmark Detection. Turk J Orthod. 2022;35(2):94–100.
22. Gil SM, Kim I, Cho JH, Hong M, Kim M, Kim SJ, Kim YJ, Kim YH, Lim SH, Sung SJ, et al. Accuracy of auto-identification of the posteroanterior cephalometric

landmarks using cascade convolution neural network algorithm and cephalometric images of different quality from nationwide multiple centers. Am J Orthod Dentofac Orthop. 2022;161(4):e361–71.
23. Lee H, Cho JM, Ryu S, Ryu S, Chang E, Jung YS, Kim JY. Automatic identification of posteroanterior cephalometric landmarks using a novel deep learning algorithm: a comparative study with human experts. Sci Rep. 2023;13(1):15506.
24. Major PW, Johnson DE, Hesse KL, Glover KE. Landmark identification error in posterior anterior cephalometrics. Angle Orthod. 1994;64(6):447–54.
25. Leonardi R, Annunziata A, Caltabiano M. Landmark identification error in posteroanterior cephalometric radiography. A systematic review. Angle Orthod. 2008;78(4):761–5.
26. Popova T, Stocker T, Khazaei Y, Malenova Y, Wichelhaus A, Sabbagh H. Influence of growth structures and fixed appliances on automated cephalometric landmark recognition with a customized convolutional neural network. BMC Oral Health. 2023;23(1):274.
27. Ghosh S, Shet R, Amon P, Hutter A, Kaup A. Robustness of Deep Convolutional Neural Networks for Image Degradations. In: 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP): 15–20 April 2018 2018; 2018: 2916–2920.
28. Fong JH, Wu HT, Huang MC, Chou YW, Chi LY, Fong Y, Kao SY. Analysis of facial skeletal characteristics in patients with chin deviation. J Chin Med Association: JCMA. 2010;73(1):29–34.

Publisher's note