Submitted: 31 October, 2023 Accepted: 19 December, 2023 Published: 03 July, 2024 DOI:10.22514/jocpd.2024.093

ORIGINAL RESEARCH

Journal of Clinical Pediatric Dentistry

Comparative analysis of deep-learning-based bone age estimation between whole lateral cephalometric and the cervical vertebral region in children

Suhae Kim[1,†], Jonghyun Shin[1,2,†], Eungyung Lee[1,2], Soyoung Park[1,2], Taesung Jeong[1,2], JaeJoon Hwang[2,3], Hyejun Seo[4,*]

[1]Department of Pediatric Dentistry, School of Dentistry, Pusan National University, 50612 Yangsan, Republic of Korea
[2]Dental and Life Science Institute & Dental Research Institute, School of Dentistry, Pusan National University, 50612 Yangsan, Republic of Korea
[3]Department of Oral and Maxillofacial Radiology, School of Dentistry, Pusan National University, 50612 Yangsan, Republic of Korea
[4]Department of Dentistry, Ulsan University Hospital, 44033 Ulsan, Republic of Korea

***Correspondence**
herrjoon@uuh.ulsan.kr
(Hyejun Seo)

† These authors contributed equally.

Abstract

Bone age determination in individuals is important for the diagnosis and treatment of growing children. This study aimed to develop a deep-learning model for bone age estimation using lateral cephalometric radiographs (LCRs) and regions of interest (ROIs) in growing children and evaluate its performance. This retrospective study included 1050 patients aged 4–18 years who underwent LCR and hand-wrist radiography on the same day at Pusan National University Dental Hospital and Ulsan University Hospital between January 2014 and June 2023. Two pretrained convolutional neural networks, InceptionResNet-v2 and NasNet-Large, were employed to develop a deep-learning model for bone age estimation. The LCRs and ROIs, which were designated as the cervical vertebrae areas, were labeled according to the patient's bone age. Bone age was collected from the same patient's hand-wrist radiograph. Deep-learning models trained with five-fold cross-validation were tested using internal and external validations. The LCR-trained model outperformed the ROI-trained models. In addition, visualization of each deep learning model using the gradient-weighted regression activation mapping technique revealed a difference in focus in bone age estimation. The findings of this comparative study are significant because they demonstrate the feasibility of bone age estimation *via* deep learning with craniofacial bones and dentition, in addition to the cervical vertebrae on the LCR of growing children.

Keywords

Bone age; Convolutional neural network; Deep learning; Lateral cephalometric radiograph; Skeletal maturity

1. Introduction

Understanding the growth and development of children is important for clinicians during diagnosis and treatment planning. Skeletal maturity is an important growth indicator and can be estimated from a variety of biological factors such as height, weight, and secondary sex characteristics [1, 2]. Individual growth and maturation are closely related to skeletal features, and radiographic imaging is a representative method for measuring skeletal maturation. Several radiographic images, such as hand-wrist radiography and lateral cephalometric radiographs (LCRs), can be used to evaluate skeletal maturation [3, 4]. Although evaluating bone age with hand-wrist radiography is a reliable method, it requires additional X-ray exposure and has the disadvantages of individual and sex differences [5]. In contrast, LCRs are widely used for orthodontic diagnosis and evaluation. Skeletal age can also be assessed by evaluating the morphological shape of the cervical vertebrae in LCRs [3].

Various studies have revealed the relationship between the shape of the cervical vertebrae and growth stage, and attempts have been made to estimate bone age in various ways using this structure [6, 7]. In these previous studies, bone age was calculated by measuring the length and height of the cervical vertebrae and applying a formula for bone age estimation. However, this bone age estimation method requires judgment based on measurements by the observer; therefore, it is highly dependent on the observer and time-consuming.

Owing to recent developments in artificial intelligence, diagnostic technologies for medical and dental radiography using deep learning have been developed [8–10]. Because deep learning with convolutional neural network (CNN) algorithms can perform tasks such as the detection, segmentation, and classification of specific structures, they have remarkable potential to assist clinicians [11]. Clinical applications stemming from these technological advances have yielded promising results, as they are less time-consuming and reduce human fatigue caused by repetitive tasks.

Previous studies have evaluated skeletal maturity using deep-learning techniques by designating the cervical vertebral area as a region of interest (ROI) on LCRs [12]. The ROI

is a subset of the entire image, and its extraction is essential to effectively reduce unnecessary space while preserving crucial information within the image [13]. However, in lateral cephalometric analysis, in addition to changes in the cervical vertebrae, the development of dentition and facial bones are growth predictors [14, 15]. Because the skeleton size and shape change with growth, there are other growth indicators such as dentition or craniofacial bones including sphenoid or temporal bones in LCRs that may assist in bone age estimation [16].

Therefore, this study aimed to develop and compare the performance of deep-learning models for bone age estimation in growing children using LCRs and ROIs. Additionally, to generalize the trained deep learning model, internal and external validations were conducted.

2. Materials and methods

2.1 Image dataset

This retrospective, multicenter study included a total of 1050 LCRs and hand-wrist radiographs. Both sets of images, which were undertaken on the same day, were acquired from each patient aged 4–18 years. The data were collected from Pusan National University Dental Hospital and Ulsan University Hospital between January 2014 and June 2023. All image data were stored in digital imaging and communications in medicine (DICOM) format. Radiographs that included metal artifacts or blurred areas, which were difficult for the examiner to interpret, were excluded.

