European Radiology (2019) 29:2322–2329
https://doi.org/10.1007/s00330-018-5791-6

FORENSIC MEDICINE

Forensic age estimation for pelvic X-ray images using deep learning

Yuan Li[1,2] · Zhizhong Huang[3] · Xiaoai Dong[2] · Weibo Liang[1] · Hui Xue[4] · Lin Zhang[1] · Yi Zhang[3] · Zhenhua Deng[2]

Received: 16 May 2018 / Revised: 6 September 2018 / Accepted: 21 September 2018 / Published online: 6 November 2018

Abstract

Purpose To develop a deep learning bone age assessment model based on pelvic radiographs for forensic age estimation and compare its performance to that of the existing cubic regression model.

Materials and method A retrospective collection data of 1875 clinical pelvic radiographs between 10 and 25 years of age was obtained to develop the model. Model performance was assessed by comparing the testing results to estimated ages calculated directly using the existing cubic regression model based on ossification staging methods. The mean absolute error (MAE) and root-mean-squared error (RMSE) between the estimated ages and chronological age were calculated for both models.

Results For all test samples (between 10 and 25 years old), the mean MAE and RMSE between the automatic estimates using the proposed deep learning model and the reference standard were 0.94 and 1.30 years, respectively. For the test samples comparable to those of the existing cubic regression model (between 14 and 22 years old), the mean MAE and RMSE for the deep learning model were 0.89 and 1.21 years, respectively. For the existing cubic regression model, the mean MAE and RMSE were 1.05 and 1.61 years, respectively.

Conclusion The deep learning convolutional neural network model achieves performance on par with the existing cubic regression model, demonstrating predictive ability capable of automated skeletal bone assessment based on pelvic radiographic images.

Key Points

- *The pelvis has considerable value in determining the bone age.*
- *Deep learning can be used to create an automated bone age assessment model based on pelvic radiographs.*
- *The deep learning convolutional neural network model achieves performance on par with the existing cubic regression model.*

Keywords Age determination by skeleton · Forensic anthropology · Pelvis · Radiography · Machine learning

Abbreviations

CA	Chronological age
CNN	Convolutional neural network
EBA-CNN	Bone age estimated by the CNN
EBA-CR	Bone age calculated by the cubic regression model
ICA	Ossification centre of the iliac crest
IW	Iliac wing
KK-SM	Kreitner and Kellinghaus ossification staging methods
MAE	Mean absolute difference
RMSE	Root-mean-squared error
ROC	Receiver operating characteristic

Introduction

Age estimation in living individuals is an area of increasing importance for both paediatric and forensic medicine. In contrast

Yuan Li and Zhizhong Huang contributed equally to this work.

✉ Yi Zhang
yzhang@scu.edu.cn

✉ Zhenhua Deng
fydzh63@163.com

[1] Department of Forensic Genetics, West China School of Preclinical and Forensic Medicine, Sichuan University, Chengdu 610041, Sichuan, People's Republic of China

[2] Department of Forensic Pathology, West China School of Preclinical and Forensic Medicine, Sichuan University, No. three, 17 South Renmin Road, Wuhou District, Chengdu City 610041, Sichuan, People's Republic of China

[3] College of Computer Science, Sichuan University, No.24 South Section 1, Yihuan Road, Chengdu 610065, China

[4] Department of Radiology, West China Hospital, Sichuan University, Chengdu 610041, Sichuan, People's Republic of China

to the former, the latter typically focuses on the age assessment of adolescents and young adults, who do not have formal age documentation to determine their legal status. In most countries, the 14th, 16th, and 18th years of life are relevant age thresholds for legal assessment. Age can be estimated very accurately during childhood, generally performed via radiological examination of the left hand using either the Greulich and Pyle method or Tanner–Whitehouse method [1]. However, with increasing age, age estimation becomes increasingly difficult. The pelvis has also been proven to have considerable value in determining skeletal age of the older adolescent due to its relatively late completion of maturation; most existing research efforts focused on radiological examination of the Risser sign and iliac crest ossification process [2–8].

Deep learning has dramatically improved the state-of-the-art techniques in image classification, object detection, and many other domains, such as drug discovery and genomics [9]. Deep learning methods have also been successfully introduced into the clinical field in various applications, including classification of skin cancer [10], detection of diabetic retinopathy in retinal fundus photographs [11, 12], and evaluation of lung nodules on chest computed tomography scans [13].

Automated assessment of radiographs for determining skeletal maturity or bone age is a logical application and ideal target for deep learning. The age estimation procedure of comparing one or several radiographic images to a reference standard has been used for several decades. It results in a single quantitative output of an estimated bone age [14–16]. Deep learning approaches for assessing bone age automatically have already been proven to be extremely successful in clinical practice, achieving accuracy similar to that of human practitioners [17–22].

