Jiang *et al. Head & Face Medicine* (2025) 21:20
https://doi.org/10.1186/s13005-025-00498-6

Head & Face Medicine

RESEARCH Open Access

Deep learning based quantitative cervical vertebral maturation analysis

Fulin Jiang[1], Abbas Ahmed Abdulqader[2], Yan Yan[2], Fangyuan Cheng[3], Tao Xiang[1], Jinghong Yu[1], Juan Li[2,4*], Yong Qiu[1*] and Xin Chen[1*]

Abstract

Objectives This study aimed to enhance clinical diagnostics for quantitative cervical vertebral maturation (QCVM) staging with precise landmark localization. Existing methods are often subjective and time-consuming, while deep learning alternatives withstand the complex anatomical variations. Therefore, we designed an advanced two-stage convolutional neural network customized for improved accuracy in cervical vertebrae analysis.

Methods This study analyzed 2100 cephalometric images. The data distribution to an 8:1:1 for training, validation, and testing. The CVnet system was designed as a two-step method with a comprehensive evaluation of various regions of interest (ROI) sizes to locate 19 cervical vertebral landmarks and classify precision maturation stages. The accuracy of landmark localization was assessed by success detection rate and *student t-test*. The QCVM diagnostic accuracy test was conducted to evaluate the assistant performances of our system for six junior orthodontists.

Results Upon precise calibration with optimal ROI size, the landmark localization registered an average error of 0.66 ± 0.46 mm and a success detection rate of 98.10% within 2 mm. Additionally, the identification accuracy of QCVM stages was 69.52%, resulting in an enhancement of 10.95% in the staging accuracy of junior orthodontists in the diagnostic test.

Conclusions This study presented a two-stage neural network that successfully automated the identification of cervical vertebral landmarks and the staging of QCVM. By streamlining the workflow and enhancing the accuracy of skeletal maturation estimation, this method offered valuable clinical support, particularly for practitioners with limited experience or access to advanced diagnostic resources, facilitating more consistent and reliable treatment planning.

Keywords Quantitative cervical vertebral maturation (QCVM), Lateral cephalogram, Automated landmark location, Artificial intelligence, Orthodontics

*Correspondence:
Juan Li
lijuan@scu.edu.cn
Yong Qiu
qiuy@cqu.edu.cn
Xin Chen
chenxin@cqu.edu.cn
[1]College of Computer Science, Chongqing University, Chongqing University Three Gorges Hospital, Chongqing 400044, China
[2]State Key Laboratory of Oral Diseases, West China School of Stomatology, West China Hospital of Stomatology, Sichuan University, Chengdu 610041, China
[3]Chengdu Boltzmann Intelligence Technology Co., Ltd, Chengdu 610095, China
[4]Department of Orthodontics, West China School of Stomatology, Sichuan University, 14#, 3rd Section, Renmin Nan Road, Chengdu, Sichuan, CN 610041, China

Introduction

Dentofacial deformities associated with abnormal maxillo-mandibular relationships, such as mandibular hypoplasia and maxillary hypoplasia, will significantly affect facial aesthetics, oral function, and mental health [1]. During the pubertal growth spurt, correcting these dentofacial deformities becomes crucial. The success of this correction depends heavily on targeted orthodontic treatments. The optimal timing for such treatments is determined by accurate skeletal maturation assessments. Here, lateral cephalograms play a pivotal role. They are the key tools in this process, specifically focusing on the maturation stages of the 2nd – 4th cervical vertebrae (C2 - C4) [2–4]. Clinical skeletal maturation determination is considered one of the daily challenges faced by orthodontists. The method's variability could impact treatment outcomes, especially in borderline cases where precise staging is critical. The evaluation of cervical vertebral maturation (CVM) was first introduced by Lamparski et al. [5]. Later, a recent meta-analysis suggested that lateral cephalometric images could be a reliable alternative to hand-wrist radiographs for assessing growth maturation, as it eliminates the need for additional X-ray exposure [6]. Previous method for CVM evaluation, characterized by its visual qualitative nature, predominantly relied on the observation of specific morphological features of the vertebrae, such as the concavity of the lower border and the shapes of distinct vertebral bodies as trapezoidal, rectangular horizontal, square, and rectangular vertical [7–10]. However, the reproducibility of this method has been the subject of debate, evidenced by divergent findings across various studies [11–15]. To address these concerns, recent propositions of Quantitative Cervical Vertebral Maturation (QCVM) include analytical techniques that employ points, angles, and equations to enhance the precision of stage classification [16, 17]. Although the QCVM method reliably classifies skeletal maturation stages, the manual localization of landmarks and calculation of measurements in QCVM are time-consuming, posing an obstacle in its adoption for routine clinical practice. Hence it is imperative to develop an advanced system for automatic QCVM evaluation in lateral cephalograms, which could complete the accurate identification and localization of landmarks [18, 19].

In recent years, numerous studies have employed deep learning techniques to explore vertebral growth analysis [19–23]. Nonetheless, these investigations predominantly focused on visual qualitative analysis, which was susceptible to variances caused by factors such as individual vertebral size, morphology, alignment, and orientation. Alongside interference from nearby structures and inconsistent image quality from different cephalography machines, the factors above significantly impacted the precision and reliability of automated analyses, as illustrated in Fig. 1.

To address these challenges, we designed a novel landmark-location-based method for automatic QCVM staging in cephalometric images. The core of our method was a two-stage neural network that first detected 19 landmarks of C2-C4 and then calculated 7 cervical vertebrae measurements to predict the QCVM stage accurately. In the first stage of landmark location, different from the traditional method that simultaneously regressed all landmarks [24], we first identified each cervical vertebra and then regressed landmarks on each individual vertebra separately. We observed that regardless of the cervical vertebrae shape, position, and orientation, the landmark position relationship among vertebral borders remained stable within a single cervical vertebra. Therefore, this relationship served as a reliable indicator to enhance the precision of landmark localization, especially in cases where cervical vertebrae exhibited significant variation in arrangement or overlapping in anatomical structures. To reliably locate C2, C3, and C4, we predicted the centroid of each vertebra and generated three regions of interest (ROIs) with proportionate sizes that were calculated based on distances between these centroids rather than using fixed pixel sizes. Within each ROI, we designed a multi-scale prediction mechanism for landmarks detection. In the second stage of QCVM staging, we initially calculated 7 measurements of the three cervical vertebral bodies based on the detected landmarks. Subsequently, all these measurements were fed into the staging prediction module to generate the QCVM stage. Furthermore, to enhance the classification accuracy of the QCVM stage, we conducted a multicenter retrospective study to collect a dataset of 2,100 lateral cephalometric images with ethical approval. These images were evenly distributed across CS1-CS6 and labeled by experts with over 10 years of orthodontic experience, specifically for network training. The newly proposed components effectively ensured the precision of QCVM stage prediction and significantly enhanced the practicality of our algorithm in real-world clinical applications. The investigation process and overview are illustrated in Fig. 2.