2.2 Data preprocessing and annotation

The LCRs were manually cropped to isolate the ROIs corresponding to the cervical vertebral region. The Med-BoneAge version 1.2.1 software program (VUNO, Seoul, South Korea) was used to perform automated bone age analysis, which applies the Greulich-Pyle method [17]. Each patient's hand-wrist radiograph was processed using deep-learning-based automatic software for bone age determination, focusing on the shape and density of each bone. The software suggested three most-likely estimated bone age based on similarities among images. Given these three highly possible bone ages, the most probable bone age was calculated, and the suggested bone age was labeled on each image to facilitate the performance comparison between LCRs and ROIs (Fig. 1).

2.3 Transfer learning and implementation

CNNs, one of the most representative deep learning models, are commonly used to analyze images in the medical field [18]. To avoid instability in the results, two pretrained CNNs, namely, InceptionResNet-v2 and NasNet-Large, were employed in this study (Table 1).

InceptionResNet-v2 [19] is an advanced CNN inspired by the ResNet [20] backbone architecture that enhances efficiency by incorporating an inception module. NasNet-Large [21] is a CNN model composed of two types of cells, normal and reduction cells, and uses a reinforcement learning search method to determine the best architecture configurations.

For bone age estimation, these CNN architectures were modified for the regression task. The final fully connected layers for classification, the softmax layer, and the classification output layer were replaced with a fully connected layer for regression and a regression output layer. The newly replaced layers were used to predict the final regression results.

The deep-learning model was trained in MATLAB 2023b (MathWorks Inc., Natick, MA, USA) using a deep-learning and parallel-computing toolbox with an NVIDIA RTX 4090 (24 GB RAM) graphical processing unit. The model was trained for up to 500 epochs using the Adam optimizer. The mini-batch size was set at 16 and randomly shuffled at each epoch. Fig. 2 illustrates the bone age estimation process based on the deep-learning model. For internal validation, the dataset was randomly divided into two subsets: a training-validation set comprised of 950 images and a test set comprised of 50 images from the Pusan National University Dental Hospital. After model training, a test set of 50 images from the Ulsan University Hospital was evaluated for external validation. All training images were subjected to a five-fold cross-validation to ensure an accurate performance comparison. The training-validation set was divided into training and validation sets at a ratio of 4:1, which rendered all available datasets as training and validation sets. Various data augmentation techniques, such as rotation at multiple angles, horizontal and vertical scaling, and horizontal and vertical translation, were used to prevent overfitting of the deep-learning models on small datasets.

2.4 Performance metrics and visualization

All statistical analyses were performed using SPSS 26.0 (SPSS Corp., Armonk, NY, USA). The mean absolute error (MAE), root-mean-square error (RMSE), and R-squared metrics were used to evaluate the prediction error rates and model performance in the regression analysis.

Using the gradient-weighted regression activation mapping (Grad-RAM) method [12], we can observe where the trained deep-learning model focuses. The Grad-RAM method is a variant of the gradient-weighted class activation mapping (Grad-CAM) method, where the classification layer of the CNN architecture is transformed into a regression layer. In general, a higher score in the related heatmap corresponds to a warmer color (red), indicating a significant contribution to the regression outcome; conversely, cooler colors (blue) represent a lower score, indicating a lower contribution [22].

TABLE 1. Properties of the pretrained convolutional neural networks (CNNs) used in this study.

CNN	Depth	Size (MB)	Parameter (Millions)	Image input size
InceptionResNet-v2	164	209	55.9	299 × 299
NasNet-Large	1041	332	88.9	331 × 331