However, existing methods all refer to the skeletal bone age assessment procedure used in paediatric radiology for diagnostic or therapeutic investigations of endocrinology problems. Their study datasets exclusively consider individuals under 18 years of age, which limits their applicability to forensic purposes, particularly for estimation of the 18-year age threshold. To date, no deep learning model has been designed for forensic age estimation purposes. For this reason, we propose a deep learning system to automatically performe age estimation by using a fine-tuned convolutional neural network. (CNN) on a dataset of 1408 pelvis X-ray radiographs from individuals between 10 and 25 years of age.

Materials and method

Data acquisition

A dataset composed of 1875 conventional pelvic anteroposterior radiography images from the West China Hospital of Sichuan University between January 2010 and March 2017 was collected to develop our deep learning model (Table 1). All participants came from the West China Han group and individuals showing evident deformity or disease in the pelvic region were intentionally excluded from our research. Date of birth and date of examination were extracted automatically from the radiological images using a Python script (Python Software Foundation). The difference between these two dates was used as the ground truth for training our model because these individuals represent a healthy population.

The subjects were all between 10 and 25 years of age. The age range of the dataset was chosen according to previous scientific studies, which reported that the iliac crest ossification process does not begin before age 12 and is completed by a maximum age of 24 [3, 22–25]. The population suitable for reliable examination included 1072 females (57.2%) and 803 males (42.8%).

We split the dataset into 1498 training and validation images and 377 test images. Regarding the training data, groups from 15 to 25 years, 1003 images, are a portion of Zhang et al's study data to establish the cubic regression model for bone age estimation [5]. Images with significant superposition of abdominal organs over the iliac crest were removed from the training set, but still used for testing.

This study was performed with the approval of the ethics committee of the West China Hospital of Sichuan University. Informed consent was waived.

Table 1 Distribution of sample images by gender, age, and dataset

Age group (year)	Male		Female		Total
	Train	Test	Train	Test	
10.00–10.99	65	14	59	10	148
11.00–11.99	40	8	42	9	99
12.00–12.99	55	13	46	11	125
13.00–13.99	48	13	44	11	116
14.00–14.99	55	18	41	11	125
15.00–15.99	62	17	28	11	118
16.00–16.99	68	14	39	11	132
17.00–17.99	52	14	30	11	107
18.00–18.99	53	11	22	10	96
19.00–19.99	45	11	27	7	90
20.00–20.99	57	11	36	10	114
21.00–21.99	39	11	44	10	104
22.00–22.99	61	14	37	11	123
23.00–23.99	45	12	41	12	110
24.00–24.99	65	14	51	12	142
25.00–25.99	54	13	47	12	126
Total	864	208	634	169	1875

Training details

Inspired by the concept of transfer learning, a modified AlexNet pre-trained using the ImageNet dataset was adopted in this study to alleviate the requirement for an extremely large dataset [26]. The original convolutional layers were maintained for feature extraction. The regression portion, which was composed of several fully connected layers in the original AlexNet, was replaced by three new fully connected layers with 2048, 1024, and 1 neurons. The first two fully connected layers are followed by ReLU layers and the final layer outputs the prediction results.

All pelvic radiographs were cropped centrally and resized to 227 × 227 pixels in order to make them compatible with the original dimensions of the AlexNet. Data augmentation was performed by rotating the training images by 0°, 90°, 180°, and 270° to enhance the performance of the model.

During the training procedure, the parameters of the convolutional layers were fixed to avoid overfitting and only the parameters of the fully connected layers were fine-tuned using our training set [27, 28]. A mean squared error loss function was adopted and training was performed utilising the Adam method. The learning rate was initialised to 10–4 and gradually decreased to 10–6. The batch size was set to 256. The proposed network was implemented using Caffe and all experiments were performed on a PC equipped with a graphical processing unit (AMD X4 870K CPU, 16 GB RAM, and GeForce GTX 1080 Ti).

We released our CNN model as an open source, it is available at: https://github.com/Hzzone/Bone-Age-Assessment.

Evaluation and statistical analysis

Two datasets were used as test sets to evaluate the performance of our model relative to the ground truth chronological age and existing cubic regression model proposed by Zhang et al [5].

The first test set, which contains 377 pelvis radiographs, was utilised to evaluate the performance of our model relative to the ground truth chronological age.

