

RESEARCH

Open Access



# Deep learning-based automated bone age estimation for Saudi patients on hand radiograph images: a retrospective study

Zuhal Y. Hamd<sup>1</sup>, Amal I. Alorainy<sup>1</sup>, Mohammed A. Alharbi<sup>2</sup>, Anas Hamdoun<sup>2</sup>, Arwa Alkhedeiri<sup>2</sup>, Shaden Alhegail<sup>2</sup>, Nurul Absar<sup>3</sup>, Mayeen Uddin Khandaker<sup>4,5</sup> and Alexander F. I. Osman<sup>6\*</sup>

## Abstract

**Purpose** In pediatric medicine, precise estimation of bone age is essential for skeletal maturity evaluation, growth disorder diagnosis, and therapeutic intervention planning. Conventional techniques for determining bone age depend on radiologists' subjective judgments, which may lead to non-negligible differences in the estimated bone age. This study proposes a deep learning-based model utilizing a fully connected convolutional neural network(CNN) to predict bone age from left-hand radiographs.

**Methods** The data set used in this study, consisting of 473 patients, was retrospectively retrieved from the PACS (Picture Archiving and Communication System) of a single institution. We developed a fully connected CNN consisting of four convolutional blocks, three fully connected layers, and a single neuron as output. The model was trained and validated on 80% of the data using the mean-squared error as a cost function to minimize the difference between the predicted and reference bone age values through the Adam optimization algorithm. Data augmentation was applied to the training and validation sets yielded in doubling the data samples. The performance of the trained model was evaluated on a test data set (20%) using various metrics including, the mean absolute error (MAE), median absolute error (MedAE), root-mean-squared error (RMSE), and mean absolute percentage error (MAPE). The code of the developed model for predicting the bone age in this study is available publicly on GitHub at <https://github.com/afiosman/deep-learning-based-bone-age-estimation>.

**Results** Experimental results demonstrate the sound capabilities of our model in predicting the bone age on the left-hand radiographs as in the majority of the cases, the predicted bone ages and reference bone ages are nearly close to each other with a calculated MAE of 2.3 [1.9, 2.7; 0.95 confidence level] years, MedAE of 2.1 years, RMAE of 3.0 [1.5, 4.5; 0.95 confidence level] years, and MAPE of 0.29 (29%) on the test data set.

**Conclusion** These findings highlight the usability of estimating the bone age from left-hand radiographs, helping radiologists to verify their own results considering the margin of error on the model. The performance of our proposed model could be improved with additional refining and validation.

\*Correspondence:  
Alexander F. I. Osman  
alexanderfadul@yahoo.com

Full list of author information is available at the end of the article



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, 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 you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. 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-nc-nd/4.0/>.

**Keywords** Deep learning, Bone age assessment, Left-hand radiographs, Pediatric medicine, Convolutional neural network

## Introduction

Bone age assessment is a widely utilized diagnostic technique in pediatric imaging and legal medicine to assess the bone age of children. Medical professionals routinely assess bone age concerning chronological age for assessing development, metabolic problems, and genetic conditions in children. The Tanner-Whitehouse (TW3) [1] or Greulich and Pyle (GP) [2] approaches are commonly employed for radiographically assessing the left hand. Approximately 76% of radiologists utilize radiograph images of the left hand to evaluate bone age by comparing them to a reference atlas that necessitates time and hard work. Forensic medicine demands the determination of criminal responsibility through age assessment, refugee age estimation, and the distinction between children and adults. Recently, novel techniques have been developed for conducting experiments using ultrasound imaging, magnetic resonance imaging, and computed tomography (CT) images. Age determination based on bone development is the method that is most frequently applied. However, environmental issues and regional variables, gender, race, endocrine disorders, dietary problems, genetic conditions, inherited syndromes, constitutional growth retardation, and nutritional disorders all influence bone age.

The rapid development of medical technologies has endowed physicians with considerable ease, thereby enhancing the quality and efficacy of healthcare. Medical imaging applications based on computer vision provide radiologists with preliminary data that improves workflow efficiency and diagnostic precision. Deep-learning algorithms, as a tool of artificial intelligence, enable the extraction of more abstract features and enhance the accuracy of predictions based on data. Deep learning methods have been used in medical fields through many approaches, such as the categorization of skin cancer, the classification of malignancy, the prediction of diseases, and the estimation of age using specific algorithms.

Several studies investigated the potential of deep learning techniques for bone age estimation using different architectures of convolutional neural networks (CNNs). These deep learning-based methods included a modified InceptionV3 [3], InceptionV3 [4], U-Net [5], VGG-16 [6, 7], multi-scale data fusion framework and CNN [8], customized CNN [9], CNN based on GP [10], ResNet50 [7], CNN [11–13], CNN based on TW3 [14, 15], and ImageNet [16]. The methods were developed using publicly available open-source data sets such as the Radiology Society of North America dataset [3–7, 13–15] or private data sets [8–10, 12, 16]. Some studies also applied

transfer learning rather than training the model from scratch [3, 12, 16]. The reported results of these methods were promising; however, further improvement is still needed. In this study, we propose a deep learning methodology utilizing a fully connected CNN for the automated determination of skeletal bone age, intending to improve and expedite bone age estimation. It is expected that deep learning models have the potential to transform pediatric radiology and enhance patient outcomes in clinical practice.

