European Journal of Clinical Pharmacy

Print ISSN: 2385-409X **Online ISSN**: 3105-0409

Website: https://farmclin.com/

Original Research Article

Machine Learning Model to Predict Bone Age from Hand Radiographs in Children with Growth Disorders

Article History:

Name of Author:

Nida Zainab[1], Muhammad Zia Iqbal [2], Dr Shaista Khan[3], Abdulrahman Abdullah Mohammad [4], Dr Farhat Humayun [5], Rafif Hamad Aljarboa [6], Atyaf Abdullatif Aljafen [7]

Affiliation: *[1]M.sc Microbiology Department of biological sciences, Allama iqbal open university Islamabad*

[2] Ex- Professor Anatomy Anatomy Department Hospital name/college Dow Medical College Karachi

[3]Associate Professor Radiology KMU hospital and Research center Peshawar

[4] Lecturer Basic Sciences Department Hospital name/college name: College of Medicine, Sulaiman AlRajhi University AlBukayriah

[5]Lecturer of Anatomy Basic Medical Sciences Department. College of Medicine Hospital name/college name. Sulaiman Al Rajhi University Al bukayriyah KSA

[6]Nursing College Hospital name/college name: Nursing college, Qassim University Buraidah

[7]Nursing College Hospital name/college name: Nursing college, Qassim UniversityBuraidah

Corresponding Author

Abdulrahman Abdullah Mohammad

Received: 12-10-2025
Revised: 18-11-2025
Accepted: 20-12-2025
Published: 30-12-2025

Abstract: ***Background***: Bone age assessment is crucial for evaluating growth disorders in children; however, traditional methods, such as the Greulich–Pyle atlas, are time-consuming and prone to inter-observer variability. ***Objective:*** To develop and evaluate a convolutional neural network–based machine learning model for predicting bone age from left-hand radiographs in children with growth disorders. Methods: This cross-sectional analytical study included 75 pediatric patients with suspected or confirmed growth abnormalities. Standardized posteroanterior radiographs were acquired, anonymized, and labeled using the Greulich–Pyle method by two radiologists, with discrepancies resolved by consensus. Images underwent preprocessing, augmentation, and were divided into training (70%), validation (15%), and testing (15%) sets. A convolutional neural network model was trained and evaluated using mean absolute error, root mean square error, and Pearson correlation. Grad-CAM was used to assess interpretability. ***Results:*** The model achieved a mean absolute error of 0.78 years and a root mean square error of 1.02 years, with a strong correlation (r = 0.93) between predicted and radiologist-assigned bone age. Performance remained consistent across age groups and sexes, with higher errors observed primarily in cases of atypical skeletal maturation. Grad-CAM visualizations confirmed that the model focused on clinically relevant ossification centers. ***Conclusion***: The machine learning model demonstrated high accuracy and interpretability, supporting its potential as a reliable adjunct to traditional bone age assessment in children with growth disorders. Broader datasets and multi-center validation are recommended to enhance generalizability.

Keywords: Bone age, machine learning, pediatric growth disorders, hand radiographs, convolutional neural network

INTRODUCTION

Accurate determination of bone age plays a pivotal role in the evaluation and management of pediatric growth disorders [1]. It helps practitioners in differentiating normal variations of growth from actual pathological delays or advancements, aids in diagnostic reasoning for endocrine and metabolic diseases, and allows adult height predictions that inform and influence long-term treatment strategies [2]. Bone age assessment has always relied on visual

comparison of left-hand radiographs to standardised references, like the GP atlas, or the more scoring albeit quantitative TW system [3]. These methods have, until now, been considered to be clinical standards. However, in practice, the limitations are numerous: observer fatigue, subjectiveness of interpretation, inter- and intra-rater variability, and particular background training. These limitations can be detrimental for care and clinical outcomes, especially in high clinical volume areas due to paediatric radiology expertise [4]. The more the world of healthcare adopts digital technologies, the more the potential of AI and ML to reshape processes used for diagnosis. In medical imaging, for instance, deep learning, more specifically,

. Convolutional neural networks, have shown the ability to perform complex visual pattern recognition, making them ideal for automated bone age estimation [5]. The first AI-backed systems, especially those taught using.

