



# Improvement of Bone Age Assessment Using a Deep Learning Model in Young Children: Significance of Carpal Bone Analysis

Sang-Un Kim <sup>1</sup>, Saelin Oh <sup>1,\*</sup>, Kee-Hyoung Lee<sup>2</sup>, Chang Ho Kang<sup>1</sup> and Kyung-Sik Ahn<sup>1</sup>

<sup>1</sup>Department of Radiology, Anam Hospital, Korea University, Seoul, Republic of Korea

<sup>2</sup>Department of Pediatrics, Anam Hospital, Korea University, Seoul, Republic of Korea

\*Corresponding author: Department of Radiology, Anam Hospital, Korea University, 73, Goryeodae-ro, Seongbuk-gu, Seoul, Republic of Korea. Tel: +82-29205588, Email: ninolin@naver.com

Received 2023 March 19; Revised 2023 June 17; Accepted 2023 June 19.

## Abstract

**Background:** Deep learning methods used for bone age assessment (BAA) mostly employ the whole hand or regional convolutional neural networks without carpal bones; therefore, their application is insufficient in young children.

**Objectives:** This study aimed to improve the accuracy of BAA in young children by integrating a carpal bone analysis and to achieve a similar BAA accuracy for all age groups.

**Patients and Methods:** A hybrid Greulich-Pyle (GP) and modified Tanner-Whitehouse deep learning model for BAA was trained by integrating an additional carpal bone analysis of an open dataset. A total of 453 hand radiographs from a single institution were selected for external validation. To create the reference standard, three human experts conducted a BAA, based on the GP Atlas, and then, interobserver agreement was evaluated. The model performance was estimated by comparing the mean absolute difference (MAD) and the root mean square error (RMSE) between the two BAA models, including one with a carpal bone analysis (M1) and one without a carpal bone analysis (M2), and the reference standard. The MAD of each model was compared between sex and age groups with respect to four major developmental stages, that is, pre-puberty, early and mid-puberty, late puberty, and post-puberty.

**Results:** The M1 model showed a higher accuracy with a lower MAD (0.366; 95% confidence interval [CI]: 0.337 - 0.395) compared to the M2 model (0.388; 95% CI: 0.358 - 0.418) for all age groups, with a significant difference ( $P < 0.001$ ). The RMSE values versus the reference standard were 0.483 and 0.505 years for the M1 and M2 models, respectively. According to sex and developmental stage distributions, the M1 model had a greater predictive ability compared to the M2 model for pre-pubertal patients, regardless of sex ( $P = 0.008$  for males and  $P = 0.022$  for females).

**Conclusion:** Based on the present findings, the integration of a carpal bone analysis into the BAA model improved its accuracy, especially in young children.

**Keywords:** Carpal Bones, Pediatrics, Comparative Study, Deep Learning

## 1. Background

Bone age assessment (BAA) is defined as a clinical procedure for evaluating the skeletal maturity of pediatric patients (1). Discrepancies between bone and chronological age indicate abnormal skeletal development. Hand-wrist radiographs, including images obtained using the Greulich-Pyle (GP) method (2) and Tanner-Whitehouse 3 (TW3) approach (3), are commonly used to determine skeletal maturity. Generally, GP is an atlas-based method, which compares a patient's radiograph with an atlas of representative age, whereas TW3 is a scoring system based on the osseous stages and events at each level.

The deep learning technology based on artificial neural networks has rapidly evolved in recent years, particularly in the field of medical imaging (4, 5). Convolutional neural networks (CNNs) allow for a successful classification of thoracic abnormalities (6), neuroanatomic structure segmentation (7), detection and classification of breast cancer (8), and performance of BAA (9-12). A BAA is an ideal method for an automated image evaluation, as a single study contains few images, and the results are relatively standardized. Therefore, various automatic and deep learning methods have been proposed for BAA.

Most of the proposed methods used for BAA employ a deep learning approach on the whole hand (9-11, 13, 14)

or an extracted region of interest (ROI) (12). Although these methods enhance the clinical efficacy and accuracy of BAA, their application in young children is insufficient, as some previous studies excluded patients under the age of 4 years (9, 11), and others lacked a carpal bone analysis (12), which is important for skeletal maturity assessment in infants and toddlers (3). Generally, the carpal bones play an important role in determining the bone age of infants and toddlers during pre-puberty before these structures begin to overlap (3). Besides, since the survival rates of infants in all preterm and birth weight categories have improved (15), a BAA is needed in young children.

## 2. Objectives

In this study, we proposed augmenting BAA by integrating a carpal bone analysis to improve its accuracy in young children and to achieve a similar BAA accuracy for all age groups.

## 3. Patients and Methods

This study was approved by the institutional review board and the ethics committee of our institution (IRB No. 2020AN0475). The need for informed consent was waived, because data were collected retrospectively and analyzed anonymously.

### 3.1. Datasets

Two public datasets were used for training in this study. The first dataset was obtained from the Radiological Society of North America (RSNA) pediatric bone age 2017 challenge and included 14,236 hand radiographs from Stanford University and Colorado University (16). The second dataset was retrieved from the digital hand atlas (17), which includes 1,375 hand radiographs from the University of Southern California. We used 10% of the datasets as the validation set and 200 images as the test set.