Our main contributions are summarized as follows:

- We proposed a novel CVnet that formulated the QCVM as two sub-problems: accurate cervical vertebral landmark location and QCVM staging prediction based on landmarks.
- We designed an inter-related strategy to locate landmarks efficiently. Besides, a centroids-regarding size calculation scheme was introduced to improve ROI accuracy.
- Extensive evaluations were conducted on a dataset collected from multicenter dental institutions. Our

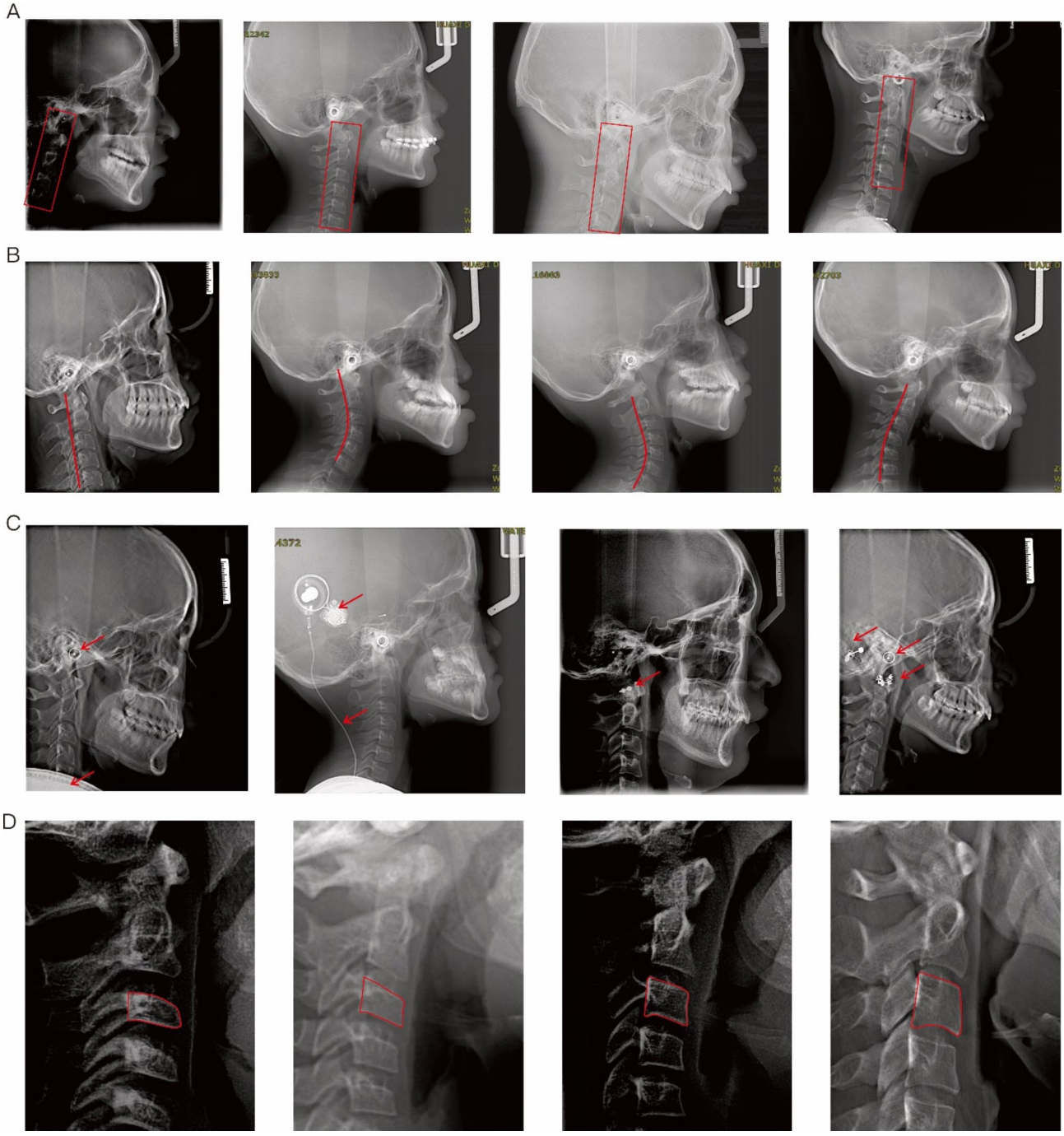

Fig. 1 Examples of various conditions that could compromise the accuracy of automatic landmark localization and CVM staging on cervical vertebrae. (**A**) Divergence on the location of cervical vertebrae between distinct images. (**B**) Various directions, alignments and curvatures of cervical vertebrae. (**C**) Interference of underlying and surrounding structures. (**D**) Morphological variation of cervical vertebrae

methods assisted junior orthodontists in achieving significantly superior results in accuracy and efficiency.

The rest of the paper is organized as follows. In the **Methodology** section, we describe the data collection, data processing, and the proposed methodology in detail. In the **Results** section, we present the quantitative and qualitative results, along with the auxiliary diagnostic effectiveness of our method. In the **Discussion** section, we discuss the effectiveness of our network and the limitations of our approach. The last section provides the **conclusion** of our study.