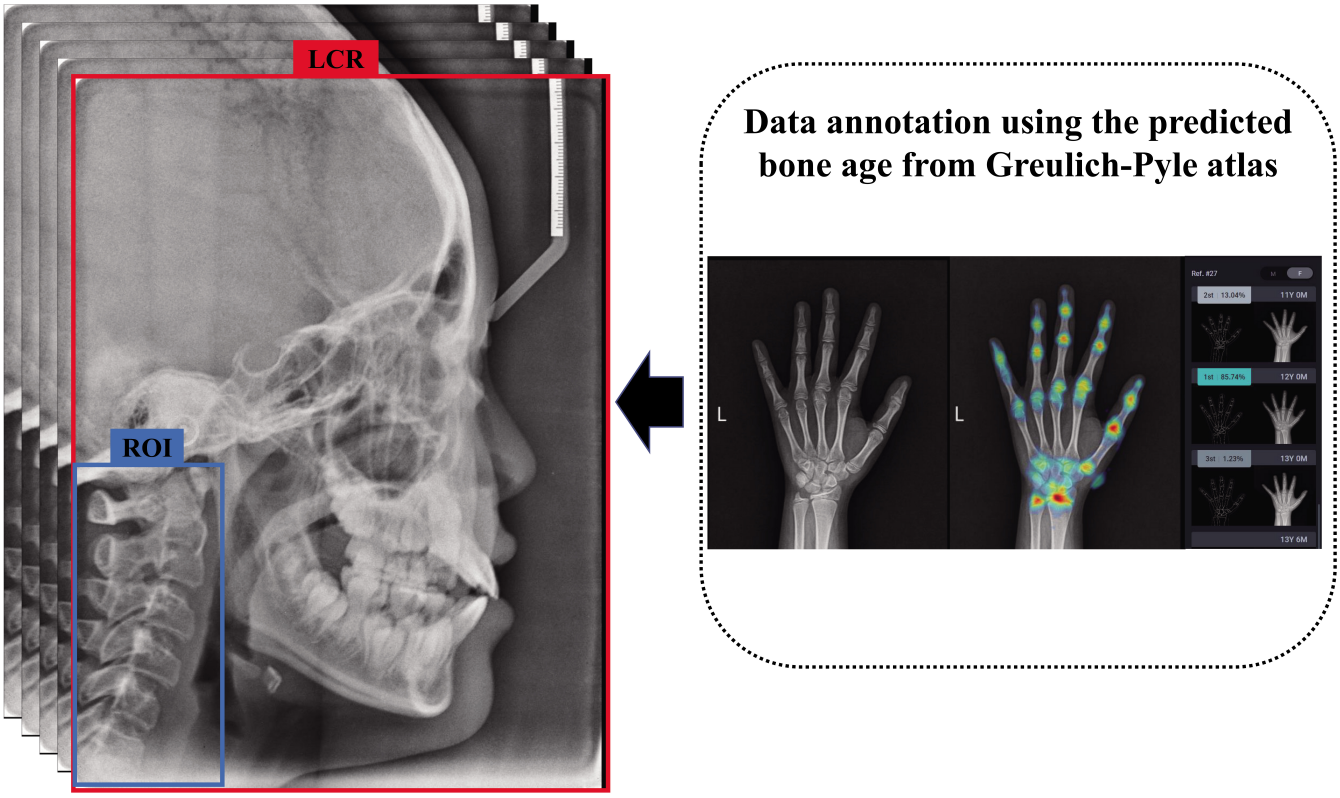

FIGURE 1. Data annotation. The LCRs and ROIs, which include the cervical vertebrae, were labeled with the bone age obtained from hand-wrist radiographs. LCR: lateral cephalometric radiograph; ROI: regions of interest.

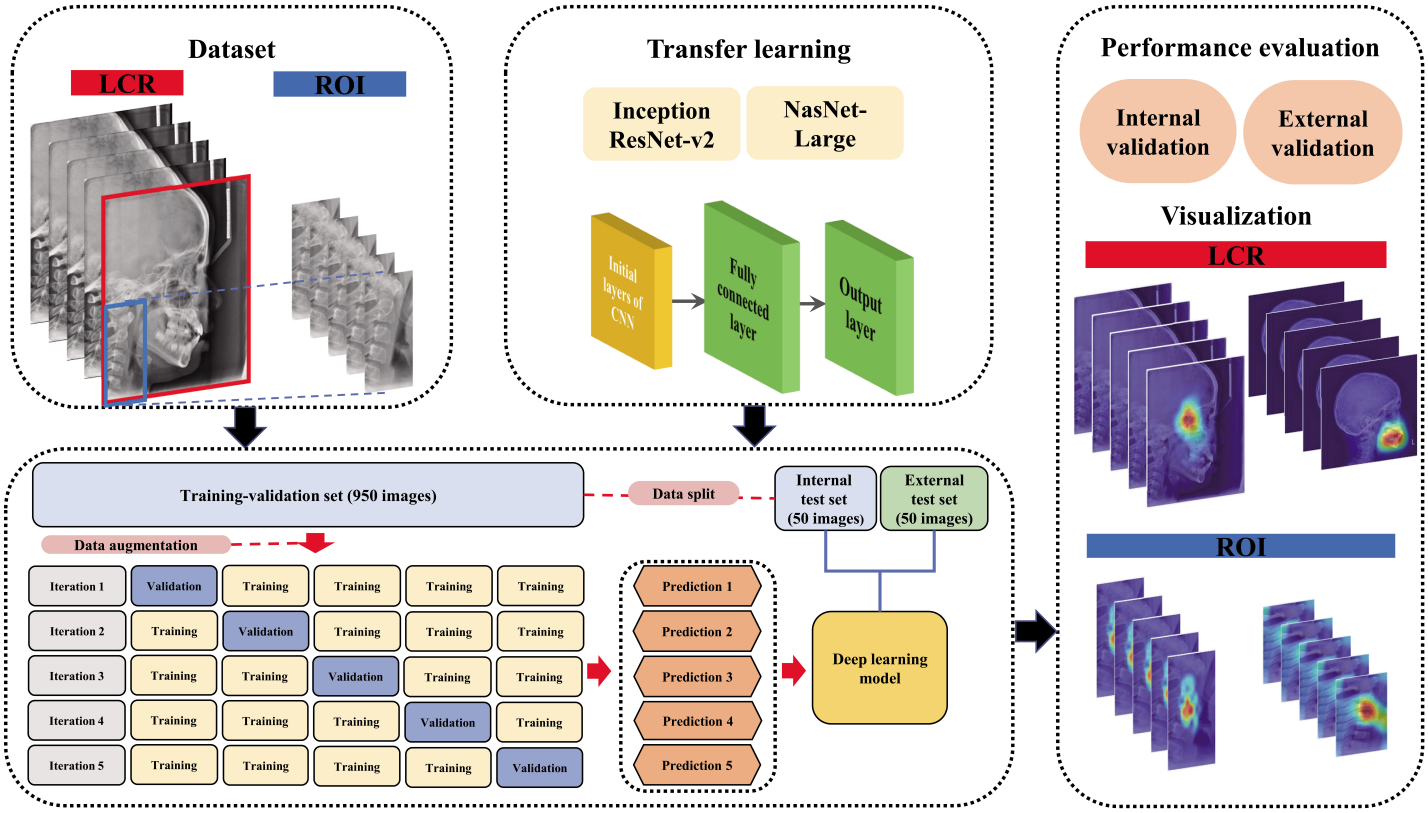

FIGURE 2. Workflow of the bone age estimation process. The dataset was divided into a training-validation set and an internal and external test set. Pretrained CNNs were used for transfer learning along with five-fold cross-validation to develop the deep-learning model which in turn was evaluated and visualized *via* internal and external validation to assess its performance. LCR: lateral cephalometric radiograph; ROI: regions of interest.