## Materials and methods

### Patient data

This a retrospective study approved by the Institutional Review Board and Ethics Committee of King Abdullah bin Abdul Aziz University Hospital and Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia (IRB No. 22–0891). The need for consent from the patient's parents or the patient was waived by the Institutional Review Board (IRB No. 22–0891). Patients with the following indications were included in this study: diagnosis and management of endocrine disorders, evaluation of metabolic growth disorders (tall / short stature), deceleration of maturity in various syndromic disorders, and assessment of treatment response in various developmental disorders. The data set used in this study was retrospectively collected from a single institution. Patients with incomplete data were excluded in this study. The bone age of the included patients was between 2 and 20 years, and the total number of the patients was 473. Each patient had only one image. Left-hand radiographs were retrieved from the PACS (Picture Archiving and Communication System) as DICOM (Digital Imaging and Communications in Medicine). The DICOM files were anonymized/de-identified to ensure confidentiality. The images were acquired in a dorsi-palmar view with digital radiograph machines. The images were two-dimensional (2D) with a 12-bit gray-scale (intensities ranged from 0 to 4095) and various dimensions with a pixel size/resolution of  $0.139 \times 0.139 \text{ mm}^2$ . The DICOM files also include the bone age information and the associated patient's data was removed after labeling. The patient's bone age was clinically determined by a pediatric radiologist or radiologist using the Greulich and Pyle (GP) method [2]. The radiologist has proven skill in skeletal radiology with several years of experience in performing bone age estimation using the GP method. Therefore, we used the bone age of the patients that determined by the radiologist as a reference to compare with our results obtained by the proposed deep learning model. The distribution of the

patient's bone ages in the whole data set is displayed in Fig. 1, with values ranging from 2 to 20 years and an average value of 10.3 years. As shown in the figure; only a few samples lie near the two borders, whereas most samples lie around 10 years.

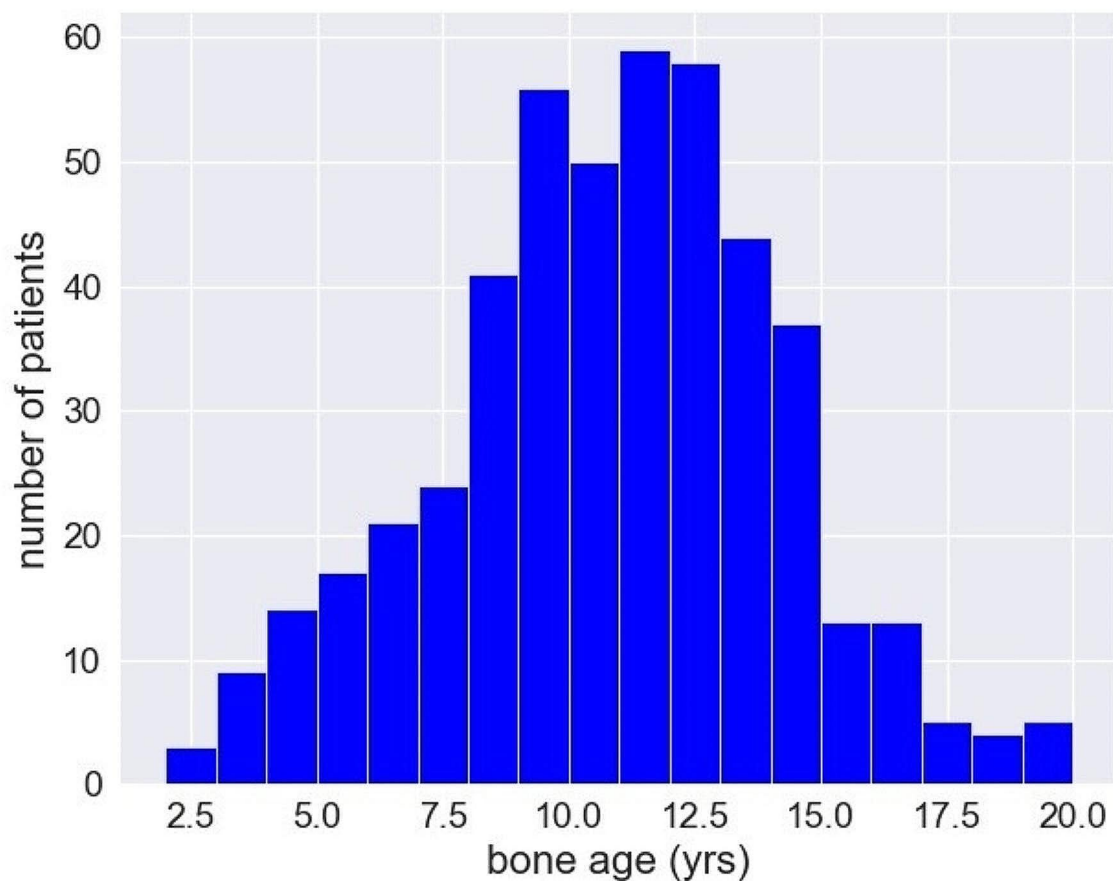
### Preprocessing

Preprocessing the data before training any deep learning model represents a crucial step toward improved predictions. As images in the data set were acquired with different dimensions, we initially resized all images into a dimension of  $512 \times 512$  pixels to eliminate the issue associated with variations in patient's hand size among different ages. We then scaled each image intensity to 256 Gy level instead of 4096 to improve the predictions. Next, we randomly split the whole data set into the model development cohort (80%,  $n=378$ ) and testing cohort (20%,  $n=95$ ). The development cohort is further divided into the training cohort (80%,  $n=302$ ) and the validation cohort (20%,  $n=76$ ). The training data set will be used