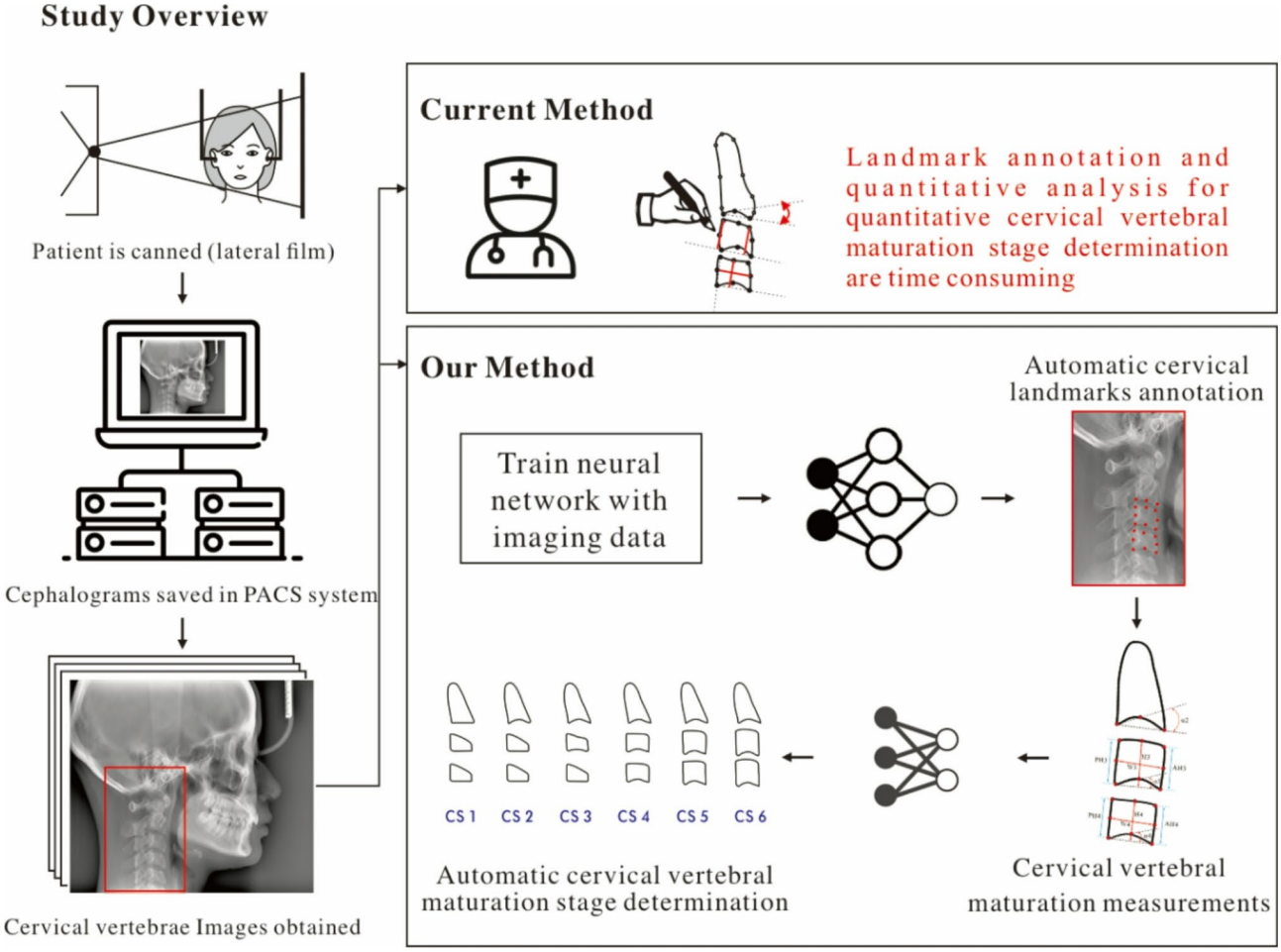

Fig. 2 Study overview explained current method for manual quantitative cervical vertebrae analysis and the proposed method for automatically quantitative cervical measurement and QCVM staging

Methodology

Ethical approval

From June 2019 to April 2021, 2100 lateral cephalometric radiographs were gathered from six hospitals and clinics. The study samples were between 6 and 19 years old, with a mean age of 12 years. This age range was chosen to ensure the inclusion of all CVM stages based on findings from previous research [15]. It was conducted in strict adherence to the World Medical Association's Helsinki Declaration, ensuring ethical standards in biomedical research involving human subjects. This study received approval from the Institutional Review Boards of Sichuan University West China Hospital of Stomatology No. WCHSIRB-D-2019-120.

Data collection and preprocessing

Data collection

This collection featured a wide variety of cases, encompassing diverse image qualities and a range of shapes and locations of vertebrae, providing a comprehensive dataset for analysis. Subjects collected for this study should exclude any conditions that might disrupt bone growth (such as systemic illnesses or delays in growth and development), and those showed congenital or acquired anomalies in the head and neck regions. It was mandatory for all lateral cephalometric radiographs to distinctly display the second (C2), third (C3), and fourth (C4) cervical vertebrae. All radiographs, originally captured in DICOM format, were converted into PNG format. The resolution range of images varied from 568 to 2,144 pixels in width and 570 to 2,600 pixels in height. Examples of original cephalometric images are shown in Fig. 1.

CVM staging

Initially, the original data had unknown CVM staging conditions and were unevenly distributed. To address this, the research team, along with 2 orthodontists (LJ and QY) with over 15 years of orthodontic experience, manually classified samples from CS1 to CS6. Then data were randomly distributed the dataset into training, validation, and test sets at a ratio of 8:1:1 for each stage. This methodological approach ensured an adequate

Table 1 Data distribution of CS1-CS6

CVM stage	Train dataset	Test dataset	Percentage
CS1	345	40	18.33%
CS2	343	40	18.24%
CS3	289	30	15.19%
CS4	275	30	14.52%
CS5	290	30	15.24%
CS6	348	40	18.48%
Total	1890	210	100.00%

representation for each stage, offering a comprehensive perspective on the scope of the study, as shown in Table 1.

Landmark selection and annotation

The selection of landmarks was primarily based on the recognized points required for QCVM assessment, as outlined by Chen et al. [17], and in accordance with the CVM Method User's Guide 2018 [25]. Specifically, the inferior borders of C2, C3, and C4, along with the morphological characteristics of C3 and C4, collectively formed the basis for CVM staging. Consequently, a total of 19 landmarks were selected, including 10 corner points of C2-C4, 3 superior-most points of the inferior borders of the bodies of C2-C4, and 6 midpoints of the posterior, anterior, and superior borders of C3 and C4. These landmarks are illustrated in Fig. 3B and Supplementary Table 1.

For landmarks annotation, two orthodontists with over five years of experience manually labeled landmarks to obtain the ground truth. Prior to the annotation process, orthodontists underwent training and evaluation conducted by the senior orthodontic professor Juan Li with over 15 years of experience. To assess the inter-rater and intra-rater agreement, 100 images were randomly selected for testing. The results indicated good agreement, with an inter-rater Cohen's kappa of 0.84 and an intra-rater Cohen's kappa of 0.87.

Data augmentation

To closely mirror the conditions encountered in actual clinical settings, our study made use of image