3. Results

3.1 Age estimation performance

An overview of the datasets used in this study is presented in Table 2. Among the 1050 patients, 530 were males and 520 were females. The mean chronological age was higher in males, while the mean bone age was higher in females.

Table 3 presents the internal validation results, where the training and testing datasets were obtained from the same hospital. The LCR-trained InceptionResNet-v2 and NasNet-Large models exhibited a better performance than the corresponding ROI-trained models, as evidenced by the regression performance metric values. The MAE and RMSE values were lower for the LCR-trained models compared to those of the ROI-trained models. Furthermore, the average R-squared values of the LCR-trained InceptionResNet-v2 and NasNet-Large models were both greater than 0.90.

Table 4 presents the external validation results, where the training and testing datasets were obtained from different hospitals. The MAE and RMSE values of the ROI-trained deep-learning models were slightly lower than those of the LCR-trained models. Whereas the R-squared values of the LCR-trained models were higher than those of the ROI-trained models.

3.2 Model regression visualization

For the LCRs, the InceptionResNet-v2 model mainly focused on the facial bones, including dentition, while the NasNet-Large model mainly focused on the cranial base and midfacial bones (Fig. 3). In contrast, for the ROIs, including the cervical vertebrae, the activation map of the InceptionResNet-v2 model exhibited a different pattern compared to that of the NasNet-Large model, which was somewhat sporadic (Fig. 4).

4. Discussion

In this study, transfer learning with pretrained CNNs was used for bone age estimation from LCRs and compared the performance of deep learning models using LCRs and ROIs. In the cervical vertebral maturation (CVM) stage assessment, previous studies have shown that the performance of deep-learning models is better when the ROI is specified compared to when the entire image is used [23]. This is because, on the LCRs, the rest of the area, except for the cervical vertebrae, is not needed when classifying the CVM stages.

The bone age estimated from hand-wrist radiographs was used as reference data for the deep-learning models for bone age estimation from LCRs. The deep-learning software used in this study was an automated evaluation program based on the Greulich-Pyle method [17]. The overall agreement rate between the estimated and reference bone ages exceeded 93%,

TABLE 2. Demographic data in this study.

	Male	Female	All
Numbers	530	520	1050
Chronological age ± SD	9.27 ± 1.95	9.29 ± 2.09	9.28 ± 2.02
Bone age ± SD	9.07 ± 2.60	9.39 ± 2.50	9.23 ± 2.55

SD: Standard Deviation.

TABLE 3. Regression performance of internal validation.

Model	Fold	MAE (yr)		RMSE (yr)		R-squared	
		LCR	ROI	LCR	ROI	LCR	ROI
InceptionResNet-v2							
	1	0.549	0.733	0.632	0.910	0.923	0.850
	2	0.611	0.823	0.700	1.068	0.915	0.802
	3	0.556	0.720	0.677	0.896	0.910	0.844
	4	0.629	0.702	0.731	0.879	0.929	0.857
	5	0.573	0.654	0.672	0.850	0.913	0.864
	Average	0.584	0.726	0.683	0.921	0.918	0.843
NasNet-Large							
	1	0.531	0.722	0.624	0.926	0.927	0.859
	2	0.569	0.698	0.666	0.893	0.930	0.849
	3	0.568	0.758	0.701	0.897	0.910	0.845
	4	0.630	0.854	0.777	1.085	0.895	0.787
	5	0.654	0.803	0.803	1.064	0.904	0.828
	Average	0.590	0.767	0.714	0.973	0.913	0.834

MAE: Mean Absolute Error; RMSE: Root Mean Square Error; LCR: Lateral cephalometric radiograph; ROI: Region of interest.

TABLE 4. Regression performance of the external validation.

Model	Fold	MAE (yr)		RMSE (yr)		R-squared	
		LCR	ROI	LCR	ROI	LCR	ROI
InceptionResNet-v2							
	1	0.846	0.691	1.028	0.880	0.883	0.859
	2	1.060	0.678	1.235	0.834	0.874	0.874
	3	0.765	0.825	0.934	1.004	0.846	0.838
	4	0.753	0.821	0.927	0.981	0.872	0.837
	5	0.737	0.756	0.937	0.940	0.844	0.840
	Average	0.832	0.754	1.012	0.928	0.864	0.850
NasNet-Large							
	1	0.863	0.806	1.019	0.997	0.842	0.821
	2	0.855	0.897	1.074	1.054	0.892	0.819
	3	0.785	0.661	0.951	0.843	0.869	0.868
	4	0.842	0.836	0.989	1.001	0.872	0.843
	5	0.825	0.797	0.957	0.979	0.870	0.845
	Average	0.834	0.799	0.998	0.975	0.869	0.839

MAE: Mean Absolute Error; RMSE: Root Mean Square Error; LCR: Lateral cephalometric radiograph; ROI: Region of interest.

ranking third overall [24].