For external validation, a total of 453 participants were selected through stratified random sampling among patients who had visited the pediatric department and underwent left-hand radiography in our institution. Individuals with congenital anomalies (i.e., Down's syndrome, Noonan syndrome, congenital adrenal hyperplasia, and methylmalonic acidemia) were excluded, along with individuals with poor-quality images. The reference standard for the external validation set was the average independent bone age estimated by three reviewers (reviewer 1, a pediatric endocrinologist with 29 years of clinical experience, and reviewers 2 and 3, musculoskeletal radiologists with 19 and 12 years of

clinical experience, respectively). The three reviewers independently estimated the bone age based on the GP Atlas, with the first digit after the decimal point (y). The reviewers tried to appraise the bone age according to the GP standard if possible; nonetheless, they were allowed to use the calculated median age based on their level of experience. If there was a discrepancy of more than two years, the image was re-evaluated until consensus was reached.

### 3.2. Model Development

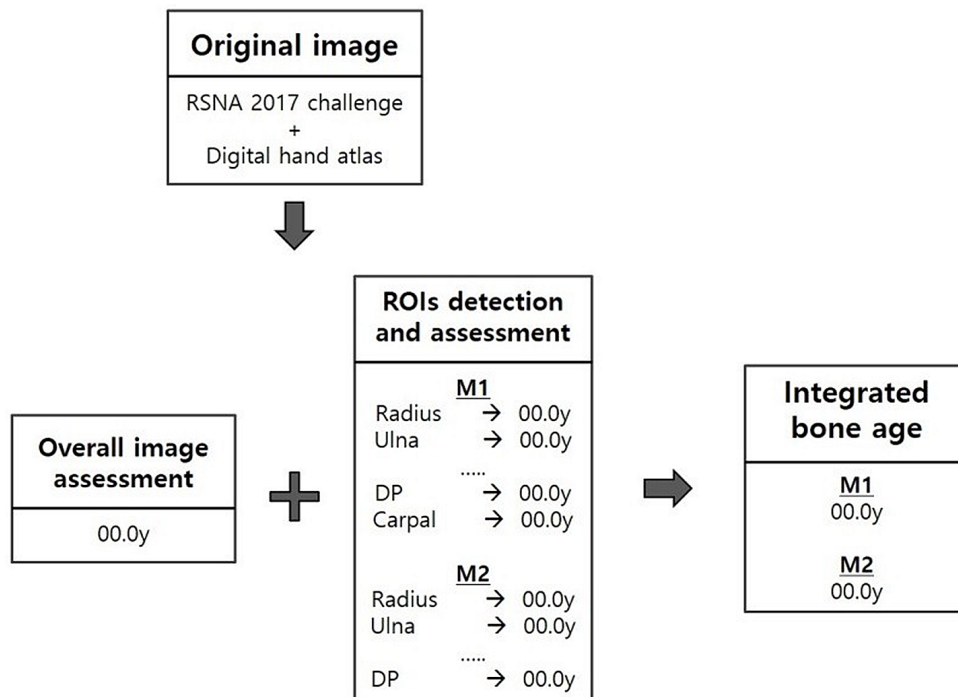
The proposed BAA model relied on hybrid TW3 and GP artificial intelligence (AI)-based automatic bone age measurements (18). The proposed BAA method is illustrated in Figure 1. First, multiple regions based on TW3 were automatically detected using the CNN algorithm. Next, each ROI and the holistic hand image were automatically classified based on the maturity level, using the CNN algorithm. To achieve greater accuracy and detailed evaluation than the nine stages of TW3, in our model, we applied 34 stages with six-month intervals of 1.5 to 18 years in age for maturity. For the maturity classification of each CNN model training, we used the physicians' ratings as the reference standard. Finally, each ROI and all holistic hand features were integrated and classified to estimate the final bone age for the input image.

To compare the accuracy of BAA, the model was applied to two separate models with different radiographic areas. Our proposed model (M1) used eight regions (radius, ulna, distal phalanges, middle phalanges, proximal phalanges, metacarpal of the third digit, metacarpal of the first digit, and carpal region), whereas the former model (M2) applied seven regions (radius, ulna, distal phalanges, middle phalanges, proximal phalanges, metacarpal of the third digit, and metacarpal of the first digit) without a carpal region (Figure 2).

The entire BAA procedure was fully automated. The model was implemented using an open-source machine-learning library (TensorFlow Version 0.9.0; Google, Mountain View, CA, USA).

### 3.3. Statistical Analysis

To validate the reference standard, the intraclass correlation coefficient (ICC) was measured for the assessment of interobserver agreement among the three reviewers by considering the following values and levels of agreement: 0 - 0.20, poor; 0.21 - 0.40, fair; 0.41 - 0.60, moderate; 0.61 - 0.80, substantial; and 0.80 - 1, almost perfect (19). The results of the two BAA models and the findings of the reviewers are summarized as



**Figure 1.** An overview of the hybrid model of Greulich-Pyle (GP) and modified Tanner-Whitehouse 3 (TW3) methods for bone age assessment with two different regions of interest (ROIs)

mean, standard deviation (SD), median, minimum, and maximum values. To measure the predictive accuracy of the models, the root mean square error (RMSE) and mean absolute difference (MAD) of each model were measured as model performance metrics; the latter was calculated based on a 95% confidence interval (CI). A BAA was considered equivalent to the reference standard when the upper limit of the 95% CI of MAD was  $< 0.5$  years. Overall, the smaller the MAD is, the closer the predictions are to the ground truth.

