



OPEN

## Prediction of Fishman's skeletal maturity indicators using artificial intelligence

Harim Kim<sup>1</sup>, Cheol-Soon Kim<sup>2</sup>, Ji-Min Lee<sup>1</sup>, Jae Joon Lee<sup>3</sup>, Jiyeon Lee<sup>3</sup>, Jung-Suk Kim<sup>2,4</sup>✉ & Sung-Hwan Choi<sup>1,4</sup>✉

The present study aimed to evaluate the performance of automated skeletal maturation assessment system for Fishman's skeletal maturity indicators (SMI) for the use in dental fields. Skeletal maturity is particularly important in orthodontics for the determination of treatment timing and method. SMI is widely used for this purpose, as it is less time-consuming and practical in clinical use compared to other methods. Thus, the existing automated skeletal age assessment system based on Greulich and Pyle and Tanner-Whitehouse3 methods was further developed to include SMI using artificial intelligence. This hybrid SMI-modified system consists of three major steps: (1) automated detection of region of interest; (2) automated evaluation of skeletal maturity of each region; and (3) SMI stage mapping. The primary validation was carried out using a dataset of 2593 hand-wrist radiographs, and the SMI mapping algorithm was adjusted accordingly. The performance of the final system was evaluated on a test dataset of 711 hand-wrist radiographs from a different institution. The system achieved a prediction accuracy of 0.772 and mean absolute error and root mean square error of 0.27 and 0.604, respectively, indicating a clinically reliable performance. Thus, it can be used to improve clinical efficiency and reproducibility of SMI prediction.

The assessment of skeletal maturation is essential not only in medical or forensic, but also in dental fields, especially in orthodontics and pediatric dentistry. In growing patients, skeletal maturity is an important factor to consider for optimal timing and success of orthodontic treatment. Especially in patients with skeletal discrepancies between the upper and lower jaws who require growth modification, the timing of orthodontic treatment is essential for the treatment outcomes, as the redirection of jaw growth is effective only during certain stages of growth<sup>1</sup>. Therefore, the evaluation of skeletal maturation is indispensable for orthodontic diagnostics and treatment planning, as well as for the evaluation of treatment progress and prognosis.

It has been reported in previous studies that maturational development shows large individual variations, leading to discrepancies between the chronological and skeletal age<sup>2,3</sup>. Hand-wrist radiographs have been widely utilized for the evaluation of skeletal maturation, given the relatively low dose of radiation exposure and quick and uncomplicated and non-invasive procedure of x-ray taking. It is based on the principle that the osseous development in the hand and wrist areas is closely related with general maturational development<sup>4</sup>. The recognition of osseous features in the hand-wrist areas characteristic of each maturation level underlies the estimation of skeletal age. Different methods have been proposed for the analysis of hand-wrist radiographs.

As one of the most widely used methods, Greulich and Pyle (GP) is based on an atlas consisting of standard hand-wrist radiographs for each sex and skeletal age. The reference images were obtained from middle-classed Caucasian children between 1931 and 1942 in Ohio, United States<sup>5</sup>. While this atlas-based comparative method is convenient for clinical use, it assumes that the sequence and pattern of ossification and the corresponding maturation level are identical in all individuals<sup>6</sup>. On the other hand, Tanner-Whitehouse (TW) method is a bone-specific technique, which evaluates and scores multiple skeletal maturity indicators and determines the skeletal age according to the total score, referred to as skeletal maturity score (SMS). There are several variations of TW, which differ in the number of evaluated sites and the relationship between SMS and the corresponding skeletal age. TW3 was published in 2001 to reflect the secular trend of maturational development<sup>7</sup>. While TW3 is less sensitive to individual variations compared with GP, its application in everyday clinical practice is rather difficult, because the evaluation of multiple maturity indicators is complicated and time-consuming.

<sup>1</sup>Department of Orthodontics, Institute of Craniofacial Deformity, Yonsei University College of Dentistry, Seoul, Republic of Korea. <sup>2</sup>Private Clinic, Seongnam-si, Republic of Korea. <sup>3</sup>Crescom, Seongnam-si, Republic of Korea. <sup>4</sup>These authors contributed equally: Jung-Suk Kim and Sung-Hwan Choi. ✉email: jsfokim@daum.net; selfexam@yuhs.ac

Skeletal maturity indicators (SMI) introduced by Fishman, an American orthodontist, in 1981 are widely used for the assessment of skeletal maturity in orthodontic patients. In contrast to the forensic or medical fields, the determination of skeletal age in number of years, or the estimation of final adult height is not of great importance in orthodontics. The main focus is rather on the timing of pubertal growth spurt and growth completion, as the course of orthodontic treatment may vary depending on these factors. Therefore, the analytic methods proposed for the application in the orthodontic field, such as Fishman's SMI<sup>4</sup> and Hägg and Taranger method<sup>8</sup> are simplified to focus on the key stages of maturational development, which are relevant for decision makings during orthodontic treatment<sup>9</sup>.