Different validation strategies exist for evaluating deep learning models. In this study, the deep-learning model performance was evaluated using the split-sample method for the internal validation, while the external validation was performed with external testing data from the original prediction model. During the internal validation, the deep-learning model using LCRs exhibited a relatively good performance compared to the deep-learning model using ROIs. The Grad-RAM images showed differences in the focal areas emphasized by each CNN-based deep-learning model. In the ROIs, there were differences in the size of the activation maps generated from the cervical vertebrae and the number of foci. The NasNet-Large model exhibited relatively small activation maps, which were focused multiple times, whereas the InceptionResNet-v2 model featured relatively large activation maps, which were focused only once. However, a larger difference was observed in the LCRs. This difference can be attributed to the presence of numerous anatomical structures that change with age and appear on the LCR. The deep-learning model estimated the bone age not only by observing the cervical vertebrae but also assessing the degree of dentition and craniofacial bone development. When observing the LCRs, the InceptionResNet-v2 model, which exhibited a relatively strong regression performance, mainly focused on the facial bones, including dentition, whereas the NasNet-Large model mainly focused on the cranial base and midfacial bones. Differently structured CNNs can exhibit differences in performance, and the deep-learning model visualization shows that they have different focuses for bone age prediction.

External validation is the process of testing the original model using a different dataset [25]. Giuste *et al.* [26] showed that when evaluating the performance of an artificial-intelligence-based model, external validation can help generalize the model. Owing to differences in the LCR standards among different hospitals, the external validation performance of the trained deep learning model was lower than that of the internal validation. However, the external validation performance remained relatively good when using ROI images. Despite the differences in specifications owing to the different settings in each hospital, the ROIs containing the cervical vertebrae were similar to each other; therefore, it is expected that they would perform relatively well.

Facial bones and dentition contain many growth indicators. The degree of tooth calcification is highly correlated with skeletal maturity. Dadgar *et al.* [27] studied relationship between chronological age, dental age, and the CVM stages in male and female children. In a study of Korean pediatric patients [28], a high correlation between the degree of tooth calcification and bone maturity was found, and the possibility of using tooth maturity evaluation from panoramic radiographs as a supplementary means of evaluating bone maturity in growing children was explored. Meanwhile, the sphenoid bone and sphenoid sinus are also anatomical structures that change with growth and maturation. As an individual grows, the length of the ventral and dorsal sides of the pterygoid process and the size of the sphenoidal sinus also increase [29].

There is a routine usage of hand-wrist radiographs among growing children, exposing them to additional radiation. Therefore, developing a deep-learning model for skeletal maturation estimation based on LCRs can reduce the radiation exposure among orthodontic patients [30]. The current skeletal maturity estimationmethod using LCRs classifies the CVM stages according to the morphology of the second to fourth cervical vertebrae [31]. However, this method does not accurately reflect the continuous growth of children. In addition, it has low reproducibility because there may