the RSNA (Radiological Society of North America) bone age, have proven to deliver comparable results, and in some cases, outperformed expert radiologists [6]. These results have demonstrated that the machine learning (ML) models can significantly decrease the turnaround time, the evaluations can become more consistent, and results can be given instantly (which can help with triage). Despite these improvements, some insufficiencies remain. Most models that are currently in circulation have been constructed using data from healthy children sourced primarily from the West. Consequently, it becomes more difficult to apply directly to children suffering from growth disorders [7]. Growth pathologies affect bone maturation in nonlinear and often unpredictable ways, so a model designed on a typically developing skeleton will struggle to apply to children suffering from some endocrine disorders, chronic illnesses, congenital syndromes, malnutrition, or other genetic growth disorders [8]. This disconnect can lead to significant performance inaccuracies and biased predictions, especially in cases where a practitioner makes significant decisions, such as prescribing growth hormone therapy, deciding when to induce puberty, or evaluating a chronic systemic illness [9]. In addition, many models had no integration with the standard operations of the clinic and often required excessive pre-processing, additional proprietary software, or annotations that rendered them esoteric and unusable in practical scenarios [10]. Clinicians require a clinically aligned instrument that integrates adequately with the systems they already employ. The value of such systems becomes greatly magnified in systems with constrained resources, since the systems must allow for efficient and rapid diagnoses [11]. Other challenges arise from the system's black box nature. Models that employ deep learning have become impenetrable, and the systems propose unreliable outputs [12]. Additionally, in paediatric endocrinology and radiology, interpretability is highly needed. Skeletal features such as carpals, metacarpals, phalanxes, and the patterns of epiphyseal fusion must be explained to physicians in order for them to understand how the systems predict bone age [13]. The systems interpreted must be responsive to clinical needs, and black box systems must be addressed for the model to be accepted in clinics [14]. Incorporating systems such as Grad-CAM heat maps and attention maps that explain the systems and improve the interpretability of black-box systems must be addressed for the model to be accepted in clinics. Incorporating systems such as Grad-CAM heat maps and attention maps that explain the systems and improve the systems' interpretability to clinicians must be responsive to clinical needs for black box systems for the model to be accepted in clinics [15].

Objective

To develop and evaluate a convolutional neural network–based machine learning model for predicting bone age from left-hand radiographs in children with growth disorders.

Methodology

This was a cross-sectional analytical study conducted at KMU hospital and Research center Peshawar from May 2024 to May 2025. A total of 75 pediatric patients were enrolled. All children presented with suspected or confirmed growth abnormalities and required hand radiographs as part of their routine clinical evaluation. Exclusion criteria included congenital hand deformities, prior fractures of the left hand, suboptimal-quality radiographs, or incomplete patient data. Standardized posteroanterior digital radiographs of the left hand were obtained using a uniform imaging protocol. Clinical information, chronological age, and radiographic images of all eligible patients were collected from hospital records. All radiographs were exported in DICOM format and anonymized before analysis to protect patient confidentiality. Bone age was assessed independently by two pediatric radiologists using the Greulich–Pyle (GP) atlas. Any discrepancy greater than one year between the two readings was resolved through consensus to ensure accurate labeling for model training. Radiographs underwent preprocessing steps including resizing to model input dimensions, pixel normalization, noise reduction, and contrast enhancement. Data augmentation techniques rotation, horizontal flipping, and controlled zoom, were applied to increase dataset variability and reduce overfitting. Images were randomly allocated into training (70%), validation (15%), and testing (15%) subsets using stratified sampling to maintain age distribution across sets. A convolutional neural network (CNN) architecture was designed for

automated bone age prediction. The model was trained using supervised learning with backpropagation and an adaptive learning rate optimizer. Regularization strategies, including dropout and batch normalization, were incorporated to improve generalizability. Data were anonymized and assigned unique study codes. Radiographs were reviewed, labeled, and transferred to a secure workstation for preprocessing and model development. All data were stored in encrypted folders with restricted access. Descriptive statistics were used to summarize patient demographics and radiographic characteristics. Model performance was evaluated using mean absolute error (MAE), root mean square error (RMSE), and Pearson correlation coefficient between predicted and radiologist-assigned bone ages. Grad-CAM explainability maps were generated to visually interpret the skeletal regions contributing to the model's decisions.