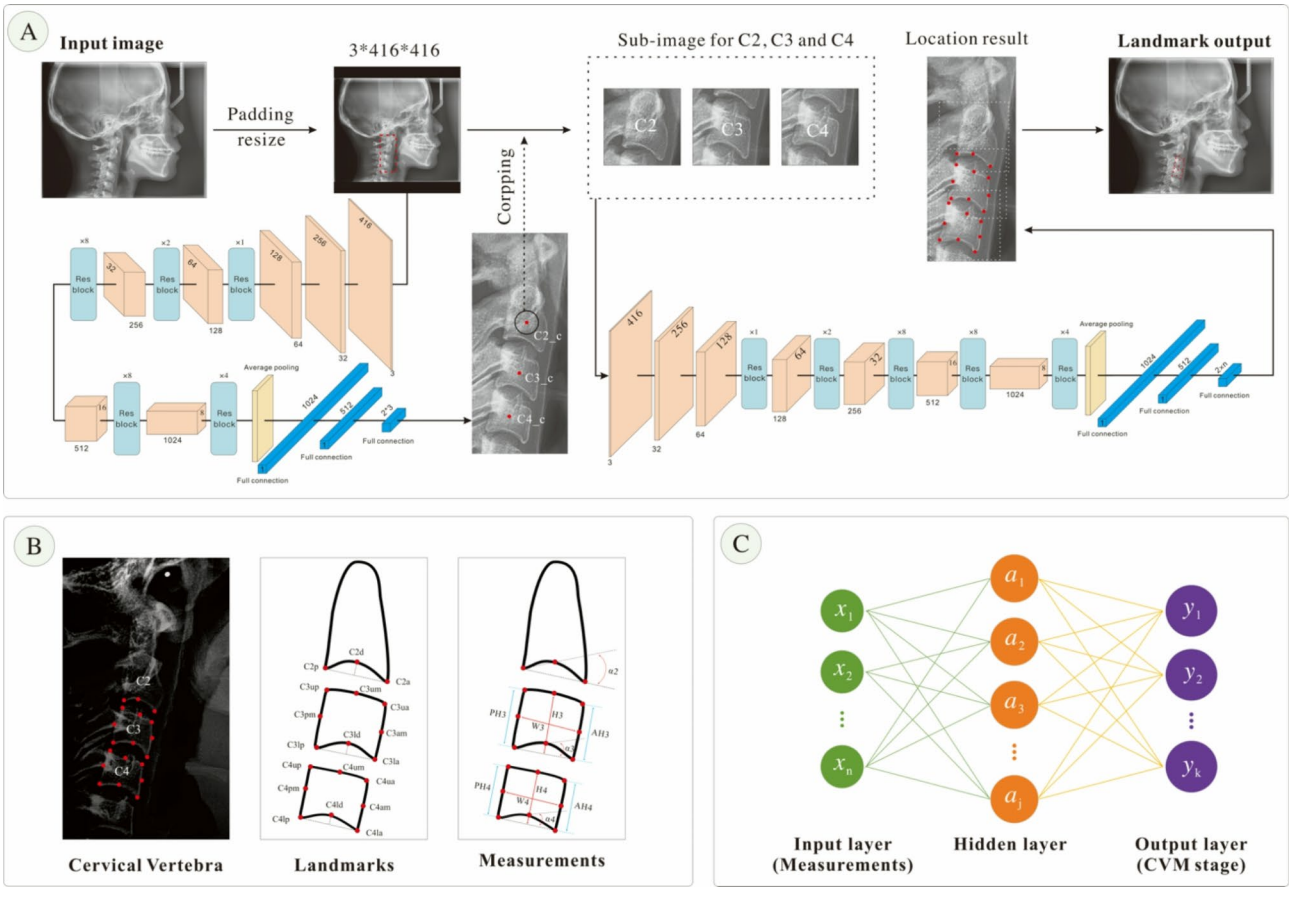

Fig. 3 The proposed framework (CVnet) for automatic quantitative cervical vertebral maturation analysis on input lateral cephalometric images. (**A**) The architecture of the landmark location network which firstly detects 3 ROIs of the second, third and fourth cervical vertebra (C2, C3, C4) and then locates landmarks on the sub-images of C2, C3 and C4. (**B**) Landmark definition and measurements for QCVM analysis. (**C**) The architecture of the QCVM determination network, the input is 7 measurements (α2, α3, α4, H3/W3, H4/W4, AH3/PH3, AH4/PH4), and the output is the classification of CS1 ~ 6

augmentation strategies to boost the performance of our model. We carefully selected a set of geometric and photometric transformation techniques relevant to clinical practices, including but not limited to rotation, solarizing, equalizing, inverting, adjustments in contrast on a random basis, and the combination of these methods. The example graphs of the data augmentation are shown in Fig. 4. These alterations were applied in a randomized manner, within set limits for the maximum degree of transformation, leading to an enlargement of our datasets by a factor of 10 to 15 times, thus reaching a substantial total of around 150,000 images.

Neural network architecture and training details

Landmark location

The landmark location network, built upon YOLOv3 [26], took cephalometric images as input and output the classification of landmarks along with their normalized coordinates. Leveraging its single-stage detection method, YOLOv3 can efficiently predict the category and location of targets in a single forward pass. Its built-in multi-scale feature fusion mechanism captures rich contextual information, making the model excel in locating interrelated landmarks, particularly suitable for anatomical structure landmark localization tasks. However, despite YOLOv3's outstanding performance in a landmark location, accurately locating cervical vertebra boundary points remains challenging due to the significant variations in cervical vertebra shapes and the blurred geometric signals near the vertebral boundaries (as shown in Fig. 1). Fortunately, the interrelationship of border landmarks within a single cervical vertebra remains relatively stable, regardless of the arrangement and morphological variations of the cervical vertebrae. Thus, we designed a cascaded landmark location network with two sub-networks, respectively, for ROI detection and landmark localization, as illustrated in Fig. 3A.

In the first step, an ROI detection network was introduced to locate three center points of C2, C3, and C4, which were C2_c, C3_c, and C4_c. Respectively, based on the center points, three sub-regions of C2, C3, and C4 were cropped. This was achieved by identifying a square for each vertebra based on the following parameters: The ROIs were three squares centered on C2_c,

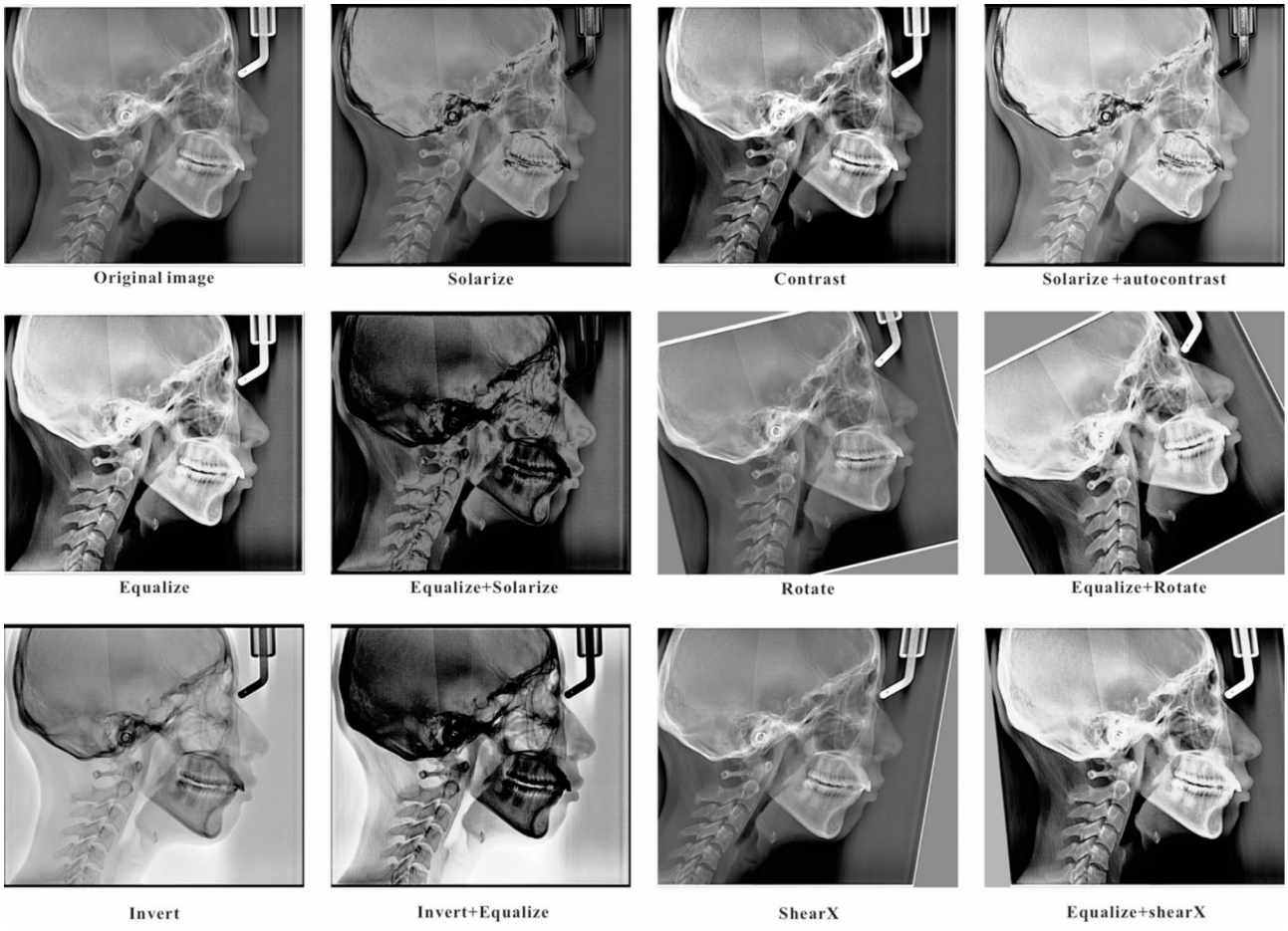

Fig. 4 The sample graphs of the data augmentation

C3_c, and C4_c, with its side length being 1.6 times the distance between C3_c and C4_c. The direction of the square was aligned with C2_c and C4_c connection. Through this method, the cervical vertebrae in different directions were aligned as parallel as possible to the axes of C2, C3, and C4. To locate C2_c, C3_c, and C4_c, the cross-entropy loss function (L_{CE}) was proposed to supervise the learning process. The loss function was defined as follows:

$$L_{CE} = -\sum_{i=1}^{N} [y_i \log \hat{y}_i + (1 - y_i) \log (1 - \hat{y}_i)]$$

Where N represented the category number, y represented the ground truth, and ŷ acted as the prediction output.

In the second step, a landmark location network was applied to accurately locate landmarks within.

each defined sub-region, ensuring a focused and detailed analysis of the vertebral landmarks in the lateral cephalometric images. The network designed for the second cervical vertebra (C2) was utilized to locate three points. Similarly, the network for C3 and C4 was employed to identify eight points, respectively. In the ROI detection network, all lateral cephalometric images were resized into the same input size of 416 × 416 pixels, and the cropped patch size of the localization network was also set as 416 × 416 pixels.

QCVM measurements calculation

Based on the 19 landmarks identified, seven quantitative measurements were derived and described in detail below, as illustrated in Fig. 3B and Supplementary Table 2. Specifically, the definition of the measurements was as follows: **α2**, Angle at C2 comparing the connection from C2d to C2p with C2p to C2a. **α3**, Angle at C3 comparing the connection from C3ld to C3lp with C3lp to C3la. **α4**, Angle at C4 comparing the connection from C4ld to C4lp with C4lp to C4la. **H3/W3**, Ratio of distance from C3um to midpoint of C3lp-C3la to distance between C3am and C3pm. **H4/W4**, Ratio of distance from C4um to the midpoint of C4lp-C4la to distance between C4am and C4pm. **AH3/PH3**, Ratio of vertical distance from C3ua to C3lp-C3la connection to the vertical distance from C3up to same connection. **AH4/PH4**: Ratio of vertical distance from C4ua to C4lp-C4la connection to the vertical distance from C4up to the same connection.

QCVM determination

Based on cervical vertebrae staging, 11 quantitative analysis indicators of the cervical vertebra were measured. Given the lack of direct standard values for QCVM staging (CS1-CS6) based on these indicators, we developed a fully connected network, as illustrated in the referenced Fig. 3C, to integrate 11 quantitative measurement indicators. This network facilitated the automatic prediction of QCVM staging from CS1 to CS6. The inputs to this fully connected network were the measurement indicators of each cervical vertebrae, and its output was the predicted staging results.

Implementation details

Python 3.7 was the primary programming language used for developing and running deep neural networks (DNNs). Implemented packages included PyTorch 1.8 (BSD-3-Clause License), an open-source machine learning library for building and training neural networks, OpenCV (Apache License 2.0) for computer vision and image processing tasks, NumPy (BSD-3-Clause License) for scientific computing, and Pillow (PIL Software License) as a friendly fork of the Python Imaging Library for image processing. The PyTorch framework was utilized to define the architecture of the DNNs, loss functions, optimizers, and training loops, enabling dynamic computation graphs for research and experimentation. Both DNNs were trained for 100 epochs using the Adam optimizer with an initial learning rate of 0.001, reduced by 10 every 40 epochs, on a single NVIDIA GeForce RTX 3090 GPU, taking 12 h for the ROI Detection Network, 14 h for the Localization Network, and 2 h for the Full Connection Layer. We have shared our code on GitHub for public access: https://github.com/Fulin-ortho/QCVM/tree/master.

Evaluation metrics

Automatic landmark location and QCVM staging accuracy

The proposed frameworks' performance was evaluated by landmark prediction accuracy and QCVM staging accuracy in the 210 test images.

Landmark prediction accuracy was assessed by calculating the distance between each predicted landmark location and its ground truth and the success detection rate (SDR) at the error ranges of 0.5, 1, 1.5, 2, 2.5, and 3 mm. For SDR_{φ}, if the radical error between a predicted landmark location and its ground truth were under φ mm, the detection was considered successful. The success detection rate for φ mm can be defined as:

$$SDR_{\varphi} = \frac{\#\{d_i : \|d_i - g_i\|_2 \leq \varphi\}}{\#\Omega} \times 100\%$$

Where d and g represented the detected location and ground truth location of a landmark, $\#(\cdot)$ represented the counting operation, $i \in \Omega$ and Ω served as the predictions on images.

For the assessment of QCVM staging performance, the predicted stages were compared with the gold standard of manual annotation, and true positives were calculated.

Our model and the manual inter-operator variability among orthodontists were compared regarding the

Table 2 Average landmark localization error and SDR by different models in the testing set (210 cephalograms)

Method	Average landmark localization error (mm)	Success detection rate (SDR) in different error ranges (%)					
		0.5 mm	1 mm	1.5 mm	2 mm	2.5 mm	3 mm
CVnet	0.66 ± 0.46	41.42	82.38	95.71	98.10	99.52	99.52
Manual	0.57 ± 0.43	52.38	87.62	96.67	98.57	99.52	99.52

Table 3 Quantitative localization results of 19 landmarks and comparison between CVnet localization error and inter-operator manual localization error. *$P < 0.05$ student t-test vs. control (Manual localization error); $n = 210$

Landmarks	Distance error (mm)			
	CVnet		Manual	
	Median	Mean ± SD	Median	Mean ± SD
C2p	0.59	0.64 ± 0.35*	0.43	0.48 ± 0.30*
C2a	0.77	0.81 ± 0.49*	0.45	0.53 ± 0.41*
C2d	0.64	0.68 ± 0.39*	0.45	0.55 ± 0.38*
C3lp	0.52	0.59 ± 1.02*	0.43	0.55 ± 1.04*
C3pm	0.53	0.62 ± 1.04*	0.49	0.61 ± 1.04*
C3up	1.07	1.20 ± 1.15	0.91	1.07 ± 1.14
C3um	0.69	0.81 ± 1.18*	0.63	0.76 ± 1.19*
C3ua	0.50	0.64 ± 1.17*	0.43	0.59 ± 1.16*
C3am	0.54	0.68 ± 1.16*	0.53	0.66 ± 1.17*
C3la	0.39	0.54 ± 1.18*	0.37	0.48 ± 1.19*
C3ld	0.43	0.55 ± 1.06*	0.39	0.51 ± 1.06*
C4lp	0.47	0.61 ± 1.04*	0.44	0.59 ± 1.03*
C4pm	0.59	0.70 ± 1.02*	0.55	0.67 ± 1.02*
C4up	1.24	1.30 ± 1.06	0.92	1.05 ± 1.09
C4um	0.79	0.87 ± 1.04	0.67	0.79 ± 1.12
C4ua	0.43	0.57 ± 1.07*	0.40	0.52 ± 1.09*
C4am	0.53	0.61 ± 1.08*	0.42	0.53 ± 1.12*
C4la	0.45	0.54 ± 1.09*	0.38	0.48 ± 1.09*
C4ld	0.44	0.55 ± 1.08*	0.37	0.49 ± 1.11*

landmark localization error. Statistical analysis was conducted using the one-way ANOVA test, with a significance level set at 0.05.