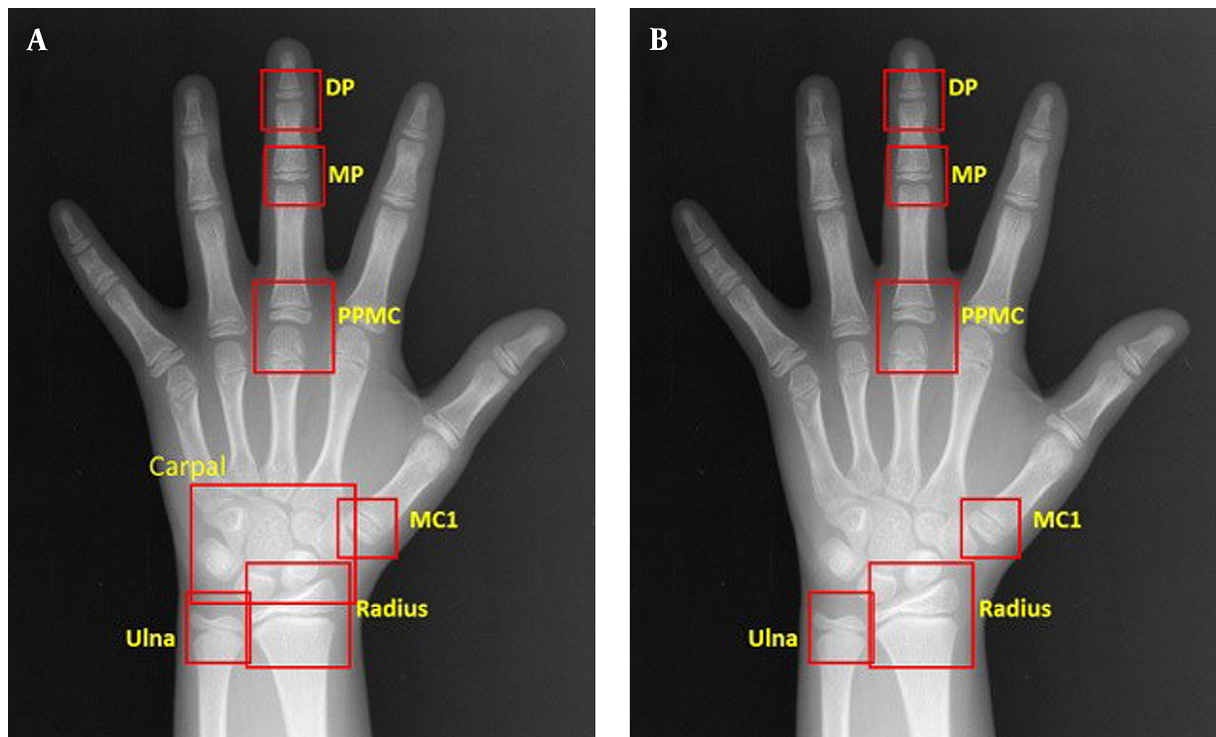
The absolute differences between the two models were compared via paired *t*-test, and Bland-Altman plots (20) were used to correlate the predictions to the ground truth. Additionally, the MAD of each model was compared between sex and age groups with respect to four major developmental stages of the skeletal system: Pre-puberty (males  $\leq 9$  years; females  $\leq 7$  years), early and mid-puberty (9 years  $<$  males  $\leq 14$  years; 7 years  $<$  females  $\leq 13$  years), late puberty (14 years  $<$  males  $\leq 16$  years; 13 years  $<$  females  $\leq 15$  years), and post-puberty (males  $> 16$  years; females  $> 15$  years) (9). P-values less than 0.05 were considered statistically significant. All other analyses were conducted in SPSS Version 20.0 for Windows (IBM SPSS Statistics for Windows Version 20.0, released in 2011; IBM Corp., Armonk,

NY, USA) and SAS Version 9.4 (SAS Institute Inc., Cary, NC, USA).

#### 4. Results

Figure 3 presents the age distribution of the participants. According to the developmental stage distribution, the largest group included 339 participants in early and mid-puberty (74.8%; 74 males and 265 females), while 53 patients (11.7%; 42 males and 11 females) in pre-puberty were included in this study. The ICC (95% CI) of reviewers 1-3 was 0.993 (95% CI: 0.990 - 0.995), which is sufficiently high, allowing for the use of the average BAA value as a reference standard.

A performance summary of the two models on the test dataset is presented in Table 1. The mean bone age  $\pm$  SD of the M1 model, M2 model, and reference standard was  $11.12 \pm 2.60$ ,  $11.14 \pm 2.59$ , and  $11.19 \pm 2.63$  years, respectively. According to the Bland-Altman plots, more than 95% of the predictions across both models were within one year of the ground-truth age (Figure 4). In the M1 model, there were 19 predictions with an absolute difference of more than one year (range, 1.01 - 1.99 years), whereas in the M2 model, there were 22 predictions (range, 1.01 - 1.95 years). The



**Figure 2.** Comparison of the regions of interest (ROIs) of two bone age assessment (BAA) models. A, The model with a carpal bone analysis, including DP, MP, PPMC, MC1, radius, ulna and carpal bones; B, The model without a carpal bone analysis, including DP, MP, PPMC, MC1, radius, and ulna. DP, distal phalanges; MP, middle phalanges; PPMC, proximal phalanges/metacarpal; MC, metacarpal

cumulative percentages of absolute differences within six months between the reference standard and the models were 72.4% and 70.6% for M1 and M2 model, respectively.