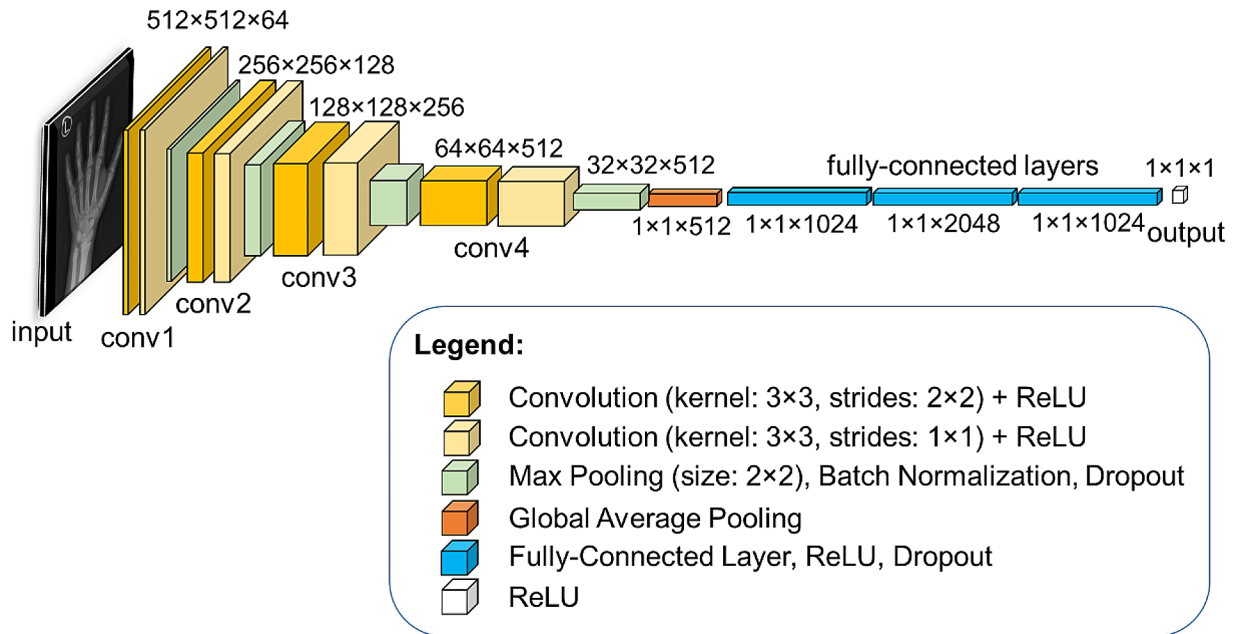
to develop the model and optimize model parameters to update the weights of the model, while the validation data set will be used to keep track of and limit overfitting (e.g., the model weights were only saved in case of improvement of validation loss) and tune hyperparameters to optimize the model. The test data set will be used to assess the model performance. Since our training data set size is relatively small and not sufficient for appropriate training of a deep learning model for improved predictions; finally, we implemented a data augmentation technique [17] by applying image translation (horizontally translated by 10% of the original image size) to double the training data set size ( $n=604$ ). Data augmentation is used to improve the generalizability of the model and avoid overfitting.

### Model architecture

The network architecture of our proposed model to learn silent features on radiograph images to estimate bone age is demonstrated in Fig. 2. It consists of four convolutional



**Fig. 1** The distributions of the bone ages in the entire data set used in this study. The minimum, average, and maximum age were 2,  $10.3 \pm 3.3$ , and 20 years, respectively



**Fig. 2** Network architecture of our proposed bone age prediction model. Boxes indicate a set of feature maps (64–1024) obtained over different operations

blocks to derive features from the input images, a global average pooling layer, and three fully connected layers. A convolutional block consists of two convolutional layers with each followed by a rectified linear unit (ReLU) activation function [18] to pass only the positive outputs. The first convolutional layer in the block has a kernel size of  $3 \times 3$  with strides of  $2 \times 2$ . The second one has the same kernel size with strides of  $1 \times 1$  instead, followed by a max-pooling operation to reduce the spatial size of the feature maps to half of its original size, a batch normalization layer, and a dropout layer. The batch normalization [19] is used to speed up the training process as well as the model convergence. The dropout mechanism [20] used in the blocks is constantly increasing over the blocks starting from 0.05 rate on the first one and reaching 0.20 on the last one. The purpose of using the dropout technique in the convolutional blocks and between the fully connected layers is to maintain the model regularization by preventing overfitting and reducing the mean average error. The number of extracted features in the convolutional blocks increased by a factor of 2 (64 features in the first block to 512 features in the fourth one) encoding various distinct patterns. After the convolution blocks, the global average pooling was utilized to flatten the feature maps into a 1D vector, and the vector donates the high-level feature of the images. The fully connected layers following the global average layer are dense layers, with each followed by a ReLU function and dropout with a rate of 0.25. The number of features in each fully

connected layer is 1024. The output layer is a dense layer with ReLU of a single neuron to propagate to a single regression predicted bone age value.