QCVM staging performance assessment with or without AI assistance

To comprehensively assess the proposed method, we incorporated subjective evaluations to gauge the practical benefits it offered orthodontists. Six junior orthodontists, each with 3–5 years of orthodontic experience, conducted visual assessments of QCVM staging for 210 lateral cephalograms, both with and without AI assistance. Initially, the orthodontists independently evaluated the QCVM staging, and we recorded the number of misdiagnoses and the time taken. After a two-week interval, the same orthodontists reassessed the lateral cephalograms, this time with the AI-predicted QCVM staging and confidence displayed on the images as a reference. The number of misdiagnoses and the time consumed were recorded again.

Results

The results of a detailed analysis of the landmark localization errors using frameworks on a test dataset of 210 images were analyzed. The average landmark localization error for CVnet was recorded at 0.66 ± 0.46 mm, closely approximating manual localization, which stood at 0.57 ± 0.43 mm, as summarized in Table 2. The SDRs identified as the best-performing model were further broken down at different error thresholds: 0.5 mm, 1 mm, 1.5 mm, 2 mm, 2.5 mm, and 3 mm. The corresponding SDRs were 41.42%, 82.38%, 95.71%, 98.10%, 99.52%, and 99.52%, respectively.

As shown in Table 3; Fig. 5, further details, including the median, mean, and standard deviation of the distance errors for each landmark, were provided. Besides, the localization error of the CVnet model was compared with the inter-operator manual tracing error using a student *t-test*. For the landmarks on cervical vertebrae C3 and C4, the model accuracy was comparable to that of manual inter-operator measurements in all landmarks except for C4 up, where some deviations were noted.

As shown in Table 4, the accuracy for each QCVM stage showed an average true positive rate of 69.52% across all six QCVM stages, with the exception of CS3 (33.33%), which displayed the lowest accuracy rate. Conversely, the highest accuracy rate was observed in CS6 (85.00%). The fully automatic model demonstrates remarkable efficiency by classifying the QCVM stages for 210 cephalograms in just 5 min.

Table 5 shows the accuracy of the QCVM staging assessment of six junior orthodontists on the 210 cephalograms with and without CVnet assistance. Without CVnet assistance, the average QCVM staging assessment accuracy was 56.9% and average assessment time was 34.6 min. However, with assistance of CVnet, the average accuracy reached 67.14% with an improvement of 10.24%, and average assessment time was decreased to 28.5 min.

Discussion

Quantitative Cervical Vertebral Maturation (QCVM) analysis is indispensable in orthodontic diagnosis and treatment planning but is often subjective, time-consuming, and labor-intensive, resulting in variability and inefficiency. This study develops a clinically applicable deep-learning model for fully automatic QCVM staging based on landmark localization. The model was trained and tested on large, real-world clinical datasets

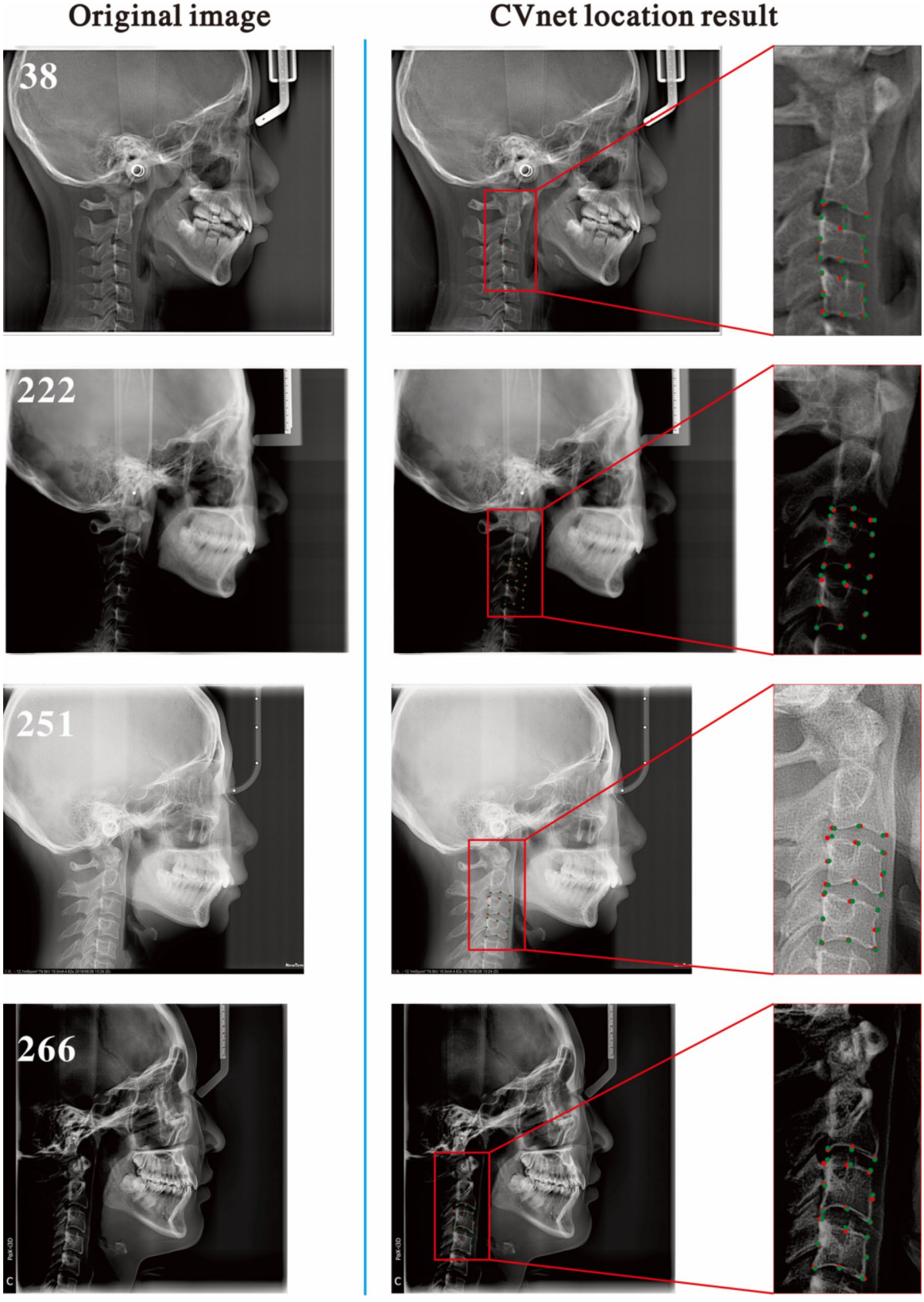

Fig. 5 Visualization of landmarks locations in lateral radiographs by CVnet. The first column represents the original image with the subject serial number and describes various cases of images from multiple centers. In the second and third columns, the superimposing of the ground truth and CVnet predicted results, while the Green points indicate the ground truth, and Red points show the predicted results. The accuracy of higher superimposition of the red and green landmarks indicates a better performance of CVnet

Table 4 CVnet classification accuracy in CS1-CS6

QCVM stage	Train dataset	Test dataset	True classification	False classification	Classification accuracy
CS1	345	40	33	7	82.50%
CS2	343	40	31	9	77.50%
CS3	289	30	10	20	33.33%
CS4	275	30	19	11	63.33%
CS5	290	30	19	11	63.33%
CS6	348	40	34	6	85.00%
Total	1890	210	146	64	69.52%

of cephalograms (i.e., 1890 images for training and 210 images for testing), demonstrating its potential for practical implementation.