The second test set, which contains 133 pelvis radiographs and is a subset of the first set, was utilised to evaluate the performance of our model relative to that of the existing cubic regression model for age estimation. This set also is a portion of Zhang et al's study data. Each image in the second dataset was labelled with (1) the sum of evaluated scores of the developmental status of ossification for the iliac crest and ischial tuberosity according to the Kreitner and Kellinghaus ossification staging methods (KK-SM), (2) bone age calculated by the cubic regression model, and (3) its chronological age. Regarding the cubic regression model, the equation for a male is age = $14.408 + 0.475X - 0.064X^2 + 0.005X^3$. The equation for a female is age = $18.196 - 1.637X + 0.204X^2 - 0.005X^3$.

Pearson correlation analysis and a Bland–Altman plot were used to assess the accuracies of the estimated bone ages from the proposed CNN model. We compare the mean absolute error (MAE) between the model estimates (CNN model and existing cubic regression model) and reference standard bone ages. Root-mean-squared error (RMSE) was also calculated as the square root of the sum of squares of paired differences. Receiver operating characteristic (ROC) analysis was performed to measure model ability to estimate of 14-, 16-, 18-, 20-, and 22-year age thresholds using the CNN.

For both methods, the differences between males and females were analysed using the Mann–Whitney U test with significance implied for values of $p < 0.05$. Statistical significance was determined by using paired t tests for comparing means (i.e. mean and MAE) and F tests for comparing variances (i.e. RMSE).

The Statistical Package for the Social Sciences (SPSS) version 22.0 (SPSS Inc.) was used for performing statistical analysis.

Results

Table 1 summarises the general data distributions, which are grouped by gender, age, and dataset (training set and test set), for CNN model training.

When evaluating CNN model performance on the first test set, when comparing the estimated bone age produced by the CNN model directly to the real chronological ages, the mean MAE and RMSE were 0.94 years and 1.30 years, respectively. The differences between the sexes were not statistically significant ($p = 0.873$). The minimum MAE and RMSE values were achieved for the age group of 11 years, followed by the age group of 17 years (Table 2). The maximum MAE and RMSE values occurred for the age group of 25 years, followed by the age groups of 24 years and 23 years (Table 2).

The estimated bone ages from our CNN model are significantly correlated to the real chronological ages ($R^2 = 0.9288$, Fig. 1). Statistical analyses indicated no differences between the sexes ($p = 0.57$). A Bland–Altman plot of the agreement between estimated bone ages and reference bone ages revealed a mean difference of 0.1 years (95% limit of agreement, ± 2.60, Fig. 1).

Table 3 and Fig. 2 present a comparison of the CNN model estimates and existing cubic regression model estimates, which utilised the second test set. When comparing the results of the CNN model to the ground truth ages, the mean MAE and RMSE were 0.91 years and 1.23 years, respectively. When we compared the cubic regression model results to the ground truth ages, the mean MAE and RMSE were 1.05 years and 1.61 years, respectively. The CNN model's MAE was slightly lower when compared with the cubic regression model, but not for the RMAE and the mean (MAE, $p = 0.001$; RMSE, $p = 0.077$; mean, $p = 0.482$).

Table 2 Comparison of MAE and RMSE values between estimated bone ages from the proposed CNN model and the real chronological ages for the first test set

Age group (year)	Male		Female		Total		EBA-CNN
	MAE	RMSE	MAE	RMSE	MAE	RMSE	Mean ± SD
10.00–10.99	1.22	1.43	0.11	0.15	0.76	1.10	10.88 ± 1.03
11.00–11.99	0.43	0.53	0.53	0.64	0.49	0.59	11.49 ± 0.59
12.00–12.99	0.96	1.31	0.76	1.12	0.87	1.23	12.19 ± 1.27
13.00–13.99	0.74	1.06	1.32	1.76	0.96	1.40	13.45 ± 1.52
14.00–14.99	0.54	0.97	0.93	1.57	0.54	0.90	14.58 ± 0.84
15.00–15.99	0.55	1.02	0.68	0.92	0.61	0.97	15.77 ± 0.91
16.00–16.99	0.75	1.06	1.04	1.36	0.89	1.20	16.82 ± 1.23
17.00–17.99	0.38	0.59	0.82	1.07	0.59	0.83	17.54 ± 0.81
18.00–18.99	0.69	0.90	0.79	1.10	0.74	1.00	18.67 ± 1.03
19.00–19.99	0.69	0.94	0.68	0.82	0.68	0.88	19.07 ± 0.96
20.00–20.99	0.74	1.04	1.16	1.48	0.86	1.24	20.34 ± 1.27
21.00–21.99	1.26	1.45	0.81	1.00	1.05	1.26	21.19 ± 1.29
22.00–22.99	1.34	1.62	1.09	1.61	1.10	1.41	21.94 ± 1.32
23.00–23.99	1.19	1.56	1.27	1.46	1.23	1.51	22.87 ± 1.49
24.00–24.99	2.11	1.95	2.71	2.47	1.86	2.17	23.05 ± 1.78
25.00–25.99	2.18	2.21	2.28	2.21	1.67	2.01	24.01 ± 1.54
Total	1.05	1.23	1.08	1.30	0.94	1.30	17.76 ± 4.28

MAE mean absolute error, *RMSE* root-mean-squared error, *EBA-CNN* estimated bone ages from the proposed CNN model

Table 4 lists that the area under the ROC curve values for the CNN model when estimating age thresholds of 14, 16, 18, and 20 years, where all values are greater than 0.91. These results indicate strong prediction performance by the CNN model.