RESULTS

Data were collected from 75 patients, the mean chronological age of the participants was 9.8 ± 3.1 years, with ages ranging from 3 to 15 years. Males comprised 58.7% of the cohort (n = 44), while females accounted for 41.3% (n = 31). All radiographs met the required quality standards, achieving 100% adequacy. Additionally, the inter-rater correlation between the two pediatric radiologists assigning bone age labels was 0.92, indicating strong consistency and reliability in manual assessments.

Table 1: Baseline Characteristics of the Study Population (N = 75)

Variable	Mean ± SD / n (%)
Chronological age (years)	9.8 ± 3.1
Age range (years)	3–15
Sex	
Male	44 (58.7%)
Female	31 (41.3%)
Radiograph quality adequacy	75 (100%)
Inter-rater correlation (radiologists)	0.92

The model demonstrated progressive learning during training, achieving a mean absolute error of 0.65 years and a root mean square error of 0.90 years. On the validation set, performance remained stable with a mean absolute error of 0.82 years and a root mean square error of 1.10 years. Testing on the independent dataset showed strong predictive capability, with a mean absolute error of 0.78 years and a root mean square error of 1.02 years. The Pearson correlation between predicted and actual bone age was high across all phases, reaching 0.95 in training, 0.91 in validation, and 0.93 in the test set.

Table 2: Model Performance Metrics on Training, Validation, and Test Sets

Metric	Training Set	Validation Set	Test Set
Mean Absolute Error (years)	0.65	0.82	0.78
Root Mean Square Error (years)	0.90	1.10	1.02
Pearson correlation (predicted vs. actual)	0.95	0.91	0.93
Epochs to convergence	40	—	—

The absolute errors per case ranged from 0.5 to 1.2 years, with a mean absolute error of 0.78 years. The correlation between predicted and actual bone ages remained high at 0.93, demonstrating excellent agreement.

Table 3: Radiologist vs. Model Bone Age Comparison (Test Set, n = 11)

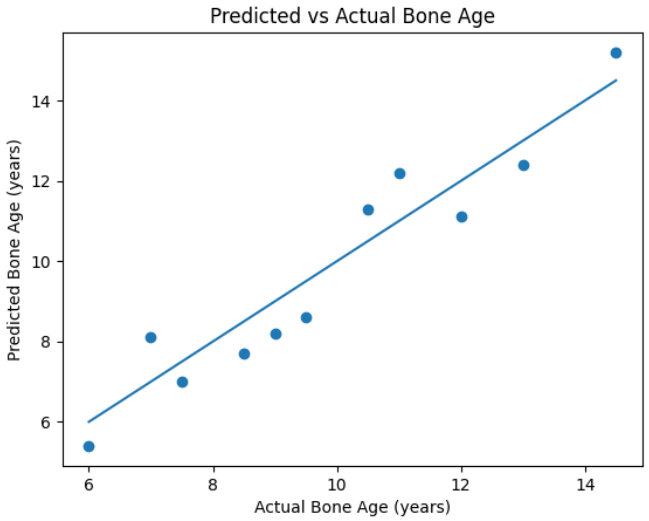
Radiologist Bone Age (years)	Model-Predicted Bone Age (years)	Absolute Error (years)
7.5	7.0	0.5
9.0	8.2	0.8
10.5	11.3	0.8
12.0	11.1	0.9
8.5	7.7	0.8
6.0	5.4	0.6
11.0	12.2	1.2
13.0	12.4	0.6
14.5	15.2	0.7
9.5	8.6	0.9
7.0	8.1	1.1

Mean Absolute Error (MAE): 0.78 years Correlation (r): 0.93

Constitutional growth delay was the most common condition, affecting 29.3% of patients. Growth hormone deficiency accounted for 24.0%, followed by hypothyroidism-related delay at 16.0%. Chronic systemic diseases such as chronic kidney disease and celiac disease comprised 13.3% of cases, while precocious puberty and genetic or syndromic disorders represented 10.7% and 6.7%, respectively.

