



---

*Research article*

## **Bone age assessment from articular surface and epiphysis using deep neural networks**

**Yamei Deng<sup>1</sup>, Yonglu Chen<sup>1,\*</sup>, Qian He<sup>1</sup>, Xu Wang<sup>2</sup>, Yong Liao<sup>3</sup>, Jue Liu<sup>1</sup>, Zhaoran Liu<sup>1</sup>, Jianwei Huang<sup>1</sup> and Ting Song<sup>1</sup>**

<sup>1</sup> Department of Radiology, Guangdong Provincial Key Laboratory of Major Obstetric Diseases, Guangdong Provincial Clinical Research Center for Obstetrics and Gynecology, The Third Affiliated Hospital of Guangzhou Medical University, Guangzhou 510150, China

<sup>2</sup> School of Automation, Guangdong University of Technology, Guangzhou 510006, China

<sup>3</sup> School of physics, electronics and electrical engineering, Xiangnan University, Chenzhou 423000, China

\* **Correspondence:** Email: 2014683060@gzhmu.edu.cn.

**Abstract:** Bone age assessment is of great significance to genetic diagnosis and endocrine diseases. Traditional bone age diagnosis mainly relies on experienced radiologists to examine the regions of interest in hand radiography, but it is time-consuming and may even lead to a vast error between the diagnosis result and the reference. The existing computer-aided methods predict bone age based on general regions of interest but do not explore specific regions of interest in hand radiography. This paper aims to solve such problems by performing bone age prediction on the articular surface and epiphysis from hand radiography using deep convolutional neural networks. The articular surface and epiphysis datasets are established from the Radiological Society of North America (RSNA) pediatric bone age challenge, where the specific feature regions of the articular surface and epiphysis are manually segmented from hand radiography. Five convolutional neural networks, i.e., ResNet50, SENet, DenseNet-121, EfficientNet-b4, and CSPNet, are employed to improve the accuracy and efficiency of bone age diagnosis in clinical applications. Experiments show that the best-performing model can yield a mean absolute error (MAE) of 7.34 months on the proposed articular surface and epiphysis datasets, which is more accurate and fast than the radiologists. The project is available at <https://github.com/YameiDeng/BAANet/>, and the annotated dataset is also published at <https://doi.org/10.5281/zenodo.7947923>.

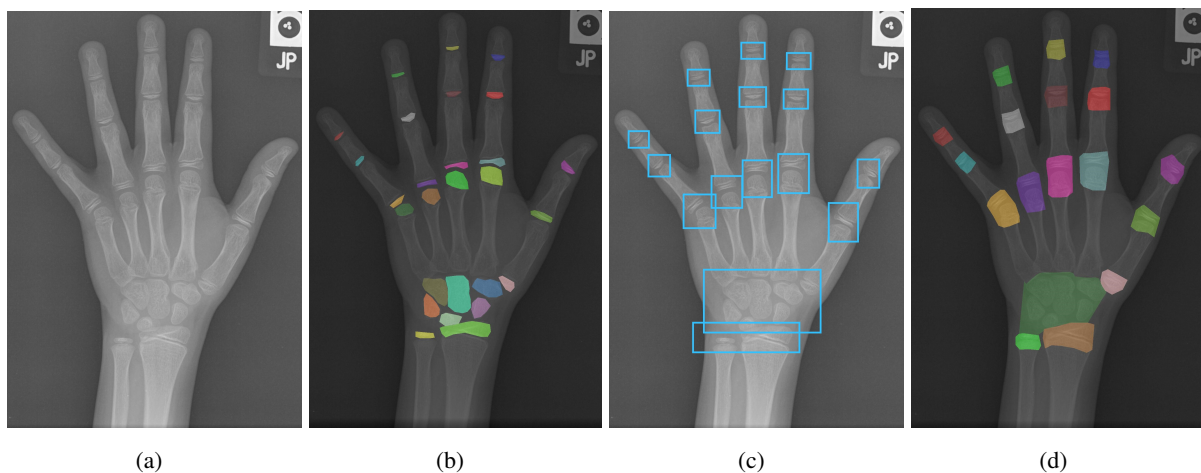
**Keywords:** bone age assessment; articular surface; epiphysis; hand radiography; deep neural networks

---

## 1. Introduction

Bone age assessment not only plays a vital role in the health care of children, but also can monitor the growth development, especially for the diagnosis of genetic and endocrine diseases [1–4], such as adrenal cortical hyperplasia [5], precocious puberty [6], and pituitary dwarfism [7]. Therefore, accurate bone age assessment is of great significance for preventing and treating developmental system diseases.

In clinical diagnosis, bone age assessment mainly depends on experienced experts examining hand radiography through Greulich-Pyle (GP) [8], Tanner-Whitehouse (TW2) [9] and RUS-CHN [10]. Specifically, the GP approach, a simple atlas-based method, can determine the bone age by matching the hand radiograph with the most similar bone region in the reference atlas [8]. The TW is a scoring method by adding the scores that vary by race and gender in a specific set of regions of interest [11]. For the RUS-CHN method, it tends to evaluate each bone in the hand radiography, including radius, ulna, metacarpal bone, and phalange [12], as shown in Figure 1(b). These traditional methods estimate skeletal age by evaluating the region of interest (ROI) in hand radiography, but they are time-consuming and always lead to subjective variability between predictions and concrete results.



**Figure 1.** Regions of interest of different methods for bone age assessment. (a) Raw hand radiograph. (b) Traditional manual diagnosis: feature bone areas. (c) Existing deep learning methods: general annotations by boxes. (d) Ours: articular surface.

Since the deep learning method has shown excellent performance in many fields [13–15], it is recommended for bone age assessment, and it can be divided into ROI-free [16–19] and ROI-based [20–22] approaches. In the ROI-free methods, Mutasa et al. [23] propose a hidden layer-customized residual network for age prediction, where the convolutional neural networks (CNNs) take images as the input and predict bone ages. These networks can estimate the skeletal age through feature extraction, but some feature regions of bone are ignored by the networks [21]. For the ROI-based approach, it can first obtain the regions of interest according to prior knowledge and then generate predicted bone age results. For example, Li et al. [24] propose a region aggregation graph convolutional network to obtain the bone age, where the convolutional neural network (CNN) is utilized to extract the key regions. Liu et al. [21] propose a novel self-supervised attention network to perform bone age assessment by discovering the informative ROIs in hand radiography, as shown in

Figure 1(c). Although these ROI-based methods show comparable performance on the bone age assessment, the specific ROIs of hand radiography are not investigated, and these regions of interest are exactly expected in clinical diagnosis.

From the above observations, the existing methods for bone age assessment from hand radiography have several problems, which can be described as follows:

- 1) The clinical manual bone age diagnosis method is based on the region of interest for bone age diagnosis, but it is very time-consuming and may produce large errors.
- 2) The ROI-free deep learning methods can predict bone age by performing feature extraction directly on the whole image, while some feature regions of bone are ignored.
- 3) Although the ROI-based approaches can estimate the skeletal age through general regions of interest, most of them have not explored the specific ROIs used in clinical applications.

To address such issues, the articular surface and epiphysis datasets are proposed for bone age assessment using deep convolutional neural networks. Specifically, for the first problem, deep learning networks are utilized to assess bone age in the regions of interest to accelerate the process and improve accuracy. For the last two issues, the specific ROIs from hand radiography, i.e., articular surface and epiphysis, are segmented and exploited to evaluate bone age. The main contributions of this article are described as follows:

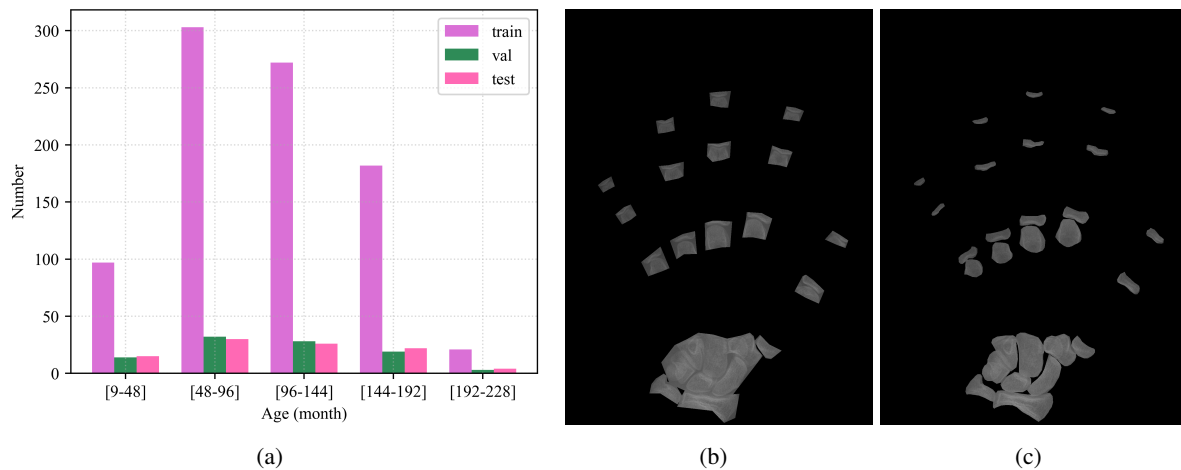
- 1) The articular surface and epiphysis datasets, derived from the RSNA dataset [25], are proposed for bone age assessment, where the key feature regions are marked by the radiologists using LableMe 4.6.0 [26].
- 2) Five deep learning models, including ResNet50 [27], SENet [28], DenseNet-121 [29], EfficientNet-b4 [30], and CSPNet [31], are trained for bone age assessment, in which the best-performing model can achieve a mean square error (MSE) of 7.23 months.
- 3) Three bone feature regions, including articular surface, epiphysis, and epiphyseal line from hand radiography, are used as the benchmarks to perform bone age assessment. Extensive experiments demonstrate that the articular surface and epiphysis from hand radiography are beneficial for bone age assessment.

The rest of this paper is organized as follows: the proposed articular surface and epiphysis datasets are described in Section 2, while Section 3 presents deep convolutional neural networks for bone age assessment on the proposed datasets. Section 4 shows the experiments and results analysis, and the conclusions are given in Section 5.

## 2. Articular surface and epiphysis datasets

Although the existing methods can predict bone age through general ROIs, they do not explore specific regions of interest, resulting in an inexplicable reference for clinical diagnosis. Inspired by it, the articular surface and epiphysis datasets, derived from the RSNA dataset, are proposed to investigate the effect of specific areas on bone age, where 1068 X-ray images with different ages (8–224 months) are randomly selected to form the new datasets. Specifically, there are 126, 365, 326, 223, and 28

images in 9–48, 48–96, 96–144, 144–192, and 192–228 months, respectively, as shown in Figure 2(a). These samples are distributed in different-group years, which guarantees that the proposed datasets are effective for bone age assessment in clinical practice. During the experiments, the proposed database is split into the training set, validation set, and test set with the numbers of 875, 96, and 97, respectively. Next, the articular surface and epiphysis datasets are described as follows.



**Figure 2.** Examples of the proposed datasets. (a) Data distribution. (b) Articular surface. (c) Epiphysis.

### 2.1. Articular surface dataset

As suggested from [32] that the epiphyseal region in the articular surface is often used by radiologists for bone age diagnosis. For example, the distal bone and joint bone are closed in a hand X-ray across the adolescent subjects [33]. Based on it, the articular surface of hand radiography is used as the feature region to assess the bone age. An example of the articular surface is shown in Figure 2(b), where the articular surface of the radius ulna, carpal, metacarpal-phalanges, phalanges, and epiphysis are marked as the feature regions. The specific implementation process is described as follows.

- During the screening process, we first sort the all samples in the RSNA dataset by age, and then randomly select images from different age groups to ensure that the proposed datasets satisfy the original distribution.
- In the labeling process, two radiologists use LabelMe4.6.0 [26] to independently mark bone feature regions on hand radiography according to [34], where the bone regions are determined by the shape, size, and color, and another expert reviews and aggregates it.
- Finally, the resulting feature area images are combined with the bone age label to establish the proposed articular surface dataset.

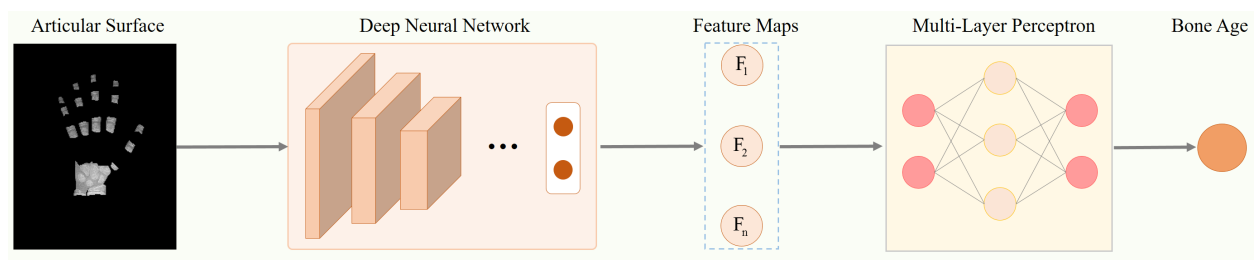
### 2.2. Epiphysis datasets

The size and shape of hand bone is also an important indicator to diagnose bone age [35]. For example, the size of the distal bone represents the development of the human epiphysis, which can

provide an effective basis for predicting the height of the human [36]. To this end, the epiphysis dataset is established for bone age assessment, where the detailed information is the same as the articular surface dataset except for the feature region. An example of the epiphysis is shown in Figure 2(c), where the radius ulna, carpal, metacarpal-phalanges, phalanges, and epiphysis are included in the new hand radiography. In addition, the creation process of the epiphysis dataset is the same as that of the articular surface dataset.

### 3. Deep convolutional neural networks for bone age assessment on articular surface and epiphysis datasets

The key regions of interest from hand radiography, i.e., articular surface and epiphysis, are often employed to perform bone age diagnosis by radiologists [8–10], but it is time-consuming and always leads to large errors between the predicted results and actual one. To solve this problem, the deep convolutional neural networks are exploited for bone age assessment on the proposed articular surface and epiphysis datasets, where five popular deep learning models, i.e., ResNet50 [27], SENet [28], DenseNet-121 [29], EfficientNet-b4 [30], and CSPNet [31], are used as the feature extractor, and then the multi-layer perceptron is exploited to produce the bone age, as shown in Figure 3, which are described as follows.



**Figure 3.** The framework of bone age assessment from hand radiography using deep neural networks, where the deep convolutional neural network is first regarded as the feature extractor to obtain the feature maps, and then the multi-layer perceptron is exploited to predict the final bone age.

#### 3.1. Bone age assessment using residual convolutional neural network–ResNet50

Residual convolutional neural network (ResNet), one of the most widely used deep learning models, can provide direct gradient feedback for network training by designing skip-connection. This can help the network provide a direct path to the very early layers in the network, making the gradient update of these layers easier. Based on it, the ResNet50 is used as a predictor for bone age assessment, defined as

$$\text{ResNet}(X) = \text{Conv}[\text{SC}(X) + X], \quad (3.1)$$

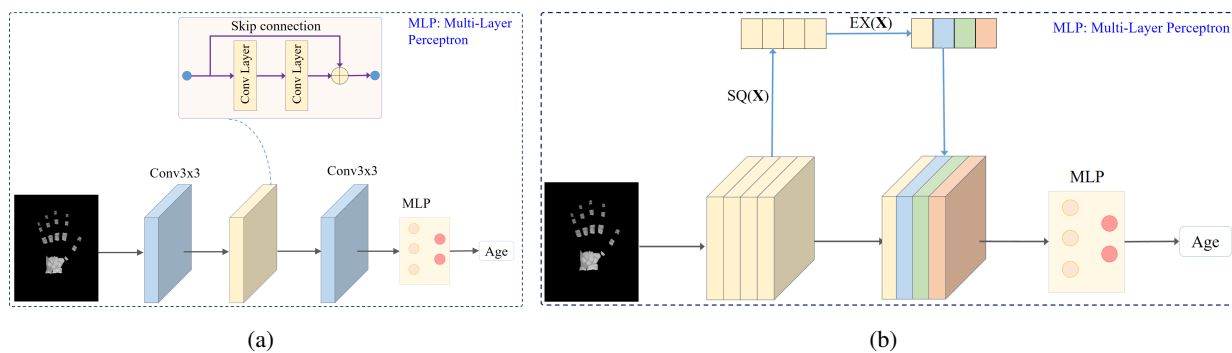
where SC and Conv denote the skip connection and convolutional layers, respectively, as shown in Figure 4(a). Specifically, the identified skip connections are embedded in the convolution layer to prevent network gradient degradation, and the last multi-layer perceptron can predict bone age by mapping the feature representation to the actual sample space.

### 3.2. Bone age assessment using attention convolutional neural network–SENet

Attention mechanism [37] has attracted extensive attention in medical image analysis, such as [38–40]. SENet [28], an effective attention mechanism, is designed to learn the details and model the global dependencies for improving the network performance. Inspired by it, the SENet is applied to obtain the bone age on the proposed dataset, defined as

$$\text{SENet}(X) = \text{EX}[\text{SQ}(X)], \quad (3.2)$$

where SQ and EX represent the Squeeze operation and Excitation operation, respectively, as shown in Figure 4(b). Specifically, the squeeze operation is utilized to compress the feature map into the feature vector, and the attention mechanism is employed to generate the attention weight matrix to obtain detailed information.



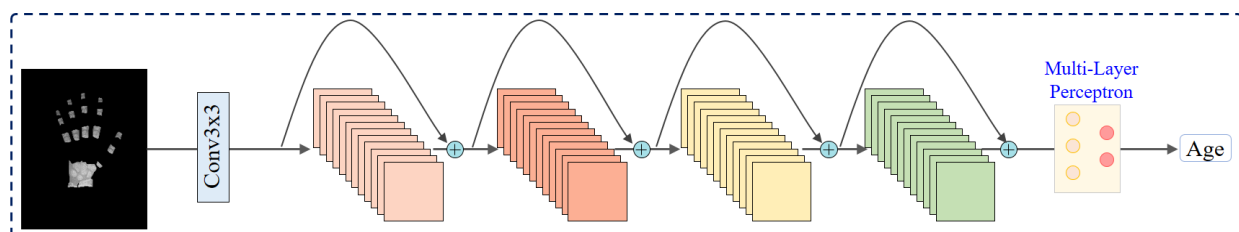
**Figure 4.** (a) Residual convolutional neural network (ResNet) for bone age assessment. (b) Squeeze and excitation network (SENet) for bone age assessment.

### 3.3. Bone age assessment using dense convolutional neural network–DenseNet

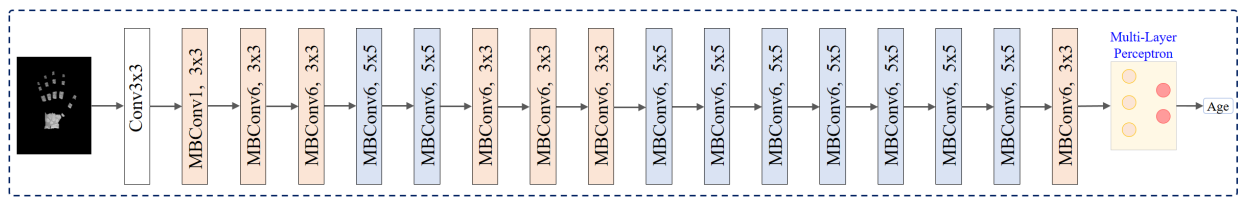
DenseNet is a convolutional neural network, which uses dense blocks to transfer the maximum information to each layer in the network, where each layer takes the information of all previous layers as input and then transmits the feature maps into all subsequent layers [29]. Similarly, the DenseNet is exploited to perform bone age assessment on the proposed dataset, defined as

$$\text{DenseNet}(X_l) = \text{F}[\text{SC}(X_0, \dots, X_{l-1}) + X_0, \dots, X_{l-1}], \quad (3.3)$$

where SC and F denote skip connection and convolution operation, respectively, as shown in Figure 5, and each layer takes the output of all previous layers as input and then transmits the feature maps into all subsequent layers.



**Figure 5.** Dense convolutional neural network (DenseNet) for bone age assessment.



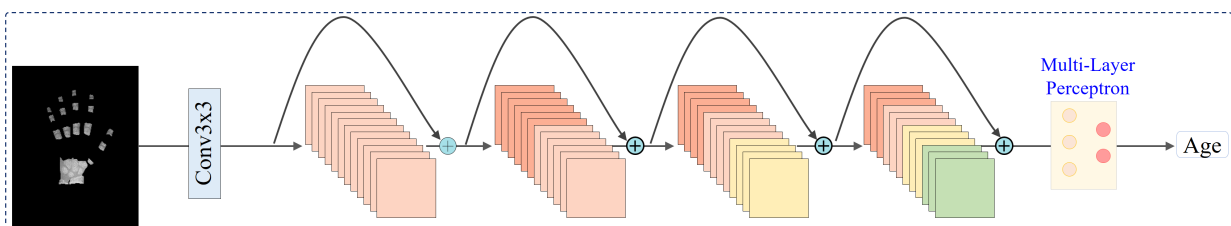
**Figure 6.** Efficient convolutional neural network (EfficientNet) for bone age assessment.

### 3.4. Bone age assessment using efficient convolutional neural network–EfficientNet

EfficientNet is an adaptive convolutional neural network, which optimizes the network by searching for the appropriate receptive field, depth, and resolution of convolutional layers [30], and it shows excellent performance on the computer vision tasks [41, 42]. Motivated by this, the EfficientNet is used as the backbone for bone age assessment on the proposed articular surface and epiphysis datasets, where the MBconv module, containing the skip connection and squeeze-excitation, is applied to generate the fine feature maps from the raw image.

### 3.5. Bone age assessment using cross stage partial network–CSPNet

The Cross Stage Partial Network (CSPNet) is designed to provide abundant gradient combination and reduce the calculation for the deep learning model by dividing gradient flow to different convolutional layers [31], where the feature maps are separated into a dense block and a transition layer, and then it will integrate with the transmitted feature map into the next layer. Thanks to its powerful excellent performance on computer vision [43], the CSPNet is employed to perform bone age on the proposed dataset, as shown in Figure 7. Specifically, the feature maps are separated into a dense block and a transition layer, and then the other one is integrated with transmitted feature maps into the next one.



**Figure 7.** Cross-stage partial network (CSPNet) based on DenseNet is applied for bone age assessment.

## 4. Experiments and results analysis

### 4.1. Evaluation metric

For all radiographs, the mean absolute error (MAE) and root mean square error (RMSE) between prediction and actual age are used as evaluation metrics [20], defined as

$$\text{MAE} = \frac{1}{N} \sum_{n=1}^N |\widehat{S}_f(n) - S_f(n)|, \quad (4.1)$$

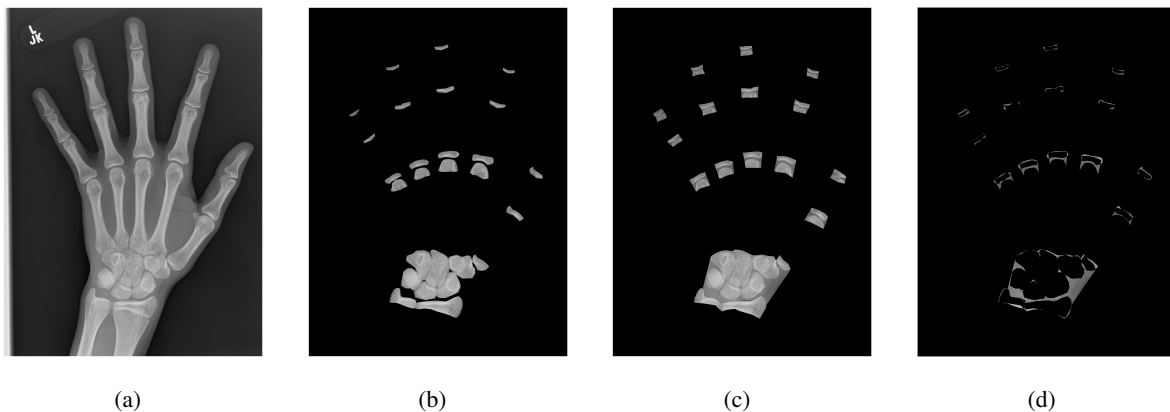
and

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{n=1}^N [\widehat{S}_f(n) - S_f(n)]^2}, \quad (4.2)$$

where  $S_f(n)$  represents the actual bone age, and  $\widehat{S}_f(n)$  represents the predicted bone age.

#### 4.2. Implementation details

These five deep learning models are conducted by Pytorch 1.8.0 [44] on a workstation with four NVIDIA GeForce GTX 1080Ti GPUs, where the SGD optimizer [45] with an initial learning rate of  $5e^{-5}$  is employed to train these networks, including ResNet50, SqueezeNet, DenseNet-121, EfficientNet-b4, and CSPNet, while the epoch, batch size, and momentum are set as 8, 100, and 0.9 respectively, the StepLR is used as the learning scheduler, and L1Loss is regarded as the loss function. In addition, all color images are scaled to  $576 \times 876$ , and they are normalized by subtracting the mean of 0.485, 0.456, and 0.406, and dividing the standard deviation of 0.229, 0.224, and 0.225, respectively at the three channels. Furthermore, the horizontal flip, vertical flip, and random rotation are applied to the proposed datasets for data augmentation.



**Figure 8.** Different hand bone feature regions. (a) Hand radiograph: the original hand radiograph. (b) Hand Epiphysis: the characteristic bone regions, including epiphysis, phalanges, metacarpal-phalanges, carpal, and radius-ulna. (c) Hand articular surface: these regions contain the characteristic bone and epiphysial regions. (d) Hand epiphyseal line: epiphysial regions between these characteristic bones.

#### 4.3. Performances on articular surface dataset

To validate the effectiveness of the proposed articular surface dataset, five networks are used for bone age assessment. The experimental results are shown in Table 1, and it can be seen that the CSPNet [31] can achieve the best performance for bone age assessment, with 7.34 months of the MAE and 9.75 months of the RMSE, respectively, demonstrating that the proposed articular surface dataset can be an effective database for bone age assessment using deep learning models. Furthermore, we also compare the proposed method with two automatic bone age assessment methods (i.e., SIMBA [46] and Chen et al. [47]) and two radiologists, and the results are listed in Table 2. It can be inferred that the



proposed method performs better and faster than automatic methods and manual diagnosis, where the average reading time of radiologists are 37.22s and 105.39s, respectively, while that of ours is 0.07s, showing that the proposed method is more accurate and fast than the radiologists.

In addition, the test results for each image are shown in Figure 9(a), and the test results of different age groups are plotted in Figure 9(c), where the first to fifth rows represent the results of the DenseNet-121 [29], ResNet50 [27], SENet [28], EfficientNet-b4 [30], and CSPNet [31], respectively. It can be noted that all models show a high sensitivity to these subjects between 9 and 192 months because the features are prone to explicitly discriminated on the radiograph. Besides, there are some slightly larger errors between 192–224 months generated by DenseNet, but the CSPNet [31] can achieve the MAE of 7.32 months, demonstrating the effectiveness of the proposed articular surface dataset.

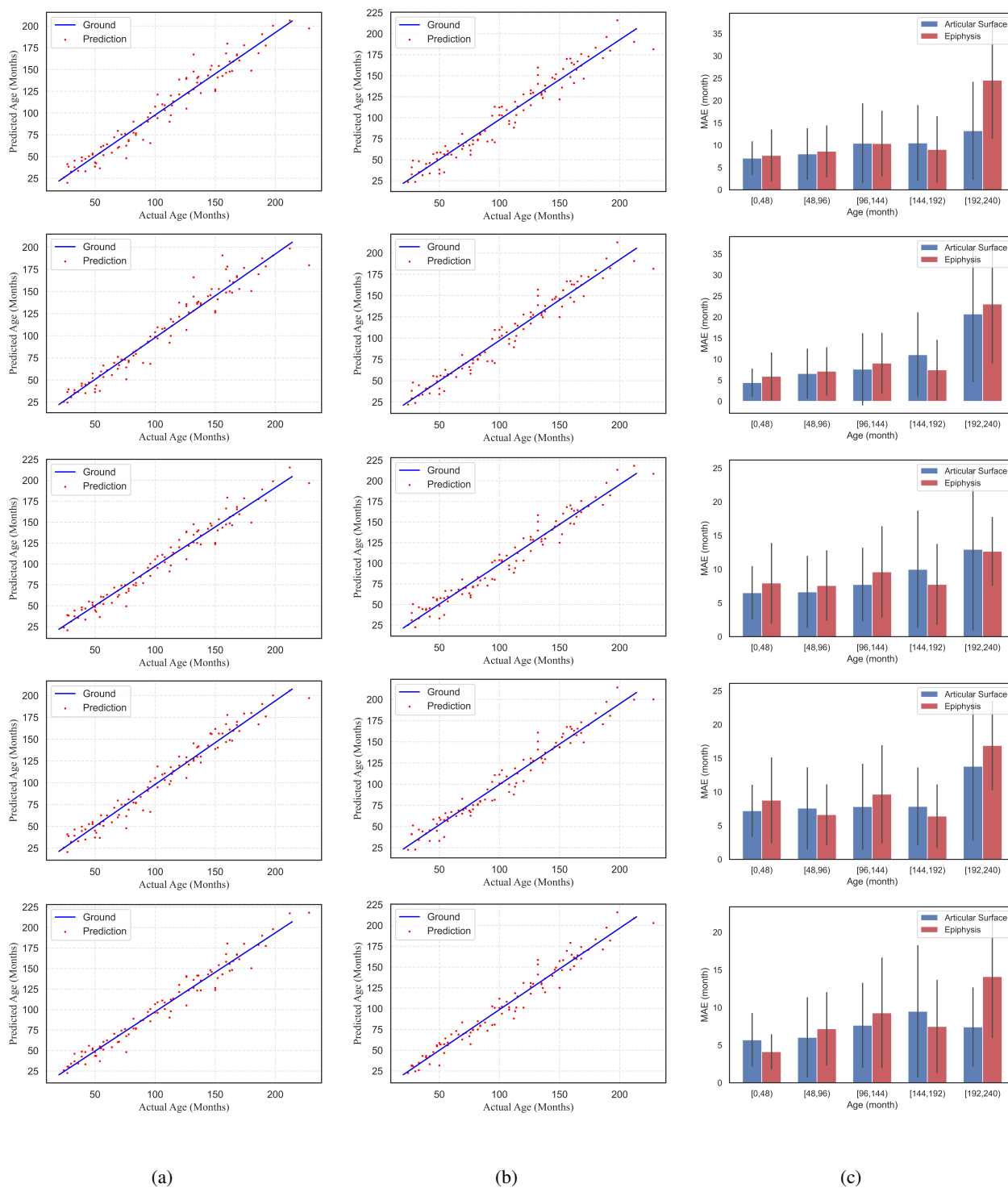
#### 4.4. Performances on epiphysis dataset

Similarly, these five popular deep-learning networks are applied for bone age assessment on the proposed epiphysis dataset. The results are given in Table 1, we can infer from it that the MAE and RMSE obtained by all models are less than 10 months, which proves that the epiphysis dataset is beneficial for bone age assessment. Moreover, the computer-aided methods of SIMBA [46] and Chen et al. [47], and two manual assessments, are used as the comparison method to validate the superiority of the proposed method on the epiphysis dataset, and the test results are given in Table 2. From it we can see that the proposed method achieves the best performance on the epiphysis dataset with the speed of 0.07 scan/s, demonstrating that it is prone to clinical practice.

**Table 1.** The quantitative comparisons on the articular surface and epiphysis datasets (mean±std).

Methods	Articular Surface		Epiphysis	
	MAE (month)	RMSE (month)	MAE (month)	RMSE (month)
DenseNet121 [28]	9.38 ± 0.37	12.13 ± 0.27	9.77 ± 0.31	12.49 ± 0.21
ResNet50 [27]	8.19 ± 0.29	12.07 ± 0.43	8.26 ± 0.18	11.32 ± 0.25
SENet [29]	7.95 ± 0.22	10.43 ± 0.34	8.32 ± 0.26	10.49 ± 0.22
EfficientNetB4 [30]	7.93 ± 0.29	10.11 ± 0.19	8.17 ± 0.17	10.29 ± 0.23
CSPNet [31]	7.34 ± 0.11	9.75 ± 0.16	7.73 ± 0.13	9.95 ± 0.17

Moreover, the specific test results conducted on the epiphysis dataset are shown in Figure 9, where the results of the DenseNet-121 [29], ResNet50 [27], SENet [28], EfficientNet-b4 [30], and CSPNet [31] are given in first to fifth rows, respectively. It can be observed from (b) in Figure 9 that most of them show comparable results across the ground truth, indicating their excellent performance for bone age assessment. Furthermore, it can be seen from (c) in Figure 9 that the CSPNet [31] can get the MAE of 13.2 months between 192 and 224 months, achieving improvements of 9.39 months compared with the DenseNet, which shows that the performance for bone age assessment on the epiphysis dataset can be advanced with the optimization of the deep learning models.



**Figure 9.** Statistical test results on the proposed datasets. (a) Articular surface dataset. (b) Epiphysis dataset. (c) Test results of different age groups on the articular surface and epiphysis datasets.

**Table 2.** Comparison to other methods (mean  $\pm$  std).

Methods	Articular Surface			Epiphysis		
	MAE (month)	RMSE (month)	Time (s)	MAE (month)	RMSE (month)	Time (s)
Radiologist1	13.41 $\pm$ 0.00	16.72 $\pm$ 0.00	37.22 $\pm$ 0.00	13.41 $\pm$ 0.00	16.72 $\pm$ 0.00	37.22 $\pm$ 0.00
Radiologist2	13.59 $\pm$ 0.00	17.34 $\pm$ 0.00	105.39 $\pm$ 0.00	13.59 $\pm$ 0.00	17.34 $\pm$ 0.00	105.39 $\pm$ 0.00
SIMBA [46]	9.28 $\pm$ 0.64	11.16 $\pm$ 0.94	0.06 $\pm$ 0.01	9.97 $\pm$ 0.86	13.36 $\pm$ 0.71	0.07 $\pm$ 0.01
Chen et al. [47]	8.37 $\pm$ 0.61	12.19 $\pm$ 0.71	0.07 $\pm$ 0.01	8.92 $\pm$ 0.59	12.23 $\pm$ 0.73	0.08 $\pm$ 0.01
Ours	7.34 $\pm$ 0.11	9.75 $\pm$ 0.16	0.07 $\pm$ 0.01	7.73 $\pm$ 0.13	9.95 $\pm$ 0.17	0.07 $\pm$ 0.01

#### 4.5. Ablation studies of skeletal feature regions

In clinical diagnosis, radiologists mainly assess bone age by comparing bone characteristic areas in hand radiographs with a reference. In order to validate the impact of these bone characteristic areas, a number of ablation experiments are conducted on the feature regions, including the hand articular surface (HAS) region, hand epiphysis (HE) region, and hand epiphyseal line (HEL) region, as shown in Figure 8.

##### 4.5.1. Impact of hand articular surface region

In medical imaging, radiologists often make bone age diagnoses by comparing hand articular surface regions. On the other hand, the CSPNet [31] shows excellent performance on the proposed datasets. Inspired by this, we explore bone age assessment in hand articular surface region using CSPNet [31]. The results are shown in Table 3. It can be seen that the CSPNet [31] can achieve the MAE of 7.34 months and the RMSE of 9.35 months in the hand bone region, which shows that the hand articular surface region can be used for bone age assessment in clinical practice.

**Table 3.** Ablation studies of different feature regions for bone age assessment.

No.	HAS	HE	HEL	MAE (month)	RMSE (month)
1	✓			7.34 $\pm$ 0.27	9.75 $\pm$ 0.21
2		✓		7.71 $\pm$ 0.24	9.95 $\pm$ 0.18
3			✓	12.28 $\pm$ 0.29	16.11 $\pm$ 0.19

##### 4.5.2. Impact of hand epiphysis region

The epiphysis is one of the important indicators for evaluating the growth and development of children. In general, the epiphyseal area has a linear relationship with age, and it can be used to assess bone age clinically [48]. Motivated by this, the bone age assessment is conducted on the hand epiphysis region using CSPNet [31] for bone features extracting. The results are given in Table 3. It can be observed that the CSPNet [31] can achieve the MAE of 7.71 months and the RMSE of 9.95 months in the hand epiphysis region, demonstrating that it is feasible to use a neural network to evaluate bone age on the hand epiphysis region.

### 4.5.3. Impact of hand epiphyseal line region

Inspired by the above two feature regions, the hand epiphyseal line region is employed for bone age assessment using CSPNet [31]. To evaluate its effectiveness, we conduct a large number of experiments on the hand epiphyseal line region, and the results are listed in Table 3. It can be inferred that the CSPNet [31] shows the best performance in the hand epiphyseal line region, achieving the MAE of 12.28 months and the RMSE of 16.11 months, indicating that bone age assessment in the hand epiphyseal line region is helpful for clinical practice.

In a word, it is considered to evaluate bone age in the hand articular surface and epiphysis region, but the performance on the articular surface is better than that of epiphysis using CSPNet [31]. Therefore, we recommend the hand articular surface region as the benchmark for bone age assessment using deep learning models.

### 4.6. Limitations and future works

As shown in the experiments, five deep convolutional neural networks are employed for bone age prediction on the proposed articular surface and epiphysis datasets, which are faster and more accurate than the radiologists. However, there are still some limitations to the proposed method. First of all, the annotations would be firstly required by radiologists, and then the bone age can be obtained from the convolutional neural networks. Second, the proposed articular surface and epiphysis datasets in this study have only 1,068 images, which is limited to the original RSNA dataset, which consists of 14,236 images. Third, the proposed datasets mainly focus on the articular surface and epiphysis of hand radiology, where some interest of regions, e.g., phalanx, are missed, which will help to improve the performance of bone age assessment.

In future works, three key measures will be conducted as follows. First, a segmentation network would be required as a pre-processing step to make the proposed method really usable. Then, more work will be performed to expand the proposed articular surface and epiphysis datasets, so as to it is beneficial to clinical practice. Finally, more characteristic regions of interest will be incorporated into our research work to improve the performance of bone age diagnosis.

## 5. Conclusions

In this paper, deep convolutional neural networks are employed for bone age prediction on the proposed articular surface and epiphysis datasets, where the convolutional neural networks are exploited to improve the accuracy and efficiency of bone age assessment. Extensive experiments on the proposed datasets demonstrate their effectiveness for bone age assessment using deep learning models, further assisting radiologists in clinical diagnosis.

### Conflicts of interest

No potential conflict of interest.

## Acknowledgements

This work was supported in part by the National Natural Science Foundation of China under Grant 51975129, Grant 6197020346, and Grant 71802055.

## References

1. P. Hao, S. Chokuwa, X. Xie, F. Wu, J. Wu, C. Bai, Skeletal bone age assessments for young children based on regression convolutional neural networks, *Math. Biosci. Eng.*, **16** (2019), 6454–6466.
2. W. Lin, F. Yang, Computational analysis of cutting parameters based on gradient voronoi model of cancellous bone, *Math. Biosci. Eng.*, **19** (2022), 11657–11674. <https://doi.org/10.3934/mbe.2022542>
3. S. L. Truesdell, M. M. Saunders, Bone remodeling platforms: Understanding the need for multicellular lab-on-a-chip systems and predictive agent-based models, *Math. Biosci. Eng.*, **17** (2020).
4. A. Nordenstrom, H. Falhammar, Management of endocrine disease: Diagnosis and management of the patient with non-classic CAH due to 21-hydroxylase deficiency, *Eur. J. Endocrinol.*, **180** (2019), R127–R145.
5. A. C. Morani, C. T. Jensen, M. A. Habra, M. M. Agrons, C. O. Menias, N. A. Wagner-Bartak, et al., Adrenocortical hyperplasia: a review of clinical presentation and imaging, *Abdominal Radiol.*, **45** (2020), 917–927. <https://doi.org/10.1007/s00261-019-02048-6>
6. J.-C. Carel, J. Léger, Precocious puberty, *New England J. Med.*, **358** (2008), 2366–2377. <https://doi.org/10.1056/NEJMcp0800459>
7. K. Kyostila, J. E. Niskanen, M. Arumilli, J. Donner, M. K. Hytönen, H. Lohi, Intronic variant in POU1F1 associated with canine pituitary dwarfism, *Human Gene.*, **140** (2021), 1553–1562. <https://doi.org/10.1007/s00439-021-02259-2>
8. K. C. Lee, K. H. Lee, C. H. Kang, K. S. Ahn, L. Y. Chung, J. J