Discussion

The iliac crest apophysis provides an excellent subject for the application of forensic age diagnostics to the living, particularly for determining age thresholds of 14, 16, and 18 years [5]. Recently, the advent and proliferation of CNN using X-ray hand radiological images have facilitated new applications to the evaluation of bone age [17–21]. For this reason, we developed a CNN model for forensic age estimation based on X-ray pelvis radiological images, particularly for individuals over 18 years old.

We developed a CNN model for the estimation of bone age based on comparisons with real chronological ages and achieved performance comparable to that of the existing cubic regression model based on a grading system of the developmental status of ossification of the iliac crest and ischial tuberosity in pelvis radiography [5]. Our model was developed based on a training set of 1498 clinical pelvis radiographs. By employing a transfer learning algorithm, our method achieved competitive performance for pelvis radiograph analysis without the need for a highly specialised deep learning machine or database with millions of example images.

To validate the CNN model, real chronological ages were used as a reference standard for true bone ages. We achieved an aggregate test set MAE and RMSE of 0.94 and 1.30, respectively (Table 3). Estimated bone ages were significantly correlated with the reference ages ($r = 0.916$; $p < 0.05$). Additionally, in a head-to-head comparison with the existing regression model for predicting bone age, the performance of

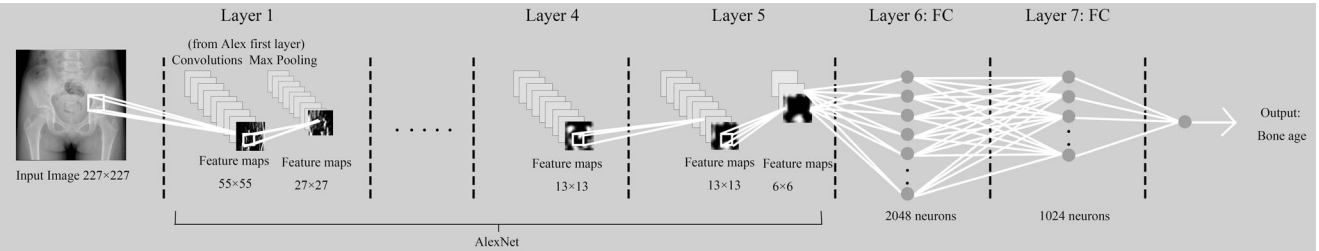

Fig. 1 The modified AlexNet consists of five convolutional layers and three new fully connected layers. The convolutional layers are frozen and only the parameters of the fully connected layers were fine-tuned

Table 3 Comparison of MAE and RMSE values between the proposed CNN model and existing regression model for the second test set

Age group (year)	CNN model			Cubic regression model		
	Mean ± SD	MAE	RMSE	Mean ± SD	MAE	RMSE
14.00–14.99	15.06 ± 1.26	0.9	1.39	15.48 ± 0.57	0.99	1.16
15.00–15.99	15.74 ± 0.97	0.83	1.18	15.22 ± 0.37	0.38	0.45
16.00–16.99	16.88 ± 1.51	1.11	1.33	16.31 ± 1.14	0.72	0.81
17.00–17.99	17.78 ± 0.94	0.84	1.09	17.66 ± 1.94	1.60	1.75
18.00–18.99	18.74 ± 1.10	0.86	1.09	19.27 ± 1.58	1.56	1.73
19.00–19.99	19.02 ± 0.98	0.68	0.92	19.82 ± 1.68	1.43	1.66
20.00–20.99	20.49 ± 1.10	0.86	1.16	20.74 ± 2.42	2.02	2.32
21.00–21.99	21.19 ± 1.29	1.05	1.26	20.82 ± 1.68	1.42	1.71
22.00–22.99	21.90 ± 1.35	1.15	1.45	22.32 ± 1.41	1.14	1.45
Total	19.42 ± 2.79	0.89	1.21	19.29 ± 2.42	1.39	1.62

MAE mean absolute error, *RMSE* root-mean-squared error

the proposed model was better than that of the existing regression model. Therefore, we can conclude that our deep learning-based model performs at a level similar to that of the existing regression model.