With the advances in artificial intelligence (AI), numerous studies have been reported concerning the use of deep learning and neural networks for the evaluation and analysis of radiographs in various medical and dental fields. While the interest in the automation for clinical efficiency is ever increasing, there are also concerns about the accuracy and clinical validity of such automated systems<sup>10</sup>. The present study aimed to evaluate the performance of an AI-based automated assessment system for SMI prediction.

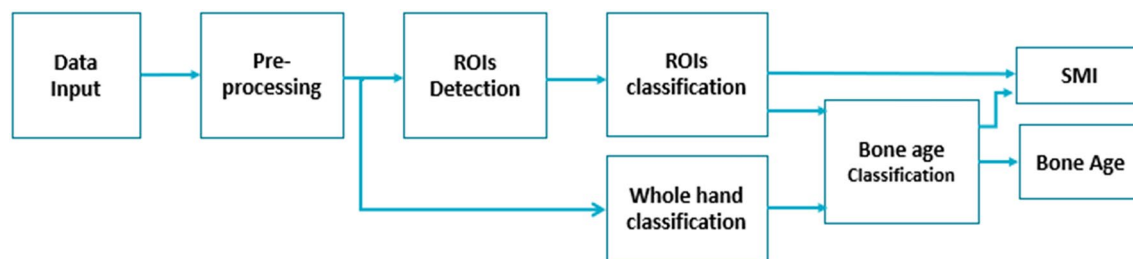
## Materials and methods

**Model development.** The existing GP and TW3-based bone age assessment system (MediAI-BA, Crescom Inc., Korea)<sup>11</sup> was further trained and modified to include the evaluation of SMI. The pre-trained system employed a hybrid approach combining the advantages of TW3 and GP methods to improve the performance of skeletal age prediction<sup>11,12</sup>. The automated assessment system for SMI was developed based on the same principle applied to the GP and TW3-based system, consisting of three major steps: (1) automated detection of region of interest (ROI); (2) automated evaluation of skeletal maturity of each region; and (3) SMI stage mapping (Fig. 1). In this version, the detection algorithm for each ROI was developed using a detection Transformer model<sup>13</sup>, which utilizes learnable queries to search image features from the output of Transformer encoders and bipartite graph matching for detection. For automated evaluation of skeletal maturity of each region, the classification algorithm was developed using Swin Transformer model<sup>14</sup>, which is a hierarchical vision transformer deep learning model with a window-shifting method. Data augmentation techniques, such as random resize, random crop, brightness adjustment, contrast adjustment, CLAHE, and random noise, were applied during deep learning training.

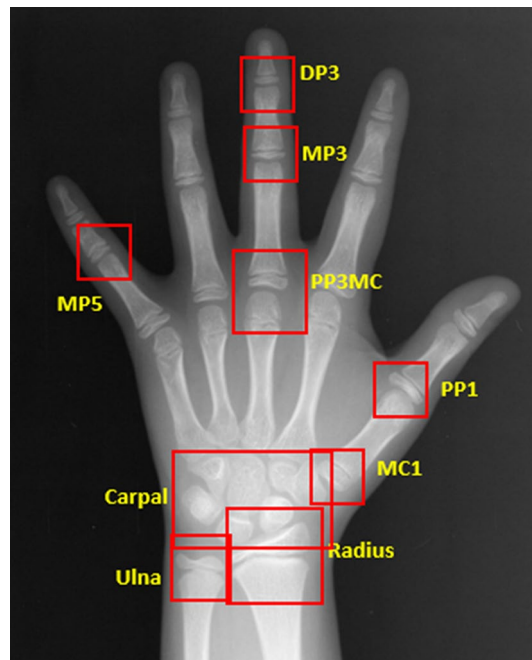
In addition to the major areas of growth plate evaluated for the TW3 method and the carpal region, all regions subject to the evaluation of SMI were included as ROIs (Fig. 2). In particular, MP5 (fifth middle phalanx) and PP1 (first proximal phalanx) were added by utilizing the detection algorithm based on the work by Li et al.<sup>13</sup> RetinaNet, a deep convolutional neural network (CNN), was utilized for the automated detection of ROIs. Subsequently, each of the detected regions was analyzed for skeletal maturity level based on the morphological changes, such as epiphyseal widening, presence of sesamoid, epiphyseal capping and epiphyseal fusion. For image classification a vision transformer-based deep learning model was utilized. The final skeletal age was calculated by analyzing and integrating the probability of skeletal maturity of each area. Irrespective of the skeletal age prediction, skeletal maturity level of each of the six regions relevant for SMI evaluation was analyzed for SMI stage mapping. The final SMI stage was calculated by integrating the skeletal maturity and age of each ROI, and the final skeletal age. Both SMI and skeletal age were outputted as results (Fig. 3).