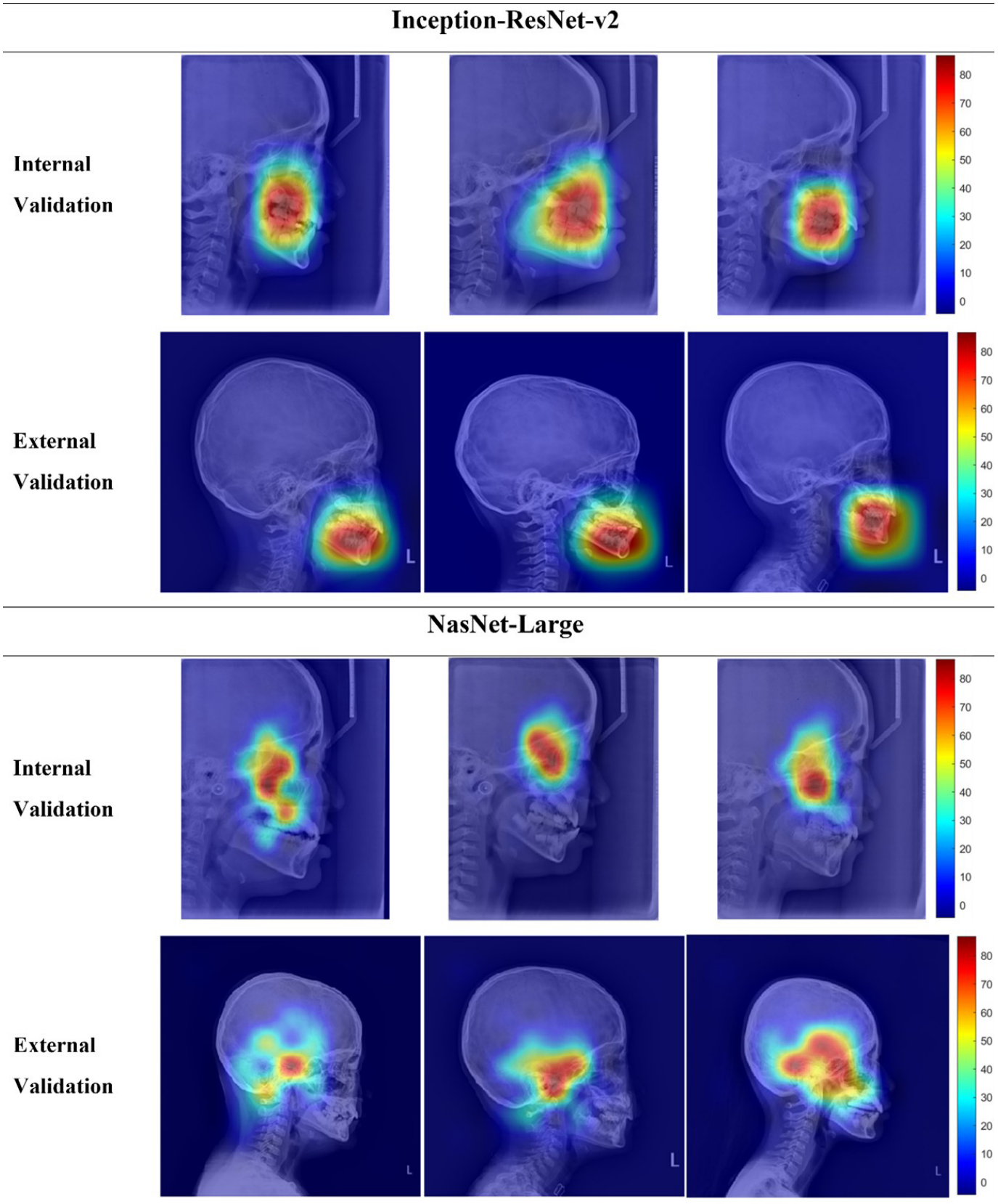

FIGURE 3. Grad-RAM images of the lateral cephalometric radiographs (LCRs) during the internal and external validation. In this case, the InceptionResNet-v2 model (top images) mainly focused on facial bones, including dentition, while the NasNet-Large model (bottom images) mainly focused on the cranial base and midfacial bones.

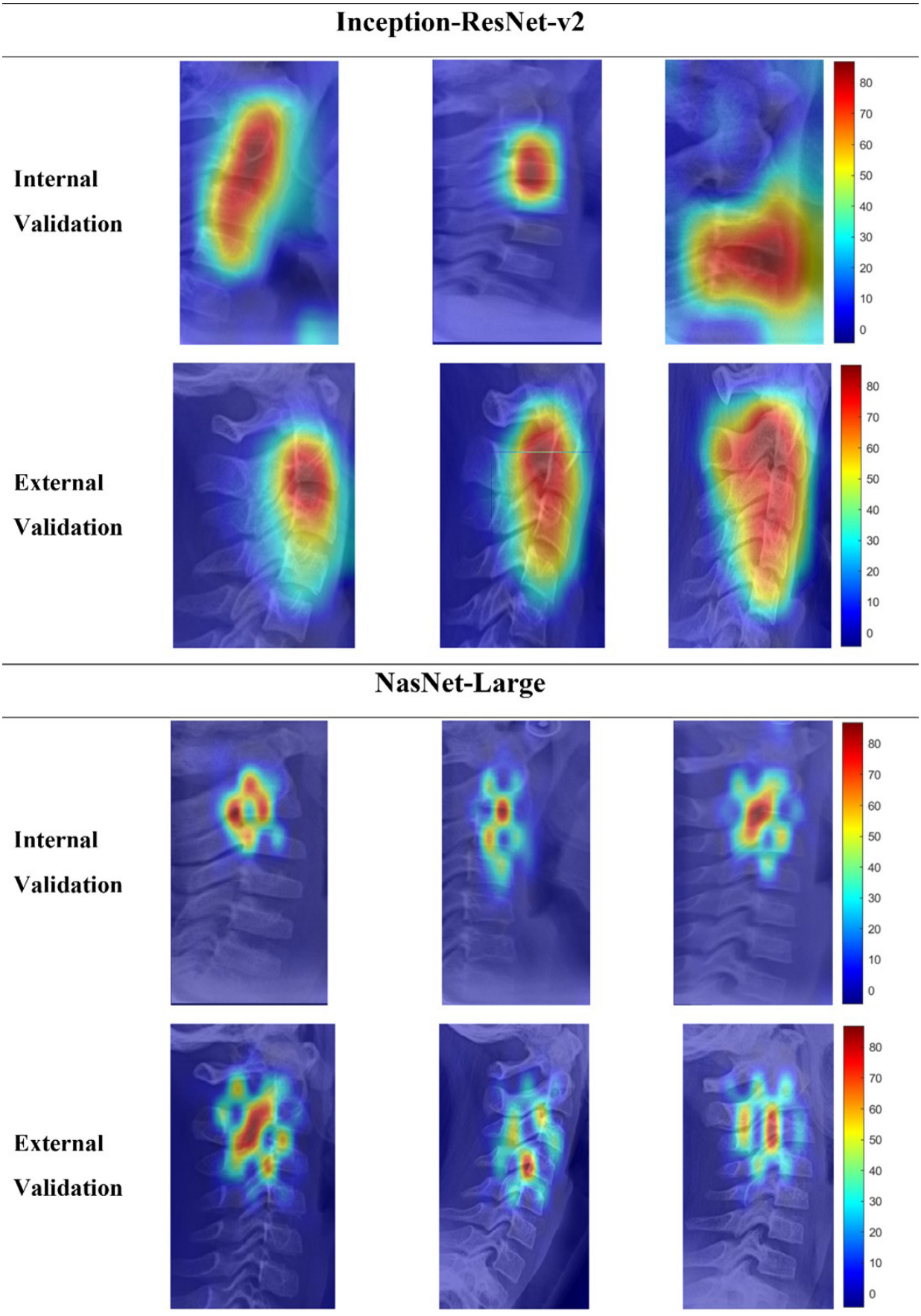

FIGURE 4. Grad-RAM images of the regions of interest (ROIs) including the cervical vertebrae during internal and external validation. In this case, the InceptionResNet-v2 model activation map exhibited a different pattern compared to that of the NasNet-Large model, which was somewhat sporadic.

be differences in the diagnosis depending on the skill of the observer [32, 33]. Furthermore, bone age identification rather than CVM stage classification has advantages in terms of establishing treatment plans, including the treatment timing. Bone age can be a relatively accurate indicator of an individual's growth compared to chronological age in that it has the advantage of predicting growth. According to Satoh *et al.* [34], the period of maximum growth is 13 years in male and 11 years in female Japanese patients. Although further research on the relationship between bone age and growth period is required, bone age can be a predictor of peak height velocity [35]. Moreover, bone age can be a predictor of adult height [36]. Recent studies that have employed deep-learning models for bone age estimation have used the model to predict the final adult height [37].