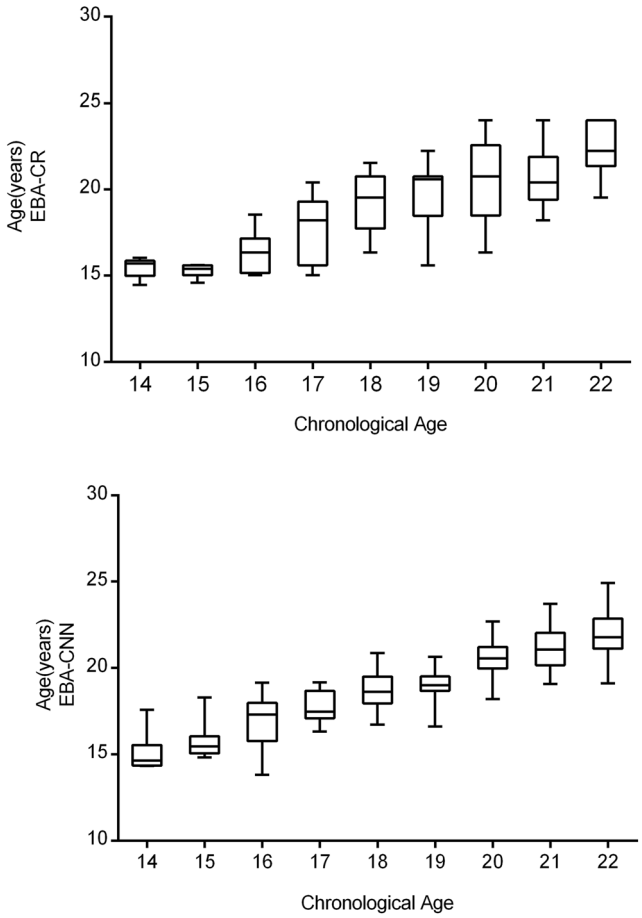

Fig. 2 Correlation between real chronological age and estimated bone age from our CNN model (top). Bland–Altman plot of real chronological age and estimated bone age from our CNN model (bottom)

Our results (MAE = 0.94 and RMSE = 1.30 for the first test set, and MAE = 0.89 and RMSE = 1.21 for the second test set) are slightly higher than those produced by deep learning methods for bone age estimation using left-hand X-ray images. In previous validation studies on deep learning architectures based on left-hand X-ray images, the MAE values between automatic age estimates by deep learning architectures and reference bone ages ranged from 0.54 to 0.80 years [17, 18, 20, 21]. Additionally, our statistical analyses indicated a high positive correlation between our estimates and real ages ($r = 0.916$; $p < 0.05$), but this value is still less significant than those reported by deep learning studies on hand X-rays ($r = 0.992$; $p < 0.001$) [18]. These results are not surprising because our subjects were older than those providing hand samples in previous studies. Humans and deep learning methods both produce more accurate results for age estimation during childhood when using hand X-ray radiographic images containing meaningful features that are strongly related to age changes.

Researchers have attempted to use regression models obtained from pelvic skeletons to validate pelvis maturation grading systems. In Cameriere's approach [29], the area of the ossification centre of the iliac crest (ICA) and area of the iliac wing (IW) were measured to derive a linear regression equation for the ICA/IW ratio, which facilitates the estimation of age. However, the ICA/IW ratio showed only moderate correlation with real chronological ages ($R^2 < 0.4$). In 2018, this moderate correlation was proved again by Viola Bartolini et al [2]. Their study revealed a Pearson's correlation coefficient of 0.52 between the ICA/IW ratio and chronological age.

In a comprehensive study by Zhang et al in 2016 [5], the cubic regression formula for estimating bone age was developed. According to an individual's developmental status, the ossification of the iliac crest and ischial tuberosity was scored by utilising the KK-MS classification scheme. This formula resulted in a higher R^2 value compared to previous methods ($R^2 = 0.744$ for females and $R^2 = 0.753$ for males). To date, this

Table 4 Area under the ROC curve for determining age thresholds of 14, 16, 18, 20, and 22 years