The primary verification of this SMI-modified system was carried out using the verification dataset, and the system underwent further modifications and adjustments according to the results. Subsequently, the accuracy of the finally modified system was tested using a new set of hand-wrist radiographs from a different institution, which had not been introduced to the system previously.

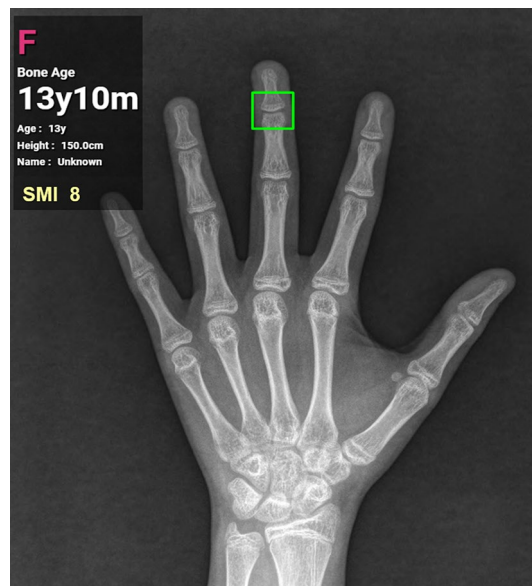
**Subjects.** The dataset used for the primary verification consisted of hand-wrist radiographs obtained from 2593 growing patients as part of orthodontic diagnosis between April 2019 and July 2022 at a private orthodontic clinic in Korea. The accuracy of the final SMI-modified system was tested using hand-wrist radiographs of 711 patients, who visited the Department of Orthodontics, Yonsei University Dental Hospital in Seoul, Korea between August 2019 and February 2021. The exclusion criteria for both of the datasets were as follows: (1) congenital craniofacial anomalies or syndromes affecting growth, (2) history of growth hormone treatment, (3)



**Figure 1.** Overview of the steps involved in the modified SMI-based skeletal maturation assessment system. In the pre-processing step, the image size of hand-wrist radiographs is adjusted, and right-hand radiographs are mirrored. In addition to the detection and classification of each ROI, the evaluation of the whole hand as a holistic image is performed to enhance the performance by including maturity features that are not parts of the ROIs.



**Figure 2.** Automated detection of regions of interest. In addition to the ROIs used in the TW3 method, all of the six ROIs used for Fishman's SMI, as well as the carpal bones are detected. (DP3, 3rd distal phalanx; MP3, 3rd middle phalanx; MP5, 5th middle phalanx; PP3MC, 3rd proximal phalanx/metacarpal; PP1, 1st proximal phalanx, MC1, 1st metacarpal).



**Figure 3.** Result screen displaying the SMI stage and bone age predicted by the SMI-modified automated skeletal maturation assessment system.

history of systemic diseases and medication, and (4) insufficient quality of radiographs affecting the evaluation of SMI.

The study protocol was approved by the Institutional Review Board of Yonsei Dental Hospital (IRB No. 2-2022-0030), and adhered to the Declaration of Helsinki (2013). The requirement of informed consent and ethical approval was waived in view of the retrospective and non-interventional design of the study. All procedures were performed irrespective of the study as part of the routine care.

**Evaluation of SMI.** SMI staging performed by the observers strictly followed the descriptions provided in the original article by Fishman. Eleven different skeletal maturity indicators, which are numbered through SMI 1 to 11, were recognized by assessing six sites of the hand and wrist areas<sup>4</sup>. In addition, SMI 0 was included to express the maturation level earlier than that of SMI 1<sup>15,16</sup>. For each dataset, the evaluation of SMI was undertaken by three different orthodontists, who have multiple years of clinical experience in orthodontics. The ground-truth SMI was determined as the stage that was the most frequently chosen by the observers. If no agreement could be reached, a fourth observer was asked for additional evaluation. To minimize any bias, the observers were unaware of the chronological age of the patients and the SMI stage determined by other observers at the time of evaluation.

**Statistical analysis.** For statistical analysis and data visualization, Microsoft Excel for Windows (version 2019; Microsoft Corp., WA, USA) and Scikit-learn library in Python were used. The performance of the automated SMI assessment system was evaluated by computing mean absolute error (MAE) and root mean square error (RMSE) using the following formulas.