One of the key attributes of our model is its fully automatic Quantitative Cervical Vertebral Maturation (QCVM) analysis across all six stages, offering robust performance. Radwan et al. [19] and Khazaei et al. [21] have consolidated the six CVM stages into three broader categories by merging every two stages. Although they achieved high classification performance, their method is limited in clinical application, particularly for precise mandibular growth prediction. Most AI methods, including those proposed by Kim et al. [12], Radwan et al. [19], and Rahimi et al. [18], have relied on segmentation and classification techniques rather than the QCVM method. While widely used in medical image analysis, segmentation, and classification approaches face challenges in CVM classification accuracy, often due to high computational demands and the necessity for precise annotations. Instead, our method adopts an anatomical landmark localization approach to develop quantitative measurements. To improve landmark accuracy, we focused on single vertebra ROI localization, leveraging the interrelationship of landmarks within individual vertebrae. Unlike previous studies [20, 21, 27] that employed fixed-pixel-resolution ROIs spanning the entire C2-C4 region, our approach dynamically calculated ROI sizes and orientations based on predicted centroid points. A recent study developed a PSC-CVM system based on visual qualitative analysis, which achieved an accuracy of 70.42% in CVM assessment by training on a large dataset [22]. Our method, with an accuracy comparable to PSC-CVM, demonstrates that a quantitative approach has the potential to enhance performance when applied to similarly large datasets, highlighting the robustness and scalability of our approach.

Another important contribution of this study lies in our comprehensive evaluations and assessments of clinical applicability that we have conducted on 2,100 images collected from multi-center clinics with balanced distribution across CS1-CS6, demonstrating great potential in dental clinical scenarios. Previous studies mostly included a restricted amount of single-center data, lacking rigorous validation of the model's robustness and generalization capabilities. For example, Akay et al. [23] and Liao et al. [28] included small datasets comprising 588 and 900 images, respectively, which exhibited an imbalanced distribution across various CVM stages. Consequently, the performance of these models on multi-center datasets remains unverified, as they have not been tested on diverse data obtained using different imaging protocols or scanning parameters. In addition, to validate the clinical applicability of CVnet, we not only assessed the accuracy of landmark localization in comparison to inter-operator manual localization error of orthodontists but also evaluated the auxiliary role of CVnet in QCVM staging, addressing a gap in previous research [23, 28, 29]. With the assistance of CVnet, six junior orthodontists achieved an average of 10.24% QCVM staging accuracy improvement and a 6.1-minute analysis time reduction in 210 images. These findings highlight the potential of CVnet to enhance clinical decision-making and streamline orthodontic workflows effectively.

The low accuracy of CS3 can be attributed to several factors. First and foremost, the morphological distinctions among CS2, CS3, and CS4 are insufficiently pronounced. Specifically, the primary divergence between CS2 and CS3 is manifested in the extent of the depression along the lower margin of C4. Additionally, the biological ages corresponding to these three cervical vertebra maturity stages are relatively close [15], which further contributes to the difficulty in accurately differentiating CS3.

Table 5 Performance of junior orthodontists (3–5 years experience) with or without CVnet assistance for QCVM stage classification in the testing set (210 cephalograms)

Methods		Ortho 1	Ortho 2	Ortho 3	Ortho 4	Orth 5	Ortho 6	Average
Without CVnet assistance	False	104	81	100	78	88	92	90.5
	True	106	129	110	132	122	118	119.5
	Time (min)	40	37	30	40	28	33	**34.67**
	Accuracy	50.48%	61.43%	52.38%	62.86%	58.10%	56.19%	**56.90%**
With CVnet assistance	False	81	57	74	59	71	72	69
	True	129	153	136	151	139	138	141
	Time (min)	30	27	30	32	25	27	**28.5**
	Accuracy	61.43%	72.86%	64.76%	71.90%	66.19%	65.71%	**67.14%**

These factors make this transition challenging and need further inclusion to this point in future works.

Although this work has achieved overall promising QCVM staging results, it still has flaws in crucial secondary factors like hand-wrist radiographs and comprehensive analysis of chronological age, which were essential for a holistic understanding of growth stages. In future work, as a post-processing step of our current method, we plan to collect paired hand-wrist radiographs and cephalograms with chronological age and develop a comprehensive system to better leverage the relationship between skeletal maturity and craniofacial development. Such an approach would provide a more accurate AI system for a comprehensive skeletal maturity analysis.

Conclusion

This study has developed a fully automated, precise, robust, and clinically viable AI model for CVM staging in cephalograms. By leveraging meticulously designed ROIs, it boosts the precision of cervical vertebral landmark localization to obtain reliable vertebral measurements as classification indicators. After quantitative evaluation and diagnostic assistance assessment on a large multicenter dataset, our model shows great promise for clinical applications, paving the way for more efficient CVM diagnosis and better patient care.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13005-025-00498-6.

Supplementary Material 1

Acknowledgements
Not applicable.

Author contributions
J. F. L. contributed to data acquisition conception, analysis, and critically revised the manuscript. A. A. A. contributed to the interpretation of data and critically revised the manuscript. Y. Y. contributed to data analysis and critically revised the manuscript. C. F. Y. contributed to data analysis, specifically in deep learning model training, and critically revised the manuscript. Y. J. H. and L. J. contributed to supervision and critically revised the manuscript. Q. Y. and C. X. contributed to supervision, provided guidance on the project's direction, ensured adherence to ethical research practices, and critically revised the manuscript. All authors gave final approval and agreed to be accountable for all aspects of the work, ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Funding
This research was funded by the Chongqing Wanzhou District PhD Direct-Acess Research Project [grant number wzstc20230409] and Chongqing Wanzhou District Science and Health Joint Medical Research Plan [grant number wzstc-kw2022016]. The funders had no role in the design of the study, the collection, analysis, and interpretation of data, the writing of the report, or the decision to submit the paper for publication.

Data availability
The data that support the findings of this study are available from the corresponding author, Dr. Qiu Yong, upon reasonable request. Due to privacy concerns, the data are not publicly available. For further inquiries, please contact Dr. Qiu at qiuy@cqu.edu.cn.