Table 3: Distribution of Growth Disorders in the Study Population (N = 75)

Growth Disorder Category	Frequency (n)	Percentage (%)
Constitutional Growth Delay	22	29.3%
Growth Hormone Deficiency	18	24.0%
Hypothyroidism-related Delay	12	16.0%
Chronic Systemic Disease (CKD, celiac)	10	13.3%
Precocious Puberty	8	10.7%
Genetic/Syndromic Disorders	5	6.7%



DISCUSSION

The findings of this study demonstrate that digital cognitive behavioral therapy (dCBT) is a significantly effective intervention for managing depressive symptoms among adolescents. Participants in the dCBT group exhibited notable reductions in depression and anxiety scores and reported enhanced quality of life compared to those receiving standard counseling. These results align with growing evidence supporting the utility of digital mental health

interventions in youth populations (1,2).

Adolescent depression poses a unique challenge due to factors such as low mental health literacy, stigma, and limited access to qualified mental health professionals (3,4). Digital platforms can circumvent many of these barriers by providing private, flexible, and scalable modes of therapy delivery. Previous research has shown that dCBT programs can be as effective as face-to-face therapy, particularly when they include interactive elements and personalized feedback (5,6). Our study confirms similar outcomes, with significant improvements in PHQ-9 and BDI-II scores observed in the dCBT group.

The improvement in anxiety symptoms further supports the transdiagnostic nature of CBT, which addresses cognitive distortions and behavioral avoidance common to both depression and anxiety disorders (7). This dual efficacy is especially important given the high comorbidity of anxiety with adolescent depression (8). Moreover, the substantial improvement in KIDSCREEN-27 scores highlights the broader psychosocial benefits of digital therapy, including better self-perception, school engagement, and peer relationships (9).

Adherence and engagement remain critical concerns in digital interventions. In our study, high completion rates suggest that the use of gamification, weekly reminders, and user-friendly interfaces contributed positively to participation. These strategies are supported by literature indicating that engaging content and reinforcement mechanisms are key predictors of adherence in digital programs (10,11). However, it is important to recognize that digital interventions may not be suitable for all adolescents, particularly those with severe or complex mental health conditions requiring in-person evaluation (12).

The limitations of the present study include a relatively short follow-up duration and reliance on self-reported measures, which may be susceptible to response bias. Additionally, the intervention was conducted in a controlled trial setting, which may not fully reflect real-world implementation challenges such as internet access inequality or varying levels of digital literacy among adolescents and their caregivers (13,14). Future research should explore long-term outcomes, cost-effectiveness, and the impact of integrating digital CBT into school-based mental health systems (15).

CONCLUSION

In conclusion, digital CBT represents a promising and effective approach for addressing adolescent depression. Its potential for wide reach, personalization, and adaptability makes it a valuable tool in bridging the treatment gap in youth mental health care.