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - x_i| \quad (1)$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - x_i)^2} \quad (2)$$

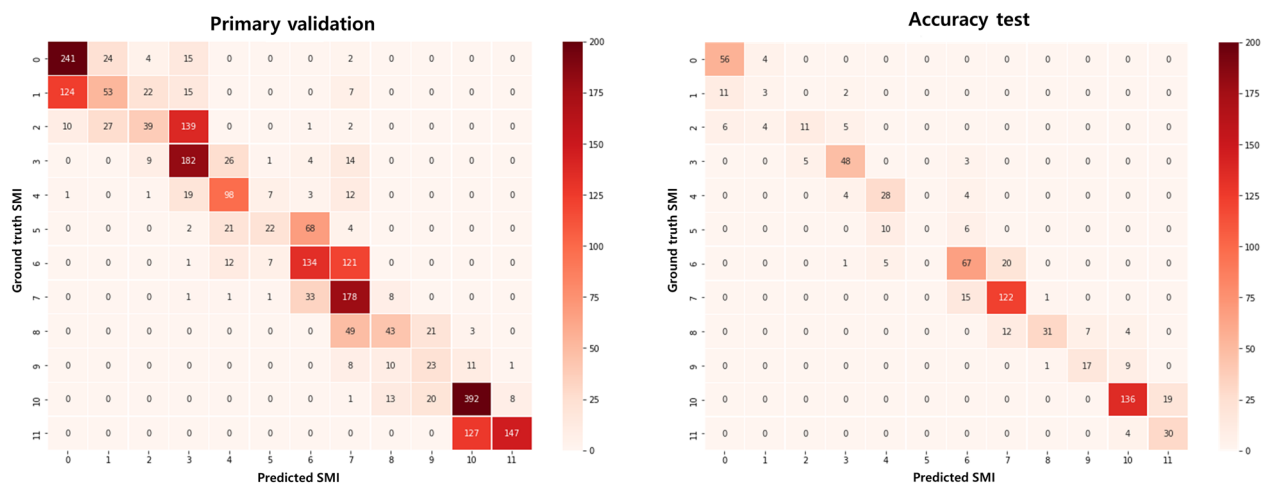
In both formulas,  $n$  is the number of evaluated hand-wrist radiographs, and  $y_i$  and  $x_i$  represent the ground-truth and AI-predicted SMI, respectively. MAE and RMSE were calculated for the overall performance of the system, as well as for each of the SMI stages.

Furthermore, confusion matrix was generated for visual presentation of the correctly and incorrectly predicted SMI. The overall prediction accuracy was calculated by dividing the number of correct predictions by the total number predictions. In addition, sensitivity and specificity of each SMI stage, as well as the overall sensitivity and specificity were obtained.

## Results

The overall prediction accuracy obtained following the primary validation of the automated SMI assessment system was 0.599 with MAE of 0.499. The confusion matrix exhibited a diagonal pattern, which implies an overall correct prediction of SMI (Fig. 4). However, the AI-predicted SMI 7 showed a large range of incorrect predictions. According to the results of the primary validation, the algorithm of SMI mapping underwent further adjustments to improve the performance of the system.

Upon adjusting the algorithm, the performance of the final model was evaluated using the test dataset consisting of hand-wrist radiographs of 711 patients (361 males and 350 females). The patient age at the obtention of radiograph ranged from 6 to 18 years, with the mean age of  $11.93 \pm 2.43$  years. As can be seen in the confusion matrix, most of the predictions remained in the diagonal, with only a few outliers (Fig. 4). The overall prediction accuracy was 0.772, indicating that the AI-predicted SMI corresponded to the ground-truth SMI in 77.2% of the cases. Considering the uneven distribution of the SMI stages in the dataset, balanced accuracy was obtained additionally. It was slightly smaller than the prediction accuracy with 0.704. When the tolerance range was expanded to within 1 SMI stage, the prediction accuracy increased to 0.963. The overall MAE and RMSE were 0.27 and 0.604, respectively. As a linear solution, MAE weighs errors equally in the average regardless of their magnitude. On the other hand, RMSE weighs large errors more heavily. Both MAE and RMSE were the smallest



**Figure 4.** Confusion matrices showing the results of primary validation and accuracy test following fine tuning.

for SMI 0 with 0.067 and 0.258, respectively. The largest MAE was calculated for SMI 5 (1.0) followed by SMI 1 (0.938) and 2 (0.846). On the other hand, RMSE was bigger for SMI 2 (1.177) than SMI 1 (1.09) and SMI 5 (1.0). Accordingly, the highest sensitivity was found for SMI 0, and the lowest for SMI 5, followed by SMI 1. The overall specificity was 0.978, with the highest values found for SMI 5 (1.0) and SMI 8 (0.997), and the lowest for SMI 7 (0.944) (Table 1).