#### Model implementation

The proposed model was developed and trained in a supervised manner using the left-hand radiograph images as inputs ( $512 \times 512$ ) and the bone age values ( $1 \times 1$ ) determined by a radiologist using the Greulich and Pyle (GP) method [2] as outputs on a training data set ( $n=604$ ). During the training process, the model learns the silent features on the images to predict the bone age of the patient and constantly updates about 7.3 million trainable parameters. The model was trained using the Adam optimization algorithm (which achieves faster convergence than most optimization algorithms) and the mean-squared error (MSE) as a cost function to minimize the difference between the predicted and reference bone age values:

$$Loss(x, y) = \frac{1}{n} \sum_{i=1}^n (y_i - x_i)^2$$

where,  $n$  is the number of images,  $x_i$  is the reference bone age value determined by a radiologist using the Greulich and Pyle (GP) method [2], and  $y_i$  the predicted bone age value obtained by the proposed deep learning model.

The learning rate was initially set to  $1e-4$  and gradually decreased by a factor of 0.10 if there was no improvement after 10 consecutive iterations until reaching  $1e-7$ . A batch size of 8 images was found to be optimal for this task. The weight parameters were initialized with uniform distribution during the training. The model was regularly validated on the validation dataset ( $n=76$ ) to estimate the generalization error in the training and update the hyper-parameters. The model was trained for 120 epochs, which was sufficient to converge. The developed model in this study is fully automated. It takes the left-hand radiograph image and instantly predicts the bone age. The model was developed using Keras API (version 2.10) with a Tensorflow (version 2.10) platform as the backend in Python (version 3.10, Python Software Foundation, Wilmington, DE, USA) on 16 GB RAM CPU with 4 GB GPU support.

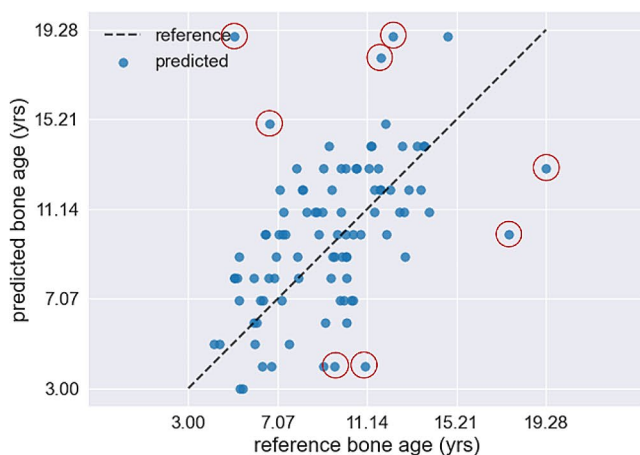
### Evaluation

The trained model is used to make predictions on the test dataset ( $n=95$ ) to assess its performance using various metrics. These metrics include the mean absolute error (MAE), median absolute error (MedAE), root-mean-squared error (RMSE), and mean-squared percentage error (MSPE). The MAE is computed as:

$$MAE(x, y) = \frac{1}{n} \sum_{i=1}^n |y_i - x_i|$$

where,  $n$  is the number of images,  $x_i$  and  $y_i$  are the reference bone age values determined by a radiologist using the Greulich and Pyle (GP) method [2] and the predicted bone age value obtained by our proposed deep learning model, respectively. The RMSE is defined as:

$$RMSE(x, y) = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - x_i)^2}$$

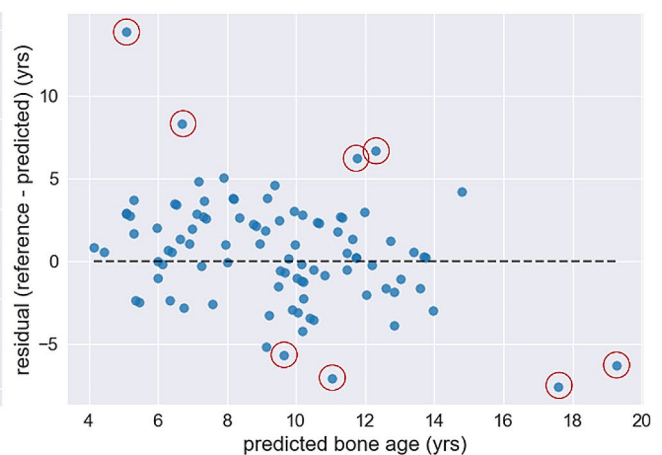


The MAPE can be written in this form:

$$MAPE(x, y) = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - x_i}{y_i} \right|$$