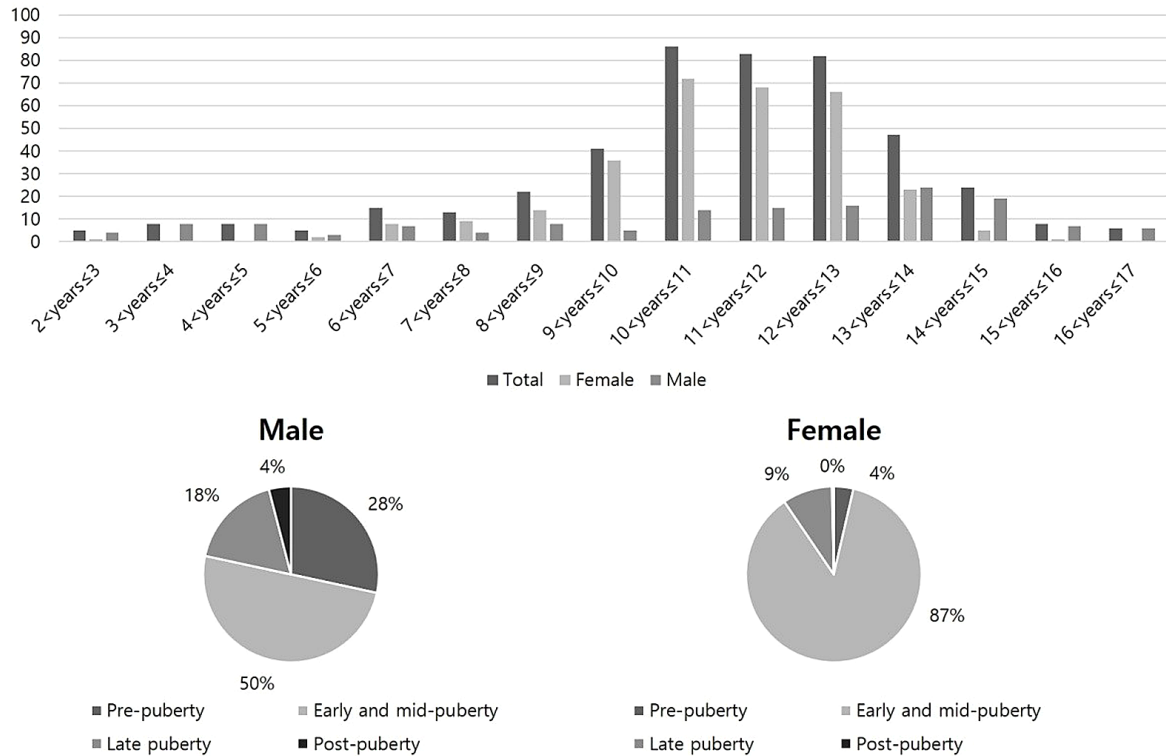
The MAD between the models and the reference standard was 0.366 year for the M1 model (95% CI: 0.337 - 0.395) and 0.388 year for the M2 model (95% CI: 0.358 - 0.418), both of which were below 0.5 year (Table 2). The RMSE value was 0.483 for the M1 model and 0.505 for the M2 model. The M1 model achieved a higher accuracy with a lower MAD compared to the M2 model for all age groups ( $P < 0.001$ ). The M1 model yielded better predictions than the M2 model in all age groups, regardless of sex ( $P = 0.002$  for males and  $P = 0.003$  for females).

According to the developmental stage and sex distribution, the M1 model showed a lower MAD compared to the M2 model in pre-pubertal males, pre-pubertal females, and early and mid-pubertal females ( $P = 0.008$ ,  $P = 0.022$ , and  $P = 0.008$ , respectively). Particularly, the M1 model reduced the 95% CI upper limit of MAD to  $< 0.5$  for pre-pubertal males (MAD = 0.364; 95% CI: 0.362 - 0.465), which could be considered equivalent to the reference standard.

## 5. Discussion

The present study investigated the clinical significance of carpal bone analysis in improving the accuracy of BAAs. Additionally, a hybrid model was established by incorporating a carpal bone analysis using an open dataset. Both models demonstrated human expert-level accuracy, with an upper 95% CI of  $< 0.5$  for the MAD between the BAA models and the reference standard. The MAD values for the M1 and M2 models were 0.366 and 0.388 years, respectively, which are similar to or lower than those reported in previous research (21-23). The current study showed that the model integrating a carpal bone analysis was more accurate than the model without a carpal bone analysis. Particularly, the model using a carpal bone analysis showed significantly improved accuracy in pre-pubertal patients.

The classical method of BAA in children is based on bone growth in the phalanges, carpal bones, and wrist joints. Carpal bone maturity varies widely, and the analysis of this variable does not provide accurate or significant information in patients  $> 7$  years of age (24). In girls aged  $> 6$  years or boys  $> 8$  years, a phalangeal feature analysis yields more reliable information than a carpal



**Figure 3.** Age distribution of the external validation set with respect to four major developmental stages of the skeletal system: Pre-puberty (males ≤ 9 years; females ≤ 7 years), early and mid-puberty (9 years < males ≤ 14 years; 7 years < females ≤ 13 years), late puberty (14 years < males ≤ 16 years; 13 years < females ≤ 15 years), and post-puberty (males > 16 years; females > 15 years)