	Sensitivity (CI)	Specificity (CI)	AUC (CI)
M-14	98.63 (95.1–99.8)	91.67 (80.0–97.7)	0.951 (0.911–0.977)
F-14	98.41 (94.4–99.8)	87.80 (73.8–95.9)	0.931 (0.881–0.964)
M-16	98.33 (94.1–99.8)	91.89 (83.2–97.0)	0.951 (0.911–0.977)
F-16	97.12 (91.8–99.4)	90.48 (80.4–96.4)	0.938 (0.890–0.969)
M-18	95.83 (89.7–98.9)	96.94 (91.3–99.4)	0.964 (0.927–0.985)
F-18	97.56 (91.5–99.7)	88.24 (79.4–94.2)	0.929 (0.879–0.963)
M-20	92.00 (83.4–97.0)	97.48 (92.8–99.5)	0.947 (0.906–0.974)
F-20	86.36 (75.7–93.6)	97.03 (91.6–99.4)	0.917 (0.864–0.954)
M-22	62.26 (47.9–75.2)	97.16 (92.9–99.2)	0.797 (0.734–0.851)
F-22	57.78 (42.2–72.3)	98.36 (94.2–99.8)	0.781 (0.710–0.841)

ROC receiver operating characteristic, *AUC* area under ROC curve

study achieved the best performance in terms of direct estimation of age, although its results are not as accurate as those reported by studies on dental mineralisation [29–31], carpus ossification, and epiphysis of the ulna and radius [32, 33].

However, our proposed CNN model outperforms the cubic regression model on average (MAE = 0.89, RMSE = 1.21 and MAE = 1.39, RMSE = 1.62, respectively) (Fig. 3). These values were achieved after removal of images outside the age range of 14–22 years to make the analysis comparable to previous studies. These ages were discarded in previous studies because there is little change in the mean score of ossification among individuals aged 23 to 25 years using the KK-MS grading system.

We also noted that the age groups of 23, 24, and 25 years presented the maximum MAE and RMSE values in the CNN model, which leads to slightly higher MAE and RMSE values for the first test set compared to the second test set (MAE = 0.94, RMSE = 1.30 and MAE = 0.89, RMSE = 1.21, respectively). However, this change in MAE and RMSE values between the two test sets is negligible, which further validates the reliability of our model.

We demonstrated the effectiveness of a pre-trained ImageNet fine-tuned CNN by using a process known as transfer learning to perform age estimation based on pelvis radiography images. The transfer learning technique significantly speeds up training and helps prevent overfitting, particularly when examining domains with limited data. Although medical images are considerably different from natural images, transfer learning can be utilised by applying using generic filter banks trained on large datasets and adjusting their parameters to render high-level features for specific medical applications. Recent works [34, 35] have demonstrated the effectiveness of transfer learning from general pictures to age estimation by fine-tuning several (or all) network layers using hand X-ray radiographic images [17, 19, 20]. Additionally, Spampinato et al [20] and Mutasa et al [21] designed and

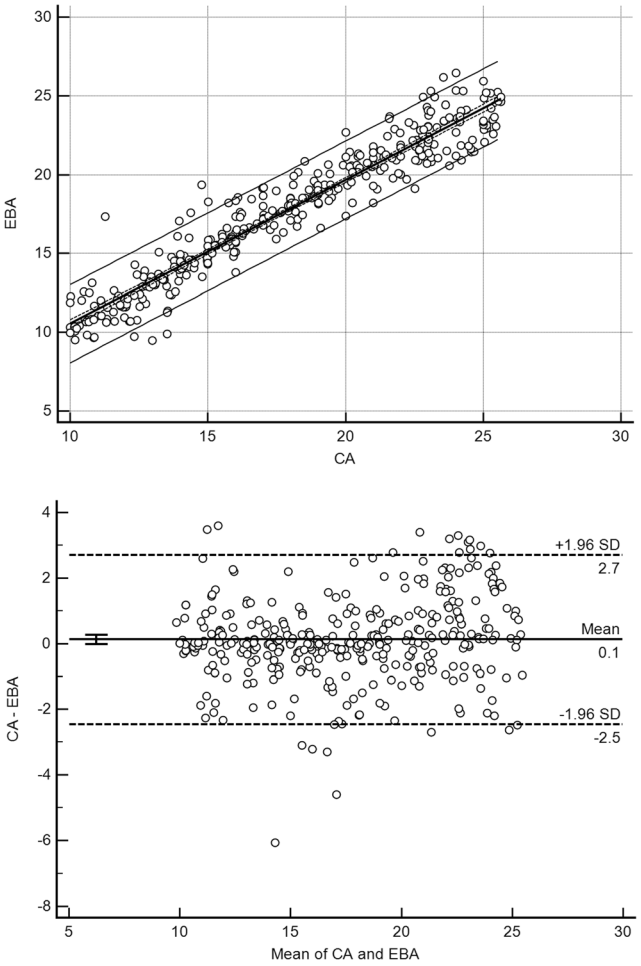

Fig. 3 Box plot for EBA-CR and EBA-CNN compared to the ground truth (EBA-CR represents the bone age calculated by the cubic regression model, CA represents the real chronological age, and EBA-CNN represents the bone age estimated by the CNN model). The CNN outperforms the cubic regression model on average

trained customised neural network algorithms from scratch and carefully calibrated them for the evaluation of bone age. The purpose-built neural networks designed by Spampinato et al provided improved performance compared to networks derived from pre-trained imaging datasets. Furthermore, Mutasa et al achieved the best published performance by utilising deep learning for bone age assessment (MAE = 0.54). Therefore, they concluded that a small, customised architecture incorporating advanced CNN strategies can indeed be trained from scratch to realise significant improvements in algorithm accuracy. These promising results inspired further study on custom CNN designed and trained from scratch to improve their effectiveness at assessing bone age.

Our study faced several limitations. First, our cases were limited to patients of a single ethnicity, but previous studies have reported the differences in growth patterns at certain ages among different populations [36]. Second, although our test data included samples with large areas of superposition of abdominal organs over the iliac crest, we attempted to

standardise our samples by excluding all images with evident artefacts, diseases, or major projection errors. Third, our model may be not effective in predicting the bone age of individuals with clinical indications or in low socio-economic status. Finally, as with most other current CNN for medical imaging, our algorithm was trained to evaluate a single diagnosis and cannot provide information regarding ancillary findings.

Conclusion

We developed a deep learning system using transfer learning techniques to perform automatic bone age estimation. It can effectively handle all possible cases of automated skeletal bone age assessment, even for samples from individuals aged 19, 20, and 21 years. However, it may not see practical application in determining ages over 22 years. Compared to the existing cubic regression model, our CNN model achieves better average performance but is less accurate than deep learning architectures based on hand X-ray radiographic images.

Funding The authors state that this work has not received any funding.

Compliance with ethical standards

Guarantor The scientific guarantor of this publication is Zhen-hua Deng.

Conflict of interest The authors declare that they have no conflict of interest.

Statistics and biometry No complex statistical methods were necessary for this paper.

Informed consent Informed consent was waived.

Ethical approval This study was performed with the approval of the ethics committee of the West China Hospital of Sichuan University.

Study subjects or cohorts overlap Some study subjects or cohorts have been previously reported in Zhang K, Dong XA, Fan F, Deng ZH (2016) Age estimation based on pelvic ossification using regression models from conventional radiography. *International Journal of Legal Medicine* 130:1143–1148.

Methodology
• Diagnostic or prognostic study
• Performed at one institution

References

1. Schmeling A, Grundmann C, Fuhrmann A et al (2008) Criteria for age estimation in living individuals. Int J Legal Med 122:457
2. Bartolini V, Pinchi V, Gualco B et al (2018) The iliac crest in forensic age estimation: evaluation of three methods in pelvis X-rays. Int J Legal Med 132:1–10
3. Wittschieber D, Schmeling A, Schmidt S, Heindel W, Pfeiffer H, Vieth V (2013) The Risser sign for forensic age estimation in living individuals: a study of 643 pelvic radiographs. Forensic Sci Med Pathol 9:36–43
4. Zhang K, Dong XA, Chen XG, Li Y, Deng ZH (2015) Forensic age estimation through evaluation of the apophyseal ossification of the iliac crest in Western Chinese. Forensic Sci Int 252:192.e191–192.e195
5. Zhang K, Dong XA, Fan F, Deng ZH (2016) Age estimation based on pelvic ossification using regression models from conventional radiography. Int J Legal Med 130:1143–1148
6. Lottering N, Reynolds MS, Macgregor DM et al (2016) Apophyseal ossification of the iliac crest in forensic age estimation: computed tomography standards for modern Australian subadults. J Forensic Sci 62:292–307
7. Schmidt S, Schiborr M, Pfeiffer H, Schmeling A, Schulz R (2013) Sonographic examination of the apophysis of the iliac crest for forensic age estimation in living persons. Sci Justice 53:395–401
8. Schmidt S, Schmeling A, Zwiesigk P, Pfeiffer H, Schulz R (2011) Sonographic evaluation of apophyseal ossification of the iliac crest in forensic age diagnostics in living individuals. Int J Legal Med 125:271
9. Lecun Y, Bengio Y, Hinton G (2015) Deep learning. Nature 521:436
10. Esteva A, Kuprel B, Novoa RA et al (2017) Dermatologist-level classification of skin cancer with deep neural networks. Nature 542:115–118
11. Kermany DS, Goldbaum M, Cai W et al (2018) Identifying medical diagnoses and treatable diseases by image-based deep learning. Cell 172:1122–1131.e1129
12. Gulshan V, Peng L, Coram M et al (2016) Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. JAMA 316:2402
13. Setio AA, Ciompi F, Litjens G et al (2016) Pulmonary nodule detection in CT images: false positive reduction using multi-view convolutional networks. IEEE Trans Med Imaging 35:1160–1169
14. Mansourvar M, Ismail MA, Herawan T, Raj RG, Kareem SA, Nasaruddin FH (2013) Automated bone age assessment: motivation, taxonomies, and challenges. Comput Math Methods Med 2013:391626–391626
15. Wang YH, Liu TA, Wei H, Wan L, Ying CL, Zhu GY (2016) Automated classification of epiphyses in the distal radius and ulna using a support vector machine. J Forensic Sci 61:409–414
16. Van Rijn RR, Thodberg HH (2013) Bone age assessment: automated techniques coming of age? Acta Radiol 54:1024
17. Larson DB, Chen MC, Lungren MP, Halabi SS, Stence NV, Langlotz CP (2017) Performance of a deep-learning neural network model in assessing skeletal maturity on pediatric hand radiographs. Radiology 287:170236
18. Kim JR, Shim WH, Yoon HM et al (2017) Computerized bone age estimation using deep learning based program: evaluation of the accuracy and efficiency. AJR Am J Roentgenol 209:1
19. Lee H, Tajmir S, Lee J et al (2017) Fully automated deep learning system for bone age assessment. J Digit Imaging 30:427–441
20. Spampinato C, Palazzo S, Giordano D, Aldinucci M, Leonardi R (2016) Deep learning for automated skeletal bone age assessment in X-ray images. Med Image Anal 36:41–51
21. Mutasa S, Chang PD, Ruzal-Shapiro C, Ayyala R (2018) MABAL: a novel deep-learning architecture for machine-assisted bone age labeling. J Digit Imaging 9:1–7
22. Wittschieber D, Vieth V, Wierer T, Pfeiffer H, Schmeling A (2013) Cameriere's approach modified for pelvic radiographs: a novel method to assess apophyseal iliac crest ossification for the purpose of forensic age diagnostics. Int J Legal Med 127(4):825–829
23. Wittschieber D, Vieth V, Domnick C, Pfeiffer H, Schmeling A (2013) The iliac crest in forensic age diagnostics: evaluation of the apophyseal ossification in conventional radiography. Int J Legal Med 127:473–479

24. Diedrichs V, Wagner UA, Seiler W, Schmitt O (1998) Reference values for development of the iliac crest apophysis (Risser sign). Z Orthop Ihre Grenzgeb 136:226
25. Coqueugniot H, Weaver TD (2007) Brief communication: Infracranial maturation in the skeletal collection from Coimbra, Portugal: new aging standards for epiphyseal union. Am J Phys Anthropol 134:424–437
26. Krizhevsky A, Sutskever I, Hinton GE (2012) ImageNet classification with deep convolutional neural networks. International Conference on Neural Information Processing Systems, pp 1097–1105
27. Pan SJ, Yang Q (2010) A survey on transfer learning. IEEE T Knowl Data En 22:1345–1359
28. Yosinski J, Clune J, Bengio Y, Lipson H (2014) How transferable are features in deep neural networks?. International Conference on Neural Information Processing Systems, pp 3320–3328
29. Cameriere R, Ferrante L, Belcastro MG, Bonfiglioli B, Rastelli E, Cingolani M (2007) Age estimation by pulp/tooth ratio in canines by mesial and vestibular peri-apical X-rays. J Forensic Sci 52: 1151–1155
30. Corradi F, Pinchi V, Barsanti I, Garatti S (2013) Probabilistic classification of age by third molar development: the use of soft evidence. J Forensic Sci 58:51–59
31. Pinchi V, Norelli GA, Pradella F, Vitale G, Rugo D, Nieri M (2012) Comparison of the applicability of four odontological methods for age estimation of the 14 years legal threshold in a sample of Italian adolescents. J Forensic Odontostomatol 2:17–25
32. Pinchi V, Luca FD, Focardi M et al (2016) Combining dental and skeletal evidence in age classification: pilot study in a sample of Italian sub-adults. Leg Med (Tokyo) 20:75–79
33. Pinchi V, Luca FD, Ricciardi F et al (2014) Skeletal age estimation for forensic purposes: a comparison of GP, TW2 and TW3 methods on an Italian sample. Forensic Sci Int 238:83–90
34. Hoochang S, Roth HR, Gao M et al (2016) Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. IEEE Trans Med Imaging 35:1285
35. Tajbakhsh N, Shin JY, Gurudu SR et al (2016) Convolutional neural networks for medical image analysis: fine tuning or full training? IEEE Trans Med Imaging 35:1299–1312
36. Zhang A, Sayre J, Vachon L, Liu B, Huang H (2009) Racial differences in growth patterns of children assessed on the basis of bone age. Radiology 250:228