### Results

The results of the predicted bone ages with our proposed model versus the reference bone ages estimated by a radiologist using the Greulich and Pyle (GP) method [2] on the test set are presented in Fig. 3. The solid line depicts the reference bone age, while the dot points represent the predicted values. The figure shows that the majority of the predicted values align along the solid line, indicating a moderate agreement between the reference bone age and the predicted one. There are a few points, eight cases, showing a very high discrepancy larger than 5 years that are not clinically acceptable. One odd point, with the highest difference of 14 years, lies at/near the lower boundary of the bone age in the patient's population (Fig. 1) where there are only a few samples in this patient group and under-represented. Given the fact that the development and testing data sets were randomly divided, these samples may not be represented at all in the development set used to train and validate the model which would certainly fail to perform well on these cases. The rest of the points, seven cases, have a deviation between 5 and 10 years. Almost half of these points, three cases, have a discrepancy between 6 and 8 years and lie at/near the two borders of the population distribution (Fig. 1); therefore, it is expected to observe a high discrepancy in the predicted bone age for these patient groups and the model would fail to perform satisfactorily due to under-representation issue. The other four points, with a discrepancy between 5 and 7 years, lie around the average bone age in the patient population. These moderate deviations, but still high, could be interpreted as



**Fig. 3** Predicted bone age versus reference bone age on the test set

a failure of the model to perform well due to the small size of the entire data set used in this study to develop the model. Training the model using a larger data set size would resolve this issue and lower these discrepancies.

The quantitative performance of the proposed model is presented here on all patients in the test set ( $n=89$ ). The MAE, MedAE, RMSE, and RMSPE metrics were reported for the evaluation. The proposed model achieved a MAE of 2.3 [1.9, 2.7; 0.95 confidence level] years, a MedAE of 2.1 years, a RMSE of 3.0 [1.5, 4.5; 0.95 confidence level] years, and a MAPE of 0.29 (29%).

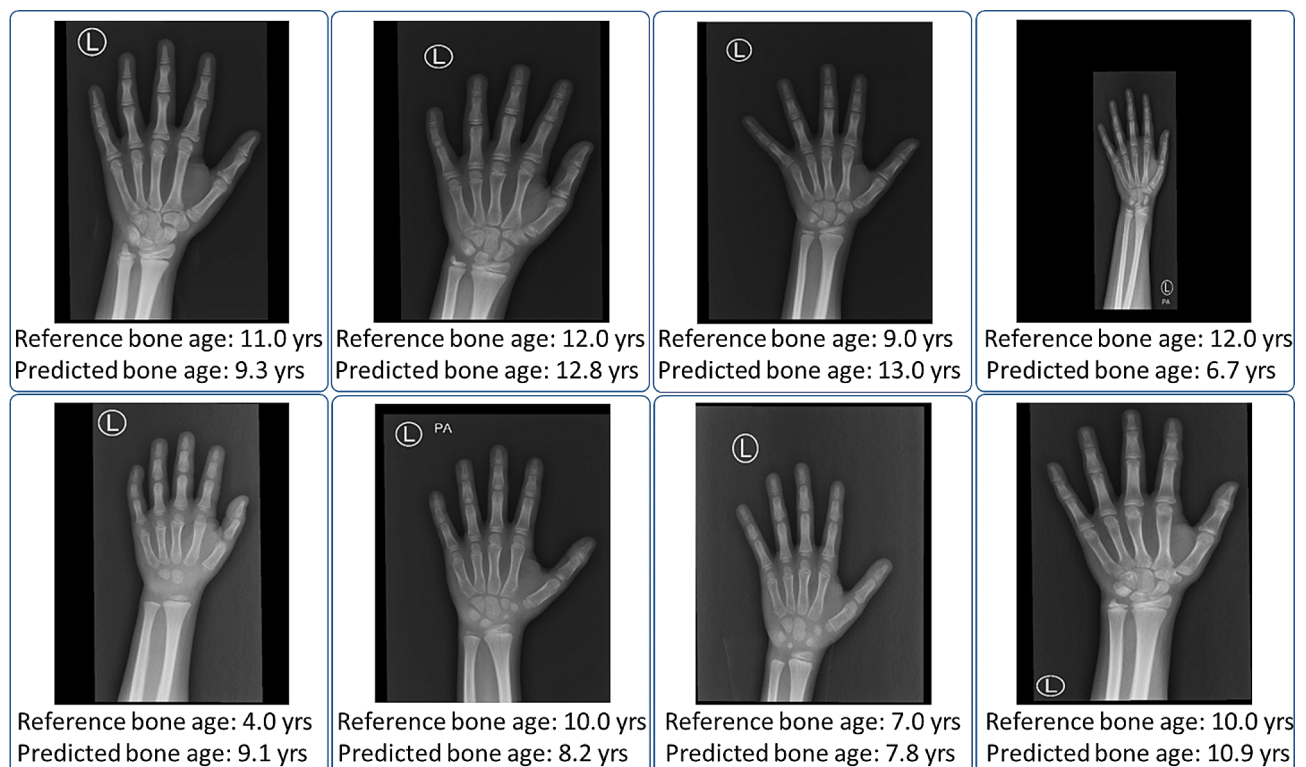
Examples of predicted bone age values compared to reference bone age values are presented in Fig. 4 for patients on the test data set. The results show that predicted bone ages are close to the reference bone ages, except for a few cases where the model was incapable of predicting the bone age accurately with a discrepancy of up to 5.3 years.

The model was trained until no further improvement in its performance has been observed as shown in Fig. 5. The figure indicates progressively minimized error (MES) over iterations/epochs, which reflects improved performance. The training and validation curves have also a comparable performance on both training and validation sets, indicating good generalizability of the model.