Declarations

Ethics approval and consent to participate
This study was conducted in strict adherence to the World Medical Association's Helsinki Declaration, ensuring ethical standards in biomedical research involving human subjects. This study received approval from the Institutional Review Boards of Sichuan university west China hospital of stomatology No. WCHSIRB-D-2019-120.). Informed consent was obtained from all individual participants included in the study.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 31 May 2024 / Accepted: 12 March 2025
Published online: 26 March 2025

References

1. de Ávila ÉD, de Molon RS, Loffredo LCM, Massucato EMS, Hochuli-Vieira E. Health-related quality of life and depression in patients with dentofacial deformity. Oral Maxillofacial Surg. 2013;17:187–91.
2. Mohammad-Rahimi H, Motamadian SR, Nadimi M, Hassanzadeh-Samani S, Minabi MAS, Mahmoudinia E, Lee VY, Rohban MH. Deep learning for the classification of cervical maturation degree and pubertal growth spurts: A pilot study. Korean J Orthod. 2022;52(2):112–22.
3. Navlani MP, Makhija. Evaluation of skeletal and dental maturity indicators and assessment of cervical vertebral maturation stages by height/width ratio of third cervical vertebra. J Pierre Fauchard Acad (India Section). 2013;27(3):73–80.
4. Xie L, Tang W, Izadikhah I, Chen X, Zhao Z, Zhao Y, Li H, Yan B. Intelligent quantitative assessment of skeletal maturation based on multi-stage model: a retrospective cone-beam CT study of cervical vertebrae. Oral Radiol. 2022;38(3):378–388.
5. Lampalski D. Skeletal age assessment utilizing cervical vertebrae. Master of Science Thesis, University of Pittsburgh, 1972.
6. Cericato G, Bittencourt M, Paranhos L. Validity of the assessment method of skeletal maturation by cervical vertebrae: a systematic review and meta-analysis. Dentomaxillofacial Radiol. 2015;44(4):20140270.
7. Baccetti T, Franchi L, McNamara JA Jr. An improved version of the cervical vertebral maturation (CVM) method for the assessment of mandibular growth. Angle Orthod. 2002;72(4):316–23.
8. Baccetti T, Franchi L, McNamara JA Jr. *The cervical vertebral maturation (CVM) method for the assessment of optimal treatment timing In dentofacial orthopedics*. In *Seminars In orthodontics*. Elsevier; 2005.
9. Alhamady AM, Ishaq RAR, Alhammadi MS, Almashraqi AA, Alhashimi N. Evaluation of an objective staging system for assessment of cervical vertebral maturation. BMC Oral Health. 2024;24(1):1–8.
10. Sella Tunis T, Masarwa M, Finkelstein T, Grinstein Koren O, Kats L, Manor Y, Shpack N, Reiter S. The reliability of a modified three-stage cervical vertebrae maturation method for estimating skeletal growth in males and females. BMC Oral Health. 2024;24(1):1255.
11. Nestman TS, Marshall SD, Qian F, Holton N, Franciscus RG, Southard TE. Cervical vertebrae maturation method morphologic criteria: poor reproducibility. Am J Orthod Dentofac Orthop. 2011;140(2):182–8.
12. Gabriel DB, Southard KA, Qian F, Marshall SD, Franciscus RG, Southard TE. Cervical vertebrae maturation method: poor reproducibility. Am J Orthod Dentofac Orthop, 2009. 136(4): p. 478. e1-478. e7.
13. Uysal T, Ramoglu SI, Basciftci FA, Sari Z. Chronologic age and skeletal maturation of the cervical vertebrae and hand-wrist: is there a relationship? Am J Orthod Dentofac Orthop. 2006;130(5):622–8.
14. Lai EH-H, Liu J-P, Chang JZ-C, Tsai S-J, Yao C-CJ, Chen M-H, Chen Y-J, Lin C-P. Radiographic assessment of skeletal maturation stages for orthodontic patients: hand-wrist bones or cervical vertebrae? J Formos Med Assoc. 2008;107(4):316–25.

15. Schoretsaniti L, Mitsea A, Karayianni K, Sifakakis I. Cervical vertebral maturation method: reproducibility and efficiency of chronological age Estimation. Appl Sci. 2021;11(7):3160.
16. Mito T, Sato K, Mitani H. Cervical vertebral bone age in girls. Am J Orthod Dentofac Orthop. 2002;122(4):380–5.
17. Chen L, Liu J, Xu T, Long X, Lin J. Quantitative skeletal evaluation based on cervical vertebral maturation: a longitudinal study of adolescents with normal occlusion. Int J Oral Maxillofac Surg. 2010;39(7):653–9.
18. Al Arif SMR, Knapp K, Slabaugh G. Fully automatic cervical vertebrae segmentation framework for X-ray images. Comput Methods Programs Biomed. 2018;157:95–111.
19. Kanchanapiboon P, Tunksook P, Tunksook P, Ritthipravat P, Boonpratham S, Satravaha Y, Chaweewannakorn C, Peanchitlertkajorn S. Classification of cervical vertebral maturation stages with machine learning models: leveraging datasets with high inter-and intra-observer agreement. Prog Orthodont. 2024;25(1):35.
20. Kim E-G, Oh I-S, So J-E, Kang J, Le VNT, Tak M-K, Lee D-W. Estimating cervical vertebral maturation with a lateral cephalogram using the convolutional neural network. J Clin Med. 2021;10(22):5400.
21. Mohammad-Rahimi H, Motamadian SR, Nadimi M, Hassanzadeh-Samani S, Minabi MA, Mahmoudinia E, Lee VY, Rohban MH. Deep learning for the classification of cervical maturation degree and pubertal growth spurts: A pilot study. Korean J Orthod. 2022;52(2):112–22.
22. Li H, Li H, Yuan L, Liu C, Xiao S, Liu Z, Zhou G, Dong T, Ouyang N, Liu L. The psc-CVM assessment system: A three-stage type system for CVM assessment based on deep learning. BMC Oral Health. 2023;23(1):557.
23. Akay G, Akcayol MA, Özdem K, Güngör K. Deep convolutional neural network-the evaluation of cervical vertebrae maturation. Oral Radiol. 2023;39(4):629–638.
24. Chen L-L, Xu T-M, Jiang J-H, Zhang X-Z, Lin J-X. Quantitative cervical vertebral maturation assessment in adolescents with normal occlusion: a mixed longitudinal study. Am J Orthod Dentofac Orthop. 2008;134(6):720. e1-720. e7.
25. McNamara Jr JAL, Franchi. The cervical vertebral maturation method: A user's guide. Angle Orthod. 2018;88(2):133–43.
26. Redmon J. Yolov3: An incremental improvement. arXiv preprint arXiv:1804.02767. 2018.
27. Radwan MT, Sin Ç, Akkaya N, Vahdettin L. Artificial intelligence-based algorithm for cervical vertebrae maturation stage assessment. Orthod Craniofac Res. 2023;26(3):349–55.
28. Liao N, Dai J, Tang Y, Zhong Q, Mo S. ICVM: an interpretable deep learning model for CVM assessment under label uncertainty. IEEE J Biomedical Health Inf. 2022;26(8):4325–34.
29. Khazaei M, Mollabashi V, Khotanlou H, Farhadian M. Automatic determination of pubertal growth spurts based on the cervical vertebral maturation staging using deep convolutional neural networks. J World Federation Orthodontists. 2023;12(2):56–63.

Publisher's note