## Discussion

The present study is the first to report an AI-based automated system for the assessment of Fishman's SMI, which is widely used in dental fields, especially in orthodontics. In recent years, an increasing number of studies have been reported on the automation of skeletal maturation assessment using AI for the improvement of clinical efficiency and reproducibility. The vast majority of these proposed systems are based on TW3 method<sup>12,17–19</sup>, as it has several advantages over other methods. Firstly, it takes into account the variability of skeletal maturation pattern in different populations and ethnic groups. Unlike its predecessor, TW2, which derived from a sample of British children<sup>20</sup>, TW3 was developed using additional data from multiple ethnic groups and populations<sup>7</sup>. Later studies have also shown its validity in various populations, including Korean children<sup>21–23</sup>. Furthermore, TW3 method offers a comprehensive evaluation of skeletal maturity by assessing multiple bones, leading to a more thorough analysis of skeletal development and improved reliability. On the other hand, a major drawback of TW3 is the complexity and time required to obtain results. This limitation has been addressed through the use of AI-based systems. Another limitation of TW3 is that it does not account for changes in the trend of maturational development over time. This could lead to potential obsolescence of the method, much like its predecessor TW2. Therefore, periodic updates are necessary to maintain the validity and reliability of the TW method<sup>7</sup>.

Similarly, the GP method is susceptible to changes in the trends of maturational development, as well as to variations resulting from differences in ethnicity, regional factors, and environmental influences<sup>24–26</sup>. Unlike the Tanner-Whitehouse method, which has undergone revisions since its inception, the GP method has not been updated since its introduction in 1959<sup>5</sup>. In other words, it is solely based on the initial reference hand-wrist radiographs of Caucasian children obtained over 80 years ago. According to the study by Mansouvar et al., the GP method is reliable for Caucasian and Hispanic children, but not for African/American and Asian groups<sup>27</sup>. A study with the sample of Korean population also concluded that the rates of skeletal development provided by GP is not applicable to Korean children<sup>28</sup>.

In contrast to TW3 and GP, Fishman's SMI provides a staging system for assessing maturation levels, which is not reliant on skeletal age<sup>4</sup>. Consequently, SMI is not subject to fluctuations in the trend of maturational development and differences arising from factors such as ethnicity. In simpler terms, SMI allows for an intuitive determination of an individual's skeletal maturity level, without necessitating consideration of additional factors. When assessing skeletal maturity for orthodontic purposes, the level of maturation in relation to chronological age is of little importance. As a result, obtaining skeletal age is not typically necessary, in contrast to its importance in medical or forensic fields. Nevertheless, SMI, like other methods, has its limitations. Since a single skeletal indicator is assigned for each stage, variations in the appearance sequence of skeletal maturity indicators or unclear indications of these indicators can lead to misstaging.

Clinicians often encounter hand-wrist radiographs with osseous maturational characteristics that are ambiguous to be classified as a certain SMI stage<sup>29</sup>. In case of individual variations in the sequence of skeletal maturation that do not comply with the descriptions by Fishman, SMI stage may be over- or underestimated depending on the observer, and reproducibility and reliability of SMI may be affected. It has been reported in a previous study that the prediction accuracy was relatively low for SMI stages 5 and 6. Insufficient amount of data, and large inter-observer variabilities were considered as the possible reasons<sup>15</sup>. Similarly, in the present study, the prediction accuracy for SMI stage 5 was found to be lower compared to other SMI stages. This may be related with the

SMI	Sensitivity	Specificity	MAE	RMSE
0	0.933	0.972	0.067	0.258
1	0.188	0.990	0.938	1.09
2	0.423	0.993	0.846	1.177
3	0.856	0.982	0.25	0.756
4	0.778	0.978	0.333	0.745
5	0.000	1	1	1
6	0.720	0.955	0.355	0.726
7	0.884	0.944	0.116	0.341
8	0.574	0.997	0.5	0.805
9	0.630	0.990	0.37	0.609
10	0.877	0.969	0.123	0.350
11	0.882	0.972	0.118	0.343
Overall	0.646	0.978	0.27	0.604
Prediction accuracy 0.772			Balanced accuracy 0.646	

**Table 1.** Performance of automated SMI assessment system. SMI, skeletal maturity indicators; MAE, mean absolute error; RMSE, root mean square error.

fact that there was no radiograph in the dataset used for the accuracy test that was predicted as SMI 5. However, the prediction accuracy of SMI 5 was calculated to be only 0.19 also in the primary validation, which was carried out with a larger dataset. In other words, SMI stages that are more likely to deviate from the proposed sequence are more prone to higher inter-observer variability and lower prediction accuracy. The system introduced in the present study is a hybrid approach that evaluates maturational level by integrating the GP, TW3, and SMI methods. This approach compensates for the known limitations of each system and enhances the accuracy and reproducibility of SMI predictions.