**Table 1.** The Results of Bone Age Assessment

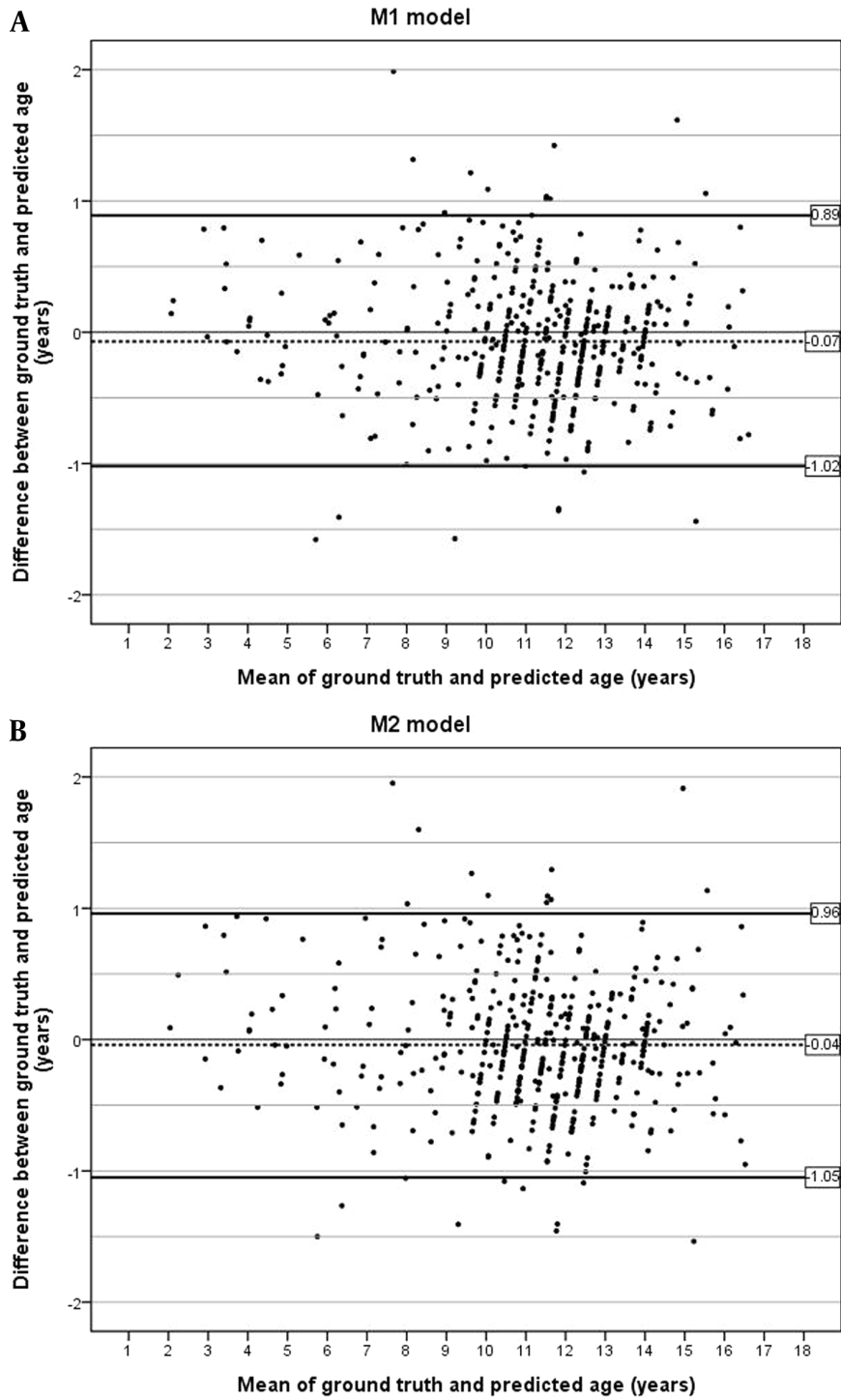
Bone age, years	Total (n = 453)	Males (n = 148)	Females (n = 305)
<b>Reference standard bone age by the three reviewers</b>	11.19 ± 2.63	11.10 ± 3.83	11.23 ± 1.78
Median	11.5	12.0	11.5
Range, min-max	2.0 - 17.0	2.0 - 17.0	3.0 - 16.0
<b>Automatic BAA using the model with a carpal bone analysis (M1 model)</b>	11.12 ± 3.05	11.10 ± 3.79	11.13 ± 1.76
Median	11.41	12.11	11.34
Range, min-max	2.14 - 16.8	2.14 - 16.88	2.97 - 15.22
<b>Automatic BAA using the model without a carpal bone analysis (M2 model)</b>	11.14 ± 2.59	11.13 ± 3.76	11.15 ± 1.77
Median	11.46	12.11	11.31
Range, min-max	2.09 - 16.86	2.09 - 16.86	2.85 - 15.39

Abbreviations: BAA, bone age assessment; SD, standard deviation.

<sup>a</sup> Values are expressed as mean ± SD unless otherwise indicated.

bone analysis; accordingly, most recent deep learning methods do not utilize the carpal bones for analysis (9-14). However, carpal bones are a vital and distinctive feature of skeletal maturity, which provide accurate and significant

information for determining the bone age of young children (3). Combined with the existing phalangeal and wrist joint analyses, a carpal bone analysis significantly improves the accuracy of BAAs in young children.



**Figure 4.** The Bland-Altman plots indicate agreement between each model estimate and the ground truth. A, M1 model with a carpal bone analysis; and B, M2 model without a carpal bone analysis. Black lines represent 1.96 standard deviation of the difference.