Despite its many advantages, this study has several limitations. First, data were collected retrospectively from two institutions; therefore, the amount of data was not uniform across age groups. The patient age distribution from both institutions does not exactly match each other. The generalization of the model would be possible under a similar age distribution from both institutions. Second, this study used data from South Korean children and adolescents; therefore, the generalizability of the deep-learning model is limited. In future work, multi-institutional and multi-country datasets should be obtained to create more generalized deep-learning models. In addition, although it was possible to estimate bone age through craniofacial bones and dentition in LCRs and to visualize the deep-learning models *via* the Grad-RAM method, it was not possible to quantitatively describe it. Therefore, further research is needed in terms of comparative analyses between specific bones *via* cone-beam computed tomography.

Despite these limitations, this study showed that LCRs can be used to predict the bone age for the growth and development of growing children without the need for conventional hand-wrist radiographs. The estimated bone age can be used to formulate a diagnostic plan for the patient and determine treatment timing.

5. Conclusions

In this study, we devised deep-learning models based on pre-trained CNNs for bone age estimation using LCRs and evaluated and generalized them through internal and external validations using multi-institutional data. The LCR-trained deep-learning models exhibited a better regression performance than the ROI-trained models, which included the cervical vertebrae. Despite the differences in performance, we found that bone age estimation is possible for craniofacial bones and dentition, in addition to the cervical vertebrae visible in LCRs. The findings of this study suggest that deep-learning models using LCRs could be a feasible alternative to hand-wrist radiographs for predicting bone age.

AVAILABILITY OF DATA AND MATERIALS

Not applicable.

AUTHOR CONTRIBUTIONS

JS and TJ—designed the study. SK and HS—performed the experiments; prepared the manuscript. JH and HS—analyzed the data. JS, EL and SP—provided critical feedback and approved the final manuscript. All authors have read and approved the final version of the manuscript.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

This study was approved by the Institutional Review Boards (IRB) of the Pusan National University Dental Hospital (IRB approval number: PNUDH-2023-09-007) and Ulsan University Hospital (IRB approval number: UUH-2023-07-026). By the Institutional Review Board of Pusan National University Dental Hospital and Ulsan University Hospital, patient consent was waived the need for individual informed consent for this study had a non-interventional retrospective design and all the data were analyzed anonymously; therefore, no written/verbal informed consent was obtained from the participants.

ACKNOWLEDGMENT

Not applicable.

FUNDING

This work was supported by an Ulsan University Hospital Research Grant (UUH-2023-10).

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

[1] Hunter CJ. The correlation of facial growth with body height and skeletal maturation at adolescence. The Angle Orthodontist. 1966; 36: 44–54.
[2] Hägg U, Taranger J. Maturation indicators and the pubertal growth spurt. American Journal of Orthodontics. 1982; 82: 299–309.
[3] Lamparski DG. Skeletal age assessment utilizing cervical vertebrae. American Journal of Orthodontics. 1975; 67: 458–459.
[4] Hägg U, Taranger J. Skeletal stages of the hand and wrist as indicators of the pubertal growth spurt. Acta Odontologica Scandinavica. 1980; 38: 187–200.
[5] Flores-Mir C, Nebbe B, Major PW. Use of skeletal maturation based on hand-wrist radiographic analysis as a predictor of facial growth: a systematic review. The Angle Orthodontist. 2004; 74: 118–124.
[6] Mito T, Sato K, Mitani H. Cervical vertebral bone age in girls. American Journal of Orthodontics and Dentofacial Orthopedics. 2002; 122: 380–385.
[7] Utama V, Soedarsono N, Yuniastuti M. Assessment of agreement between cervical vertebrae skeletal and dental age estimation with chronological age in an Indonesian population. The Journal of Forensic Odonto-stomatology. 2020; 38: 16.
[8] Lin L, Tang B, Cao L, Yan J, Zhao T, Hua F, *et al*. The knowledge, experience, and attitude on artificial intelligence-assisted cephalometric analysis: survey of orthodontists and orthodontic students. American Journal of Orthodontics and Dentofacial Orthopedics. 2023; 164: e97–e105.
[9] Law M, Seah J, Shih G. Artificial intelligence and medical imaging:

applications, challenges and solutions. Medical Journal of Australia. 2021; 214: 450.
[10] Barragán-Montero A, Javaid U, Valdés G, Nguyen D, Desbordes P, Macq B, *et al*. Artificial intelligence and machine learning for medical imaging: a technology review. Physica Medica. 2021; 83: 242–256.
[11] Schwendicke F, Golla T, Dreher M, Krois J. Convolutional neural networks for dental image diagnostics: a scoping review. Journal of Dentistry. 2019; 91: 103226.
[12] Seo H, Hwang J, Jung Y, Lee E, Nam OH, Shin J. Deep focus approach for accurate bone age estimation from lateral cephalogram. Journal of Dental Sciences. 2023; 18: 34–43.
[13] Sun S, Zhang R. Region of interest extraction of medical image based on improved region growing algorithm. 2017 International Conference on Material Science, Energy and Environmental Engineering (MSEEE 2017). Atlantis Press. 2017; 471–475.
[14] Jangam D, Kale P, Fatema S. Age determination using lateral cephalogram and orthopantomograph: a comparative study. Scholars Journal of Applied Medical Sciences. 2014; 2: 987–990.
[15] Murali K, Nirmal RM, Balakrishnan S, Shanmugam S, Altaf SK, Nandhini D. Age estimation using cephalometrics—a cross-sectional study among teenagers of salem district, Tamil Nadu. Journal of Pharmacy and Bioallied Sciences. 2023; 15: S725–S728.
[16] Zhang Z, Liu N, Guo Z, Jiao L, Fenster A, Jin W, *et al*. Ageing and degeneration analysis using ageing-related dynamic attention on lateral cephalometric radiographs. npj Digital Medicine. 2022; 5: 151.
[17] Lee SY, Im SA. Comparison of bone agesin early puberty: computerized greulich-pyle based bone age *vs.* sauvegrain method. Journal of the Korean Society of Radiology. 2022; 83: 1081.
[18] Seo H, Hwang J, Jeong T, Shin J. Comparison of deep learning models for cervical vertebral maturation stage classification on lateral cephalometric radiographs. Journal of Clinical Medicine. 2021; 10: 3591.
[19] Yu H, Yang LT, Zhang Q, Armstrong D, Deen MJ. Convolutional neural networks for medical image analysis: state-of-the-art, comparisons, improvement and perspectives. Neurocomputing. 2021; 444: 92–110.
[20] Wightman R, Touvron H, Jégou H. Resnet strikes back: an improved training procedure in timm. arXiv:2110.00476. 2021; 1–22.
[21] Zoph B, Vasudevan V, Shlens J, Le QV. Learning transferable architectures for scalable image recognition. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2018; 15: 8697–8710.
[22] Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D. Grad-CAM: visual explanations from deep networks *via* gradient-based localization. 2017 IEEE International Conference on Computer Vision (ICCV). 2017; 618–626.
[23] Kim EG, Oh IS, So JE, Kang J, Le VNT, Tak MK, *et al*. Estimating cervical vertebral maturation with a lateral cephalogram using the convolutional neural network. Journal of Clinical Medicine. 2021; 10: 5400.
[24] Kim JR, Shim WH, Yoon HM, Hong SH, Lee JS, Cho YA, *et al*. Computerized bone age estimation using deep learning based program: evaluation of the accuracy and efficiency. American Journal of Roentgenology. 2017; 209: 1374–1380.
[25] Ramspek CL, Jager KJ, Dekker FW, Zoccali C, van Diepen M. External validation of prognostic models: what, why, how, when and where? Clinical Kidney Journal. 2021; 14: 49–58.
[26] Giuste F, Shi W, Zhu Y, Naren T, Isgut M, Sha Y, *et al*. Explainable artificial intelligence methods in combating pandemics: a systematic review. IEEE Reviews in Biomedical Engineering. 2023; 16: 5–21.
[27] Dadgar S, Hadian H, Ghobadi M, Sobouti F, Rakhshan V. Correlations among chronological age, cervical vertebral maturation index, and Demirjian developmental stage of the maxillary and mandibular canines and second molars. Surgical and Radiologic Anatomy. 2021; 43: 131–143.
[28] Kim SJ, Song JS, Kim I, Kim S, Choi H. Correlation between dental and skeletal maturity in Korean children. The Journal of the Korean Academy of Pedtatric Dentistry. 2021; 48: 255–268.
[29] Jaworek-Troć J, Zarzecki M, Bonczar A, Kaythampillai LN, Rutowicz B, Mazur M, *et al*. Sphenoid bone and its sinus: anatomo-clinical review of the literature including application to FESS. Folia Medica Cracoviensia. 2019; 59: 45–59.
[30] Lai EH, Liu J, Chang JZ, Tsai S, Yao CJ, Chen M, *et al*. Radiographic assessment of skeletal maturation stages for orthodontic patients: hand-wrist bones or cervical vertebrae? Journal of the Formosan Medical Association. 2008; 107: 316–325.
[31] Baccetti T, Franchi L, McNamara JA. The cervical vertebral maturation (CVM) method for the assessment of optimal treatment timing in dentofacial orthopedics. Seminars in Orthodontics. 2005; 11: 119–129.
[32] McNamara JA, Franchi L. The cervical vertebral maturation method: a user's guide. The Angle Orthodontist. 2018; 88: 133–143.
[33] Gabriel DB, Southard KA, Qian F, Marshall SD, Franciscus RG, Southard TE. Cervical vertebrae maturation method: poor reproducibility. American Journal of Orthodontics and Dentofacial Orthopedics. 2009; 136: 478.e1–478.e7.
[34] Satoh M, Tanaka T. Bone age at onset of pubertal growth spurt and final height in normal children. Clinical Pediatric Endocrinology. 1995; 4: 129–136.
[35] Satoh M. Bone age: assessment methods and clinical applications. Clinical Pediatric Endocrinology. 2015; 24: 143–152.
[36] Reinehr T, Carlsson M, Chrysis D, Camacho-Hübner C. Adult height prediction by bone age determination in children with isolated growth hormone deficiency. Endocrine Connections. 2020; 9: 370–378.
[37] Suh J, Heo J, Kim SJ, Park S, Jung MK, Choi HS, *et al*. Bone age estimation and prediction of final adult height using deep learning. Yonsei Medical Journal. 2023; 64: 679.

How to cite this article: Suhae Kim, Jonghyun Shin, Eungyung Lee, Soyoung Park, Taesung Jeong, JaeJoon Hwang, *et al*. Comparative analysis of deep-learning-based bone age estimation between whole lateral cephalometric and the cervical vertebral region in children. Journal of Clinical Pediatric Dentistry. 2024; 48(4): 191-199. doi: 10.22514/jocpd.2024.093.