Recent advances in technology have led a notable surge in the integration of AI into dental practice for tasks such as diagnosis, radiographic analysis and treatment planning<sup>30</sup>. Consequently, there have been attempts to streamline and accelerate the process of skeletal maturity assessment through the application of AI<sup>31</sup>, as the manual assessment has been subject to criticism for its tediousness and intra- and interobserver variabilities<sup>32,33</sup>. Previous studies have demonstrated clinically reliable performance of deep learning-based automated systems for assessing skeletal maturity<sup>11,17,18,34–36</sup>. However, the majority of introduced models are based on TW3 or GP methods. Few automated systems have been proposed for the assessment of skeletal maturity using SMI. However, previous studies that investigated SMI in relation to AI focused rather on the prediction of SMI using the radiographic images of cervical vertebrae<sup>15,37</sup>.

The performance of various automated skeletal maturation assessment systems has been evaluated in previous studies. According to these studies, the AI-predicted skeletal age was not significantly different from the skeletal age assessed by experts<sup>11,19</sup>. The range of MAE reported in the literature varies from 0.39 to 2.41 years depending on the study<sup>17,34–36,38–42</sup>. It is notable that the models proposed during the last decade<sup>17,34,35,38,42</sup> show better performance with smaller MAE compared with the models introduced earlier<sup>36,39–41</sup>. Since the MAE computed in the present study does not refer to skeletal age, but SMI stage, it cannot be directly compared with the results of previous studies. According to the data provided in Fishman's study, the mean interval between SMI stages is 0.61 years for the female and 0.64 years for the male sex<sup>4</sup>. Based on this information, MAE of 0.27 SMI stage can be converted to approximately 0.169 years. Although this conversion may not be accurate, as it does not consider the differences in the size of the interval between the stages, the results suggest that the SMI-modified automated skeletal maturation system shows a satisfactory performance compared with previous systems.

The present study has several limitations. The size of the study population as a whole was sufficient, however, the number of observations was relatively small for some of the SMI stages. This may have affected the reliability of the accuracy measured for these subgroups. Furthermore, the data were collected retrospectively. While the data used for AI-training were collected from different institutions including different ethnicities, the datasets used for the primary validation and final evaluation consisted of hand-wrist radiographs only of a single ethnic group. Therefore, the results of the present study may not reflect possible differences between various ethnicities, or populations. In order to validate the results of the present study, and to further improve the accuracy of the proposed system, future studies with larger datasets from multiple institutions and populations are required.

## Conclusion

The hybrid SMI-reinforced automated skeletal maturation assessment system introduced in the present study was shown to deliver clinically reliable prediction of SMI with a very low prediction error. Hence, it can be efficiently utilized in dental fields to enhance clinical efficiency. It can also assist clinicians in improving the reproducibility of skeletal maturation assessment.

## Data availability

The data underlying this article cannot be publicly shared to protect the privacy of the individuals participating in the study. The data will be shared at a reasonable request to the corresponding author.

Received: 16 December 2022; Accepted: 6 April 2023

Published online: 11 April 2023

## References

- Proffit, W. R. The timing of early treatment: An overview. *Am. J. Orthod. Dentofac. Orthop.* **129**, S47–S49 (2006).
- Fishman, L. S. Chronological versus skeletal age, an evaluation of craniofacial growth. *Angle Orthod.* **49**, 181–189 (1979).
- Dahlberg, P. S. *et al.* A systematic review of the agreement between chronological age and skeletal age based on the Greulich and Pyle atlas. *Eur. Radiol.* **29**, 2936–2948 (2019).
- Fishman, L. S. Radiographic evaluation of skeletal maturation: A clinically oriented method based on hand-wrist films. *Angle Orthod.* **52**, 88–112 (1982).
- Greulich, W. W. & Pyle, S. I. *Radiographic Atlas of Skeletal Development of the Hand and Wrist* (Stanford University Press, 1959).
- Gilli, G. The assessment of skeletal maturation. *Horm. Res. Paediatr.* **45**, 49–52 (1996).
- Tanner, J. M., Healy, M. J. R., Cameron, N. & Goldstein, H. *Assessment of Skeletal Maturity and Prediction of Adult Height (TW3 Method)* (Saunders, 2001).
- Hagg, U. & Taranger, J. Maturation indicators and the pubertal growth spurt. *Am. J. Orthod.* **82**, 299–309. [https://doi.org/10.1016/0002-9416\(82\)90464-x](https://doi.org/10.1016/0002-9416(82)90464-x) (1982).
- Jeon, J. Y., Kim, C. S., Kim, J. S. & Choi, S. H. Correlation and correspondence between skeletal maturation indicators in hand-wrist and cervical vertebra analyses and skeletal maturity score in Korean adolescents. *Children* <https://doi.org/10.3390/children8100910> (2021).
- Rubin, D. A. Assessing bone age: A paradigm for the next generation of artificial intelligence in radiology. *Radiology* **301**, 700–701. <https://doi.org/10.1148/radiol.202111339> (2021).
- Lee, K.-C. *et al.* Clinical validation of a deep learning-based hybrid (Greulich-Pyle and modified Tanner-Whitehouse) method for bone age assessment. *Korean J. Radiol.* **22**, 2017 (2021).
- Lee, B. D. & Lee, M. S. Automated bone age assessment using artificial intelligence: The future of bone age assessment. *Korean J. Radiol.* **22**, 792–800. <https://doi.org/10.3348/kjr.2020.0941> (2021).