**Table 2.** Comparison of Absolute Differences of Automatic Bone Age Assessment with a Carpal Bone Analysis (M1) and Without a Carpal Bone Analysis (M2)

	M1			M2			P-value
	MAD	95% CI	RMSE	MAD	95% CI	RMSE	
<b>Males</b>	0.355	0.305 - 0.406	0.471	0.389	0.334 - 0.444	0.515	0.002
Pre-puberty	0.364	0.362 - 0.465	0.485	0.441	0.326 - 0.557	0.573	0.008
Early and mid-puberty	0.319	0.246 - 0.392	0.446	0.338	0.216 - 0.415	0.473	0.064
Late puberty	0.432	0.325 - 0.540	0.505	0.431	0.316 - 0.546	0.513	0.938
Post-puberty	0.415	0.072 - 0.759	0.512	0.459	0.069 - 0.848	0.570	0.333
<b>Females</b>	0.371	0.335 - 0.407	0.488	0.388	0.352 - 0.423	0.500	0.003
Pre-puberty	0.540	0.101 - 0.979	0.824	0.618	0.211 - 1.024	0.845	0.022
Early and mid-puberty	0.370	0.335 - 0.406	0.473	0.386	0.350 - 0.421	0.484	0.008
Late puberty	0.270	0.174 - 0.366	0.364	0.275	0.177 - 0.373	0.370	0.839
<b>Total</b>	0.366	0.337 - 0.395	0.483	0.388	0.358 - 0.418	0.505	< 0.001

Abbreviations: MAD, mean absolute difference; CI, confidence interval; RMSE, root mean square error.

Several studies have used carpal bone extraction to calculate the bone age (25-27). Most of these studies focused on classical image segmentation rather than bone assessments, without applying a deep learning algorithm. Meanwhile, the accuracy of BAAs was unacceptable. In this regard, Hao et al. (28) exploited carpal bones using a regression CNN to evaluate the bone age. Although they achieved an accuracy of 90.15% within six months from the ground truth for males, their study was only performed on individuals with a chronological age of 0 - 6 years, and a limited BAA was performed for children of all ages.

Additionally, Iglovikov et al. (29) trained several deep network architectures using different parts of images (whole hand, carpal bones, and metal carpals/proximal phalanges) to evaluate the contribution of various bones to the model performance. Moreover, the linear ensemble of these regional models outperformed all the aforementioned models, which is consistent with our results. However, our study reported a higher accuracy and a lower MAD compared to the study conducted by Iglovikov et al. (29) with the lowest MAD at 6.1 months.

Recently, the survival rates of infants in all preterm and birth weight categories have improved owing to advances in obstetric and perinatal care (15). As the survival rates of preterm and low-birth-weight infants increase, concerns regarding the developmental outcomes and growth of survivors also increase. Most infants exhibit an accelerated compensatory growth pattern (known as “catch-up growth”), which is usually completed by two years of age (30). Nevertheless, in the absence of compensatory growth, infants are unlikely to grow to the extent of their peers and reach their target height in adulthood. Consequently, the need for recombinant

human growth hormone therapy has increased in these young children (31), and the need for BAA for these young children has also increased. Our model incorporating a carpal bone analysis yielded better predictions compared to the model without a similar analysis in prepubertal men, suggesting its improved accuracy in young children.

Our hybrid model overcame the limitations of GP and TW3 methods by focusing on regions that are highly related to changes in bone maturity and by applying finer-grained maturity stages compared to TW3, yielding a reliable and accurate bone age estimate (18). Similar to our approach, the BoneXpert (Visiana Aps, Holte, Denmark; <http://www.boneexpert.com>), an automated commercially available BAA system (recently released, version 3.0), is used to evaluate the bone age using both GP and alternative TW2 methods (32). However, unlike our model, the BoneXpert is based on a feature extraction technique, which reconstructs the boundaries of 15 bones (i.e., metacarpals, phalanges, distal radius, and ulna) (33). Overall, the present results showed that the accuracy of bone age predictions improved remarkably for males versus females, which is consistent with the results of a previous carpal bone analysis (25, 29), probably due to the fact that girls mature sooner than boys. This finding is evident in the dataset applied in the experiment, where the carpal bones of girls ossified at a much earlier age than those of boys.

The present study had several limitations. First, although previous studies reported ethnic differences in the growth patterns of certain age groups (34), our study was limited to patients of a single ethnicity. Therefore, a prospective multicenter study with a large sample size is required. Second, our model required radiographs of

diagnostic quality. The model itself did not evaluate the radiographic quality, such as the rotation or incomplete filming of the left hand. Third, our model could not detect traumatic or congenital deformities. Finally, the reference standard for the external validation set was based on the GP Atlas, which is inherently limited to young children. There are also differences in the nutritional and ethnic characteristics of children today compared to those during the 1930's and 1940's, which were used to generate standards.

In conclusion, the present results showed the improved accuracy of the hybrid GP and modified TW model for BAA by integrating a carpal bone analysis. Particularly, the model utilizing a carpal bone analysis showed a significantly improved predictive ability in pre-pubertal patients; therefore, it can be sufficiently used in young children.

## Footnotes

**Authors' Contributions:** Study concept and design: S. O. and C. K.; acquisition of data: K. L., K. A., and S. O.; analysis and interpretation of data: S. K. and S. O.; drafting of the manuscript: S. K.; critical revision of the manuscript for important intellectual content: K. A. and S. O.; statistical analysis: S. O.; administrative, technical, and material support: S. K., K. A., and S. O.; and study supervision: S. O.

**Conflict of Interests:** C. K. is a stockholder of Crescom Inc. (Seongnam, Korea), a startup company, the products and services of which are related to the topic of this article. C. K. only contributed to the study conceptualization. This study was technically supported by Crescom Inc. Other authors declare no conflicts of interest or organizational support for this study.

**Data Reproducibility:** The data presented in this study are available on request from the corresponding author by this journal representative at any time during submission or after publication.

**Ethical Approval:** This study was approved by the institutional review board and ethics committee of our institution (IRB No. 2020AN0475), which waived the need for obtaining informed consent.