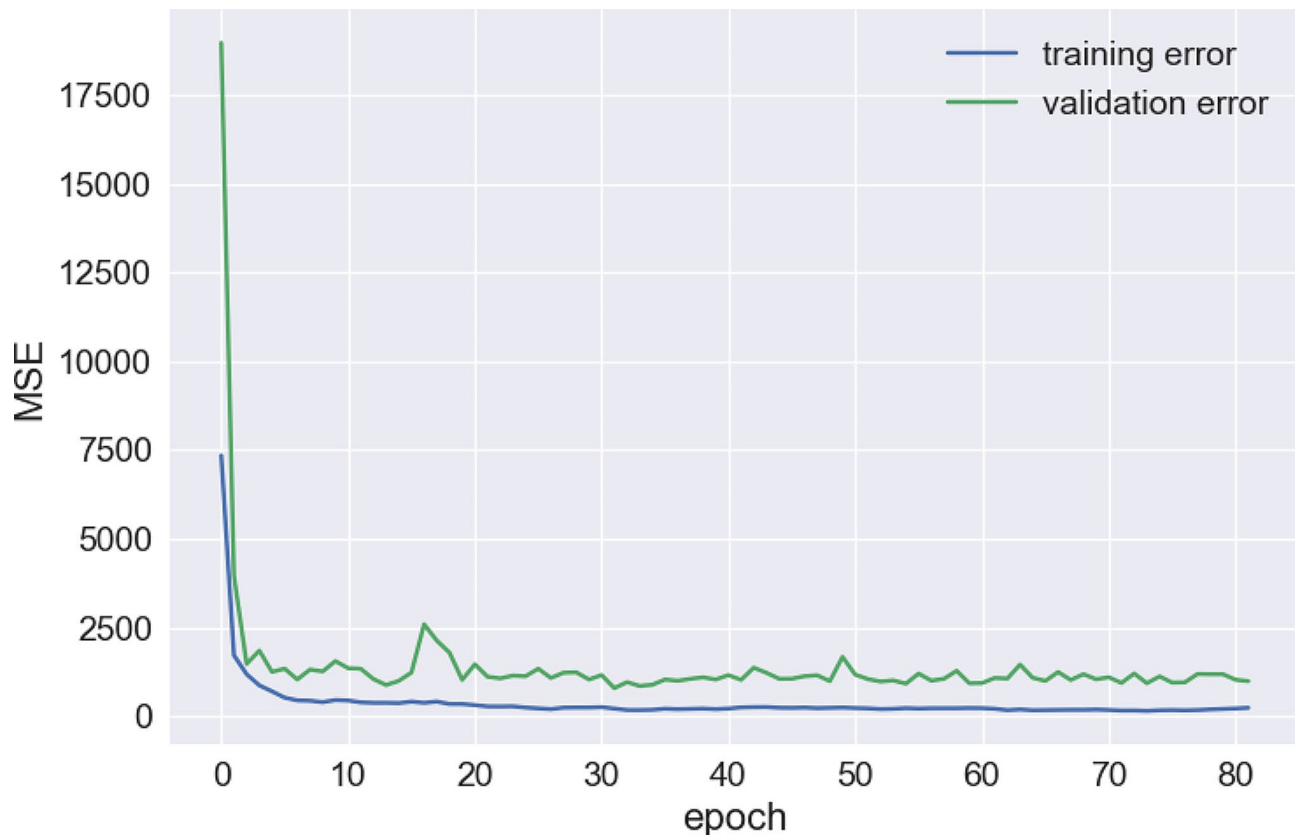
## Discussion

The work offers a novel method for determining bone age from left-hand radiographs using deep learning-based prediction. It is impossible to overestimate the importance of precisely determining bone age in pediatric medicine since it is essential for determining skeletal maturity, identifying growth abnormalities, and developing treatment plans. Bone age is traditionally identified by subjective evaluations made by radiologists, which can lead to inconsistent and variable diagnoses. Nevertheless, by using deep learning techniques to automate the bone age estimation process, the suggested learning-based model provides a possible remedy. CNNs are increasingly being used in medical imaging because of their efficaciousness in processing and analyzing intricate visual input. In this study, a deep learning-based model was developed and validated using a dataset containing 473 patients from a single institution.

The small differences, in the majority of the cases, between predicted and reference bone ages (Fig. 4) show the capability of the proposed model for bone age prediction. The proposed model reduces the subjectivity involved with conventional methods by providing a standardized and objective approach to quantifying bone age through the use of deep learning techniques. By using computerized bone age prediction to speed up the



**Fig. 4** Reference and predicted bone age (in years) of some example patients on the test data set



**Fig. 5** Model performance curve

diagnostic process, clinicians may concentrate on patient care and treatment planning.