13. Li, F. et al. Dn-detr: Accelerate detr training by introducing query denoising. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 13619–13627 (2022).
14. Liu, Z. et al. Swin transformer: Hierarchical vision transformer using shifted windows. *Proceedings of the IEEE/CVF international conference on computer vision*, 10012–10022 (2021).
15. Kim, D. W. et al. Prediction of hand-wrist maturation stages based on cervical vertebrae images using artificial intelligence. *Orthod. Craniofac. Res.* **24**, 68–75 (2021).
16. Chatzigianni, A. & Halazonetis, D. J. Geometric morphometric evaluation of cervical vertebrae shape and its relationship to skeletal maturation. *Am. J. Orthod. Dentofac. Orthop.* **136**, 481 e481. <https://doi.org/10.1016/j.ajodo.2009.04.017> (2009).
17. Spampinato, C., Palazzo, S., Giordano, D., Aldinucci, M. & Leonardi, R. Deep learning for automated skeletal bone age assessment in X-ray images. *Med. Image Anal.* **36**, 41–51. <https://doi.org/10.1016/j.media.2016.10.010> (2017).
18. Dallora, A. L. et al. Bone age assessment with various machine learning techniques: A systematic literature review and meta-analysis. *PLoS One* **14**, e0220242. <https://doi.org/10.1371/journal.pone.0220242> (2019).
19. Shin, N. Y. et al. Evaluation of the clinical efficacy of a TW3-based fully automated bone age assessment system using deep neural networks. *Imaging Sci. Dent.* **50**, 237–243. <https://doi.org/10.5624/isd.2020.50.3.237> (2020).
20. Tanner, J. M. & Whitehouse, R. H. Clinical longitudinal standards for height, weight, height velocity, weight velocity, and stages of puberty. *Arch. Dis. Child.* **51**, 170–179 (1976).
21. Kim, J. R., Lee, Y. S. & Yu, J. Assessment of bone age in prepubertal healthy Korean children: Comparison among the Korean standard bone age chart, Greulich-Pyle method, and Tanner-Whitehouse method. *Korean J. Radiol.* **16**, 201–205 (2015).
22. Yuh, Y.-S., Chou, T. Y. & Tung, T.-H. Bone age assessment: Large-scale comparison of Greulich-Pyle method and Tanner-Whitehouse 3 method for Taiwanese children. *J. Chin. Med. Assoc.* **86**, 246–253 (2023).
23. Kim, S. Y., Oh, Y. J., Shin, J. Y., Rhie, Y. J. & Lee, K. H. Comparison of the Greulich-Pyle and Tanner Whitehouse (TW3) methods in bone age assessment. *J. Korean Soc. Pediatr. Endocrinol.* **13**, 50–55 (2008).
24. Ontell, F. K., Ivanovic, M., Ablin, D. S. & Barlow, T. W. Bone age in children of diverse ethnicity. *AJR Am. J. Roentgenol.* **167**, 1395–1398. <https://doi.org/10.2214/ajr.167.6.8956565> (1996).
25. Zhang, A., Sayre, J. W., Vachon, L., Liu, B. J. & Huang, H. K. Racial differences in growth patterns of children assessed on the basis of bone age. *Radiology* **250**, 228–235. <https://doi.org/10.1148/radiol.2493080468> (2009).
26. Benjavongkulchai, S. & Pittayapat, P. Age estimation methods using hand and wrist radiographs in a group of contemporary Thais. *Forensic Sci. Int.* **287**, 218 e211–218 e218. <https://doi.org/10.1016/j.forsciint.2018.03.045> (2018).
27. Mansourvar, M. et al. The applicability of Greulich and Pyle atlas to assess skeletal age for four ethnic groups. *J. Forensic Leg. Med.* **22**, 26–29 (2014).
28. Hwang, J. et al. Re-assessment of applicability of Greulich and Pyle-based bone age to Korean children using manual and deep learning-based automated method. *Yonsei Med. J.* **63**, 683 (2022).
29. Chung, D. D. & Ghanouni, S. Comparison of abnormalities in the sequence of growth stages in the skeletal maturity index vs cervical vertebral maturation: A retrospective study. *Angle Orthod.* **92**, 353–357 (2022).