**Funding/Support:** There was no funding support for this study.

## References

- Martin DD, Wit JM, Hochberg Z, Sävendahl L, van Rijn RR, Fricke O, et al. The Use of Bone Age in Clinical Practice – Part 1. *Horm Res Paediatr*. 2011;76(1):1–9. <https://doi.org/10.1159/000329372>.
- Greulich WW, Pyle SI. *Radiographic Atlas of Skeletal Development of the Hand and Wrist*. Redwood City, CA: Stanford University Press; 1959.
- Malina RM, Beunen GP. Assessment of skeletal maturity and prediction of adult height (TW3 method). *Am J Hum Biol*. 2002;14(6):788–9. <https://doi.org/10.1002/ajhb.10098>.
- LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature*. 2015;521(7553):436–44. [PubMed ID: 26017442]. <https://doi.org/10.1038/nature14539>.
- Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, et al. A survey on deep learning in medical image analysis. *Med Image Anal*. 2017;42:60–88. [PubMed ID: 28778026]. <https://doi.org/10.1016/j.media.2017.07.005>.
- Hwang EJ, Park CM. Clinical Implementation of Deep Learning in Thoracic Radiology: Potential Applications and Challenges. *Korean J Radiol*. 2020;21(5):51–25. [PubMed ID: 32323497]. [PubMed Central ID: PMC7183830]. <https://doi.org/10.3348/kjr.2019.0821>.
- Kushibar K, Valverde S, Gonzalez-Villa S, Bernal J, Cabezas M, Oliver A, et al. Automated sub-cortical brain structure segmentation combining spatial and deep convolutional features. *Med Image Anal*. 2018;48:177–86. [PubMed ID: 29935442]. <https://doi.org/10.1016/j.media.2018.06.006>.
- McKinney SM, Sieniek M, Godbole V, Godwin J, Antropova N, Ashrafian H, et al. International evaluation of an AI system for breast cancer screening. *Nature*. 2020;577(7788):89–94. [PubMed ID: 31894144]. <https://doi.org/10.1038/s41586-019-1799-6>.
- Lee H, Tajmir S, Lee J, Zissen M, Yeshiwas BA, Alkasab TK, et al. Fully Automated Deep Learning System for Bone Age Assessment. *J Digit Imaging*. 2017;30(4):427–41. [PubMed ID: 28275919]. [PubMed Central ID: PMC5537090]. <https://doi.org/10.1007/s10278-017-9955-8>.
- 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*. 2017;36:41–51. [PubMed ID: 27816861]. <https://doi.org/10.1016/j.media.2016.10.010>.
- Tajmir SH, Lee H, Shailam R, Gale HI, Nguyen JC, Westra SJ, et al. Artificial intelligence-assisted interpretation of bone age radiographs improves accuracy and decreases variability. *Skeletal Radiol*. 2019;48(2):275–83. [PubMed ID: 30069585]. <https://doi.org/10.1007/s00256-018-3033-2>.
- Shin NY, Lee BD, Kang JH, Kim HR, Oh DH, Lee BI, et al. Evaluation of the clinical efficacy of a TW3-based fully automated bone age assessment system using deep neural networks. *Imaging Sci Dent*. 2020;50(3):237–43. [PubMed ID: 33005581]. [PubMed Central ID: PMC7506088]. <https://doi.org/10.5624/isd.2020.50.3.237>.
- Ren X, Li T, Yang X, Wang S, Ahmad S, Xiang L, et al. Regression Convolutional Neural Network for Automated Pediatric Bone Age Assessment From Hand Radiograph. *IEEE J Biomed Health Inform*. 2019;23(5):2030–8. [PubMed ID: 30346295]. <https://doi.org/10.1109/JBHI.2018.2876916>.
- Mutasa S, Chang PD, Ruzal-Shapiro C, Ayyala R. MABAL: a Novel Deep-Learning Architecture for Machine-Assisted Bone Age Labeling. *J Digit Imaging*. 2018;31(4):513–9. [PubMed ID: 29404850]. [PubMed Central ID: PMC6113150]. <https://doi.org/10.1007/s10278-018-0053-3>.
- Raju TNK, Buist AS, Blaisdell CJ, Moxey-Mims M, Saigal S. Adults born preterm: a review of general health and system-specific outcomes. *Acta Paediatr*. 2017;106(9):1409–37. [PubMed ID: 28419544]. <https://doi.org/10.1111/apa.13880>.
- Halabi SS, Prevedello LM, Kalpathy-Cramer J, Mamonov AB, Bilbily A, Cicero M, et al. The RSNA Pediatric Bone Age Machine Learning Challenge. *Radiology*. 2019;290(2):498–503. [PubMed ID: 30480490]. [PubMed Central ID: PMC6358027]. <https://doi.org/10.1148/radiol.2018180736>.
- University of Southern California. *Image processing and informatics lab: Digital hand atlas*. 2017, [cited 6 March 2020]. Available from: <https://ipilab.usc.edu/research/baaweb>.
- Lee KC, Lee KH, Kang CH, Ahn KS, Chung LY, Lee JJ, et al. Clinical Validation of a Deep Learning-Based Hybrid (Greulich-Pyle and Modified Tanner-Whitehouse) Method for Bone Age Assessment.



- Korean J Radiol.* 2021;**22**(12):2017-25. [PubMed ID: 34668353]. [PubMed Central ID: PMC8628149]. <https://doi.org/10.3348/kjr.2020.1468>.
19. Benchoufi M, Matzner-Lober E, Molinari N, Jannot AS, Soyer P. Interobserver agreement issues in radiology. *Diagn Interv Imaging.* 2020;**101**(10):639-41. [PubMed ID: 32958434]. <https://doi.org/10.1016/j.diii.2020.09.001>.
  20. Yu M, Tham YC, Rim TH, Ting DSW, Wong TY, Cheng CY. Reporting on deep learning algorithms in health care. *Lancet Digit Health.* 2019;**1**(7):e328-9. [PubMed ID: 33323206]. [https://doi.org/10.1016/S2589-7500\(19\)30132-3](https://doi.org/10.1016/S2589-7500(19)30132-3).
  21. Larson DB, Chen MC, Lungren MP, Halabi SS, Stence NV, Langlotz CP. Performance of a Deep-Learning Neural Network Model in Assessing Skeletal Maturity on Pediatric Hand Radiographs. *Radiology.* 2018;**287**(1):313-22. [PubMed ID: 29095675]. <https://doi.org/10.1148/radiol.2017170236>.
  22. Lee BD, Lee MS. Automated Bone Age Assessment Using Artificial Intelligence: The Future of Bone Age Assessment. *Korean J Radiol.* 2021;**22**(5):792-800. [PubMed ID: 33569930]. [PubMed Central ID: PMC8076828]. <https://doi.org/10.3348/kjr.2020.0941>.
  23. Bui TD, Lee JJ, Shin J. Incorporated region detection and classification using deep convolutional networks for bone age assessment. *Artif Intell Med.* 2019;**97**:1-8. [PubMed ID: 31202395]. <https://doi.org/10.1016/j.artmed.2019.04.005>.
  24. Cavallo F, Mohn A, Chiarelli F, Giannini C. Evaluation of Bone Age in Children: A Mini-Review. *Front Pediatr.* 2021;**9**:580314. [PubMed ID: 33777857]. [PubMed Central ID: PMC7994346]. <https://doi.org/10.3389/fped.2021.580314>.
  25. Somkantha K, Theera-Umporn N, Auephanwiriyakul S. Bone age assessment in young children using automatic carpal bone feature extraction and support vector regression. *J Digit Imaging.* 2011;**24**(6):1044-58. [PubMed ID: 21347746]. [PubMed Central ID: PMC3222542]. <https://doi.org/10.1007/s10278-011-9372-3>.
  26. Hsieh CW, Liu TC, Jong TL, Tiu CM. A fuzzy-based growth model with principle component analysis selection for carpal bone-age assessment. *Med Biol Eng Comput.* 2010;**48**(6):579-88. [PubMed ID: 20405228]. <https://doi.org/10.1007/s11517-010-0609-y>.
  27. Giordano D, Spampinato C, Scarciolo G, Leonardi R. An Automatic System for Skeletal Bone Age Measurement by Robust Processing of Carpal and Epiphysal/Metaphysal Bones. *IEEE Trans Instrum Meas.* 2010;**59**(10):2539-53. <https://doi.org/10.1109/tim.2010.2058210>.
  28. Hao PY, Chokuwa S, Xie XH, Wu FL, Wu J, Bai C. Skeletal bone age assessments for young children based on regression convolutional neural networks. *Math Biosci Eng.* 2019;**16**(6):6454-66. [PubMed ID: 31698572]. <https://doi.org/10.3934/mbe.2019323>.
  29. Igloukov VI, Rakhlin A, Kalinin AA, Shvets AA. Paediatric Bone Age Assessment Using Deep Convolutional Neural Networks. In: Stoyanov D, Taylor Z, Carneiro G, Syeda-Mahmood T, Martel A, Maier-Hein L, et al., editors. *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. DLIA ML-CDS 2018 2018. Lecture Notes in Computer Science.* Vol. **11045**. Cham: Springer; 2018. p. 300-8. [https://doi.org/10.1007/978-3-030-00889-5\\_34](https://doi.org/10.1007/978-3-030-00889-5_34).
  30. Clayton PE, Cianfarani S, Czernichow P, Johannsson G, Rapaport R, Rogol A. Management of the child born small for gestational age through to adulthood: a consensus statement of the International Societies of Pediatric Endocrinology and the Growth Hormone Research Society. *J Clin Endocrinol Metab.* 2007;**92**(3):804-10. [PubMed ID: 17200164]. <https://doi.org/10.1210/jc.2006-2017>.
  31. Dura-Trave T, San Martin-Garcia I, Gallinas-Victoriano F, Chueca Guindulain M], Berrade-Zubiri S. [Catch-up growth and associated factors in very low birth weight infants]. *An Pediatr (Engl Ed).* 2020;**93**(5):282-8. Spanish. [PubMed ID: 31983650]. <https://doi.org/10.1016/j.anpedi.2019.06.017>.
  32. Thodberg HH, Kreiborg S, Juul A, Pedersen KD. The BoneXpert method for automated determination of skeletal maturity. *IEEE Trans Med Imaging.* 2009;**28**(1):52-66. [PubMed ID: 19116188]. <https://doi.org/10.1109/TMI.2008.926067>.
  33. van Rijn RR, Thodberg HH. Bone age assessment: automated techniques coming of age? *Acta Radiol.* 2013;**54**(9):1024-9. [PubMed ID: 24179234]. <https://doi.org/10.1258/ar.2012.120443>.
  34. Zhang A, Sayre JW, Vachon L, Liu BJ, Huang HK. Racial differences in growth patterns of children assessed on the basis of bone age. *Radiology.* 2009;**250**(1):228-35. [PubMed ID: 18955510]. [PubMed Central ID: PMC2817832]. <https://doi.org/10.1148/radiol.2493080468>.