The quantitative results obtained in the present study with 95% confidence level (MAE=2.3 years, MedAE=2.1 years, RMSE=3.0 years, and MAPE=0.29) are promising. Compared with some studies in the literature, our results are comparable/inferior to that reported by Ozdemir et al. [3](MAE of 0.36 years, RMSE of 0.48 years), Igloukov et al. [5] (MAE=0.41 years), Castillo et al. [6] (MAE=0.86 years), Liu et al. [8] (RMSE=0.69 years), Metha et al. [4](MAE=0.50 years), Larson et al. [10](RMSE=0.63 years, MAE=0.50 years), Mutasa et al. [9](MAE=0.53 years), Xu et al. [15](MAE=0.64 years), Pan et al. [12](MAE=0.92 years), and Nabilah et al. [7] (MAE=0.41 years). We want to highlight that we used a relatively small single-institution data set ( $n=473$ ) compared to that used in these previous studies. It has been well-known that the accuracy of a CNN model is strongly dependent on the data set size during the training. Therefore, the performance of our model could be further improved by considering more data.

While the study demonstrates the potential of deep learning for bone age estimation, there are specific challenges encountered when dealing with younger children. Some of the challenges are identified as follows. (i) Younger children, especially infants and toddlers,

experience rapid bone development with significant changes occurring within short timeframes. This rapid growth makes it difficult to capture the subtle variations in bone maturity through radiographs, potentially leading to larger errors in age prediction compared to older children. (ii) In younger children, many bones haven't fully ossified, meaning they are still partially composed of cartilage. Cartilage is less visible on radiographs, making it harder for the model to identify and interpret the relevant features for accurate age estimation. (iii) Skeletal abnormalities or diseases affecting bone development can significantly deviate from the typical bone development patterns the model learns from. This can lead to inaccurate age estimations in children with such conditions.

The findings of the proposed study show the capability of our deep learning model in predicting bone age from radiographs of the left hand. However, there are some limitations to this study as well. The main concern of this study is using a small data set (patient demographics, imaging quality, and dataset composition) for training a CNN model from scratch, which is not enough for obtaining accurate predictions and may impact the model's generalizability. However, while a larger dataset would be ideal, in this study, this limitation was addressed in the following ways. (i) We employed data augmentation techniques during training and validation.

This process artificially expands the dataset by generating modified versions of existing images, which helps the model learn from a wider variety of data points and reduces the risk of overfitting to the specific training set. (ii) We used an 80/20 train-test split to ensure the model wasn't simply memorizing the training data. The model's performance was evaluated on a held-out test set (20% of the data) that the model had never seen before. This provides a more realistic assessment of generalizability to unseen data. Besides this, to mitigate the overfitting problem with a smaller dataset, the following strategies were implemented in this study. (i) We employed regularization techniques like dropout layers during training. These techniques help prevent the model from becoming overly reliant on specific features in the training data and promote learning more generalizable patterns. (ii) We implemented early stopping during training. This technique monitors the model's performance on the validation set and stops training once the validation error starts to increase. This helps prevent the model from overfitting to the training data. However, to further improve the model performance and generalizability, future research endeavors may concentrate on verifying the model among other populations by considering more data from open-source datasets and optimizing its structure.

## Conclusions

The study concludes by highlighting the potential of deep learning-based methods to transform the assessment of bone age in pediatric patients. Through the utilization of convolutional neural networks, medical professionals can obtain a dependable and effective instrument for evaluating skeletal maturity and identifying growth abnormalities, leading to better patient outcomes and improved healthcare provision. The study's findings add to the increasing amount of data that supports the use of deep learning methods in medical imaging to enhance the efficiency and accuracy of pediatric diagnosis. The model's applicability and status as a useful clinical decision support tool may be strengthened by additional validation research and clinical applications.

## Abbreviations

2D	Two-dimensional
CNNs	Convolutional neural networks
CT	Computed tomography
DICOM	Digital Imaging and Communications in Medicine
GP	Greulich and Pyle
MAE	Mean absolute error
MAPE	Mean absolute percentage error
MedAE	Median absolute error
PACS	Picture Archiving and Communication System
ReLU	Rectified linear unit
RMSE	Root mean-squared error
TW3Tanner	Whitehouse

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12880-024-01378-2>.

Supplementary Material 1

## Acknowledgements

The authors extend their appreciation to the Deputyship for Research & Innovation, Ministry of Education In Saudi Arabia for funding this research work through the project number RI -44-0055.

## Author contributions

ZH, AIA, MK, and NA contributed to conceptualization, methodology, writing the first draft, editing and reviewing the final draft. MA, AH, AA, and SA contributed to data curation, supervision, and funding acquisition. AO contributed to the methodology, developing the model, analyzing the results, writing the methods and results sections, and revising the manuscript for important intellectual content. All authors contributed to the manuscript revision and approved the submitted version.

## Funding

This work is funded by the Deputyship for Research & Innovation, Ministry of Education In Saudi Arabia for funding this research work through the project number RI -44-0055.

## Data availability

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request. The code of the model for estimating the bone age developed in this study is publicly available in the GitHub repository (<https://github.com/afosman/deep-learning-based-bone-age-estimation>).

## Declarations

### Ethical approval and consent to participate

This study was approved by the Institutional Review Board and Ethics Committee of King Abdullah bin Abdul Aziz University Hospital and Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia (IRB No. 22-0891). The need for consent to participate was waived by an Institutional Review Board (IRB No. 22-0891).

### Consent for publication

Not applicable.

### Competing interests

The authors declare no competing interests.

### Author details

<sup>1</sup>Department of Radiological Sciences, College of Health and Rehabilitation Sciences, Princess Nourah bint Abdulrahman University, P.O. Box 84428, Riyadh 11671, Saudi Arabia

<sup>2</sup>Medical Imaging Department, KAAUH, Riyadh, Saudi Arabia

<sup>3</sup>Department of Computer Science & Engineering, BGC Trust University Bangladesh, Chittagong 4301, Bangladesh

<sup>4</sup>Applied Physics and Radiation Technologies Group, CCDCU, School of Engineering and Technology, Sunway University, Bandar Sunway, Subang jaya 47500, Malaysia

<sup>5</sup>Faculty of Graduate Studies, Daffodil International University, Daffodil Smart City, Birulia, Savar, Dhaka 1216, Bangladesh

<sup>6</sup>Department of Medical Physics, Al-Neelain University, Khartoum 11121, Sudan

Received: 26 March 2024 / Accepted: 24 July 2024

Published online: 01 August 2024



## References

1. Tanner JM, Whitehouse RH, Cameron N, Marshall WA, Healy MJ, Goldstein H. Assessment of skeletal maturity and prediction of adult height (TW2 method). London: Academic; 1983.
2. Greulich WW, Pyle SI. Radiograph atlas of skeletal development of the hand and wrist. 2nd ed. California: Stanford University Press; 1959.
3. Ozdemir C, Gedik MA, Kaya Y. Age estimation from left-hand radiographs with deep learning methods. *Traitement Du Signal*. 2021;38(6):1565–74. <https://doi.org/10.18280/ts.380601>.
4. Mehta C, Ayeesha B, Sotakanal A, Nirmala SR, Desai SD, Suryanarayana KV, Ganguly AD, Shetty V. Deep Learning Framework for Automatic Bone Age Assessment. In *Proceedings of the 2021 43rd Annual Int'l Conf of the IEEE Engineering in Med & Biol Society (EMBC) Virtual 2021*;3093–3096.
5. Iglovikov VI, Rakhlin A, Kalinin AA, Shvets AA et al. Paediatric Bone Age Assessment Using Deep Convolutional Neural Networks. In: *Stoyanov Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. DLMA ML-CDS 2018*. 2018. Lecture Notes in Computer Science, vol 11045. Springer, Cham, 2018. [https://doi.org/10.1007/978-3-030-00889-5\\_34](https://doi.org/10.1007/978-3-030-00889-5_34).
6. Castillo JC, Tong Y, Zhao J, Zhu F. (2018). RSNA bone-age detection using transfer learning and attention mapping. 2018. [http://noiselab.ucsd.edu/ECE228\\_2018/Reports/Report6.pdf](http://noiselab.ucsd.edu/ECE228_2018/Reports/Report6.pdf).
7. Nabilah A, Sigit R, Fariza A, Madyono M. Human bone age estimation of carpal bone X-Ray using residual network with batch normalization classification. *Int J Inf Visualization*. 2023;7(1):105–14.
8. Liu Y, Zhang C, Cheng J, Chen X, Wang ZJ. A multi-scale data fusion framework for bone age assessment with convolutional neural networks. *Comput Biol Med*. 2019;108:161–73. <https://doi.org/10.1016/j.combiomed.2019.03.015>.
9. 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;31:513–9.
10. 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. <https://doi.org/10.1148/radiol.2017170236>.
11. Hao P, Chokuwa S, Xie X, Wu F, 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. <https://doi.org/10.3934/mbe.2019323>.
12. Pan I, Baird GL, Mutasa S, Merck D, Ruzal-Shapiro C, Swenson DW, Ayyala RS. Rethinking Greulich and Pyle: a Deep Learning Approach to Pediatric Bone Age Assessment Using Pediatric Trauma Hand Radiographs. *Radiol Artif Intell*. 2020;2(4):e190198. <https://doi.org/10.1148/ryai.2020190198>.
13. 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. <https://doi.org/10.1016/j.media.2016.10.010>.
14. Bui TD, Lee J-J, Shin J. Incorporated region detection and classification using deep convolutional networks for bone age assessment. *Artif Intell Med*. 2019;97:1–8.
15. Xu X, Xu H, Li Z. Automated Bone Age Assessment: a New Three-Stage Assessment Method from Coarse to Fine. *Healthcare*. 2022;10:2170.
16. Lee H, Tajmir S, Lee J, Zissen M, Yesiwas BA, Alkasab TK, Choy G, Do S. Fully automated deep learning system for bone age assessment. *J Digit Imaging*. 2017;30(4):427–41. <https://doi.org/10.1007/s10278-017-9955-8>.
17. Salamon J, Bello JP. Deep convolutional neural networks and data augmentation for environmental sound classification. 2016; arXiv:1608.04363. Available at <https://arxiv.org/abs/1608.04363>.
18. Nair V, Hinton GE. Rectified linear units improve restricted Boltzmann machines. *Proc. 27th Int Conf Mach Learn*. 2010;807–814.
19. Ioffe S, Szegedy C. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. 2015; arXiv:1502.03167v3. <https://arxiv.org/abs/1502.03167v3>.
20. Srivastava N, Hinton GE, Krizhevsky A, Sutskever I, Salakhutdinov R. Dropout: a simple way to prevent neural networks from overfitting. *J Mach Learn Res*. 2014;15:1929–58.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.