30. Thurzo, A. et al. Where is the artificial intelligence applied in dentistry? Systematic review and literature analysis. *Healthcare* **10**, 1269 (2022).
31. Rijn, R. V. & Thodberg, H. Bone age assessment: Automated techniques coming of age?. *Acta Radiol.* **54**, 1024–1029 (2013).
32. Roche, A., Rohmann, C. G., French, N. Y. & Dávila, G. H. Effect of training on replicability of assessments of skeletal maturity (Greulich-Pyle). *Am. J. Roentgenol.* **108**, 511–515 (1970).
33. Johnson, G. F., Dorst, J. P., Kuhn, J. P., Roche, A. F. & Dávila, G. H. Reliability of skeletal age assessments. *Am. J. Roentgenol.* **118**, 320–327 (1973).
34. Giordano, D., Kavasidis, I. & Spampinato, C. Modeling skeletal bone development with hidden Markov models. *Comput. Methods Progr. Biomed.* **124**, 138–147. <https://doi.org/10.1016/j.cmpb.2015.10.012> (2016).
35. Giordano, D., Spampinato, C., Scarciofalo, G. & Leonardi, R. An automatic system for skeletal bone age measurement by robust processing of carpal and epiphysal/metaphysal bones. *IEEE Trans. Instrum. Meas.* **59**, 2539–2553 (2010).
36. Pietka, E., Pospiech-Kurkowska, S., Gertych, A. & Cao, F. Integration of computer assisted bone age assessment with clinical PACS. *Comput. Med. Imaging Graph.* **27**, 217–228. [https://doi.org/10.1016/s0895-6111\(02\)00076-9](https://doi.org/10.1016/s0895-6111(02)00076-9) (2003).
37. Xie, L. et al. Development of a multi-stage model for intelligent and quantitative appraising of skeletal maturity using cervical vertebrae cone-beam CT images of Chinese girls. *Int. J. Comput. Assist. Radiol. Surg.* **17**, 761–773 (2022).
38. Kashif, M., Deserno, T. M., Haak, D. & Jonas, S. Feature description with SIFT, SURF, BRIEF, BRISK, or FREAK? A general question answered for bone age assessment. *Comput. Biol. Med.* **68**, 67–75. <https://doi.org/10.1016/j.compbiomed.2015.11.006> (2016).
39. Gertych, A., Zhang, A., Sayre, J., Pospiech-Kurkowska, S. & Huang, H. K. Bone age assessment of children using a digital hand atlas. *Comput. Med. Imaging Gr.* **31**, 322–331. <https://doi.org/10.1016/j.compmedimag.2007.02.012> (2007).
40. Hsieh, C. W., Jong, T. L., Chou, Y. H. & Tiu, C. M. Computerized geometric features of carpal bone for bone age estimation. *Chin. Med. J. (Engl)* **120**, 767–770 (2007).
41. Pietka, E. et al. Computer-assisted bone age assessment: Image preprocessing and epiphyseal/metaphyseal ROI extraction. *IEEE Trans. Med. Imaging* **20**, 715–729. <https://doi.org/10.1109/42.938240> (2001).
42. Son, S. J. et al. TW3-based fully automated bone age assessment system using deep neural networks. *IEEE Access* **7**, 33346–33358 (2019).

## Acknowledgements

This research was supported by a grant from the Korea Health Technology R&D Project through the Korea Health Industry Development Institute, funded by the Ministry of Health & Welfare, Republic of Korea (Grant Number: HI20C0611).

## Author contributions

H.K., J.S.K. and S.H.C. conceived and designed the study, and H.K., J.S.K., C.S.K., J.M.L. and S.H.C. performed all the measurements. J.S.K. and J.J.L. provided the data required for the study. H.K., J.Y.L. and J.J.L. interpreted and analyzed the data. H.K. and S.H.C. drafted the manuscript. J.S.K. and J.J.L. provided writing assistance and critically revised the manuscript. All authors have reviewed and approved the final manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Correspondence** and requests for materials should be addressed to J.-S.K. or S.-H.C.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023