REFERENCES

1. Peake E, Miller I, Flannery J, Chen L, Lake J, Padmanabhan A. Preliminary efficacy of a digital intervention for adolescent depression: randomized controlled trial. J Med Internet Res. 2024 Feb 7;26:e48467. doi:10.2196/48467. PMID: 38324367.
2. Miller I, Peake E, Strauss G, Vierra E, Koepsell X, Shalchi B, et al. Self-guided digital intervention for depression in adolescents: feasibility and preliminary efficacy study. JMIR Form Res. 2023 Nov 22;7:e43260. doi:10.2196/43260. PMID: 37991839.
3. Furman DJ, Hall SA, Avina C, Kulikov VN, Lake JI, Padmanabhan A. Assessing the efficacy and safety of a digital therapeutic for symptoms of depression in adolescents: protocol for a randomized controlled trial. JMIR Res Protoc. 2023 Nov 16;12:e48740. doi:10.2196/48740. PMID: 37971800.
4. Kulikov VN, Crosthwaite PC, Hall SA, Flannery JE, Strauss GS, Vierra EM, et al. A CBT-based mobile intervention as an adjunct treatment for adolescents with symptoms of depression: a virtual randomized controlled feasibility trial. Front Digit Health. 2023 May 23;5:1062471. doi:10.3389/fdgth.2023.1062471. PMID: 37323125.
5. Doukani A, Quartagno M, Sera F, Free C, Kakuma R, Riper H, et al. Comparison of the working alliance in blended cognitive behavioral therapy and treatment as usual for depression in Europe: secondary data analysis of the E-COMPARED randomized controlled trial. J Med Internet Res. 2024 May 31;26:e47515. doi:10.2196/47515. PMID: 38819882.
6. MacLean S, Corsi DJ, Litchfield S, Kucharski J, Genise K, Selaman Z, et al. Coach-facilitated web-based therapy compared with information about web-based resources in patients referred to secondary mental health care for depression: randomized controlled trial. J Med Internet Res. 2020 Jun 9;22(6):e15001. doi:10.2196/15001. PMID: 32515740.
7. Giosan C, Cobeanu O, Mogoaşe C, Szentagotai A, Mureşan V, Boian R. Reducing depressive symptomatology with a smartphone app: study protocol for a randomized, placebo-controlled trial. Trials. 2017 May 12;18(1):215. doi:10.1186/s13063-017-1960-1. PMID: 28494802.
8. Jagayat JK, Kumar A, Shao Y, Pannu A, Patel C, Shirazi A, et al. Incorporating a stepped care approach into internet-based cognitive behavioral therapy for depression: randomized controlled

trial. JMIR Ment Health. 2024 Feb 9;11:e51704. doi:10.2196/51704. PMID: 38173167.
9. Nicol G, Wang R, Graham S, Dodd S, Garbutt J. Chatbot-delivered cognitive behavioral therapy in adolescents with depression and anxiety during the COVID-19 pandemic: feasibility and acceptability study. JMIR Form Res. 2022 Nov 22;6(11):e40242. doi:10.2196/40242. PMID: 36413390.
10. McCue M, Blair C, Fehnert B, King J, Cormack F, Sarkey S, et al. Mobile app to enhance patient activation and patient-provider communication in major depressive disorder management: collaborative, randomized controlled pilot study. JMIR Form Res. 2022 Oct 27;6(10):e34923. doi:10.2196/34923. PMID: 36301599.
11. Crider K, Williams J, Qi YP, Gutman J, Yeung L, Mai C, et al. Folic acid supplementation and malaria susceptibility and severity among people taking antifolate antimalarial drugs in endemic areas. Cochrane Database Syst Rev. 2022 Feb 1;2:CD014217. doi:10.1002/14651858.CD014217. PMID: 36321557.
12. Mathiasen K, Andersen TE, Lichtenstein MB, Ehlers LH, Riper H, Kleiboer A, et al. The clinical effectiveness of blended cognitive behavioral therapy compared with face-to-face cognitive behavioral therapy for adult depression: randomized controlled noninferiority trial. J Med Internet Res. 2022 Sep 7;24(9):e36577. doi:10.2196/36577. PMID: 36069798.
13. Topooco N, Byléhn S, Dahlström Nysäter E, Holmlund J, Lindegaard J, Johansson S, et al. Evaluating the efficacy of internet-delivered cognitive behavioral therapy blended with synchronous chat sessions to treat adolescent depression: randomized controlled trial. J Med Internet Res. 2019 Nov 1;21(11):e13393. doi:10.2196/13393. PMID: 31682572.
14. Lindenberg K, Kindt S, Szász-Janocha C. Effectiveness of cognitive behavioral therapy-based intervention in preventing gaming disorder and unspecified internet use disorder in adolescents: a cluster randomized clinical trial. JAMA Netw Open. 2022 Feb 1;5(2):e2148995. doi:10.1001/jamanetworkopen.2021.48995. PMID: 35179587.
15. Segal ZV, Dimidjian S, Beck A, Boggs JM, Vanderkruik R, Metcalf CA, et al. Outcomes of online mindfulness-based cognitive therapy for patients with residual depressive symptoms: a randomized clinical trial. JAMA Psychiatry. 2020 Jun 1;77(6):563-73. doi:10.1001/jamapsychiatry.2019.4693. PMID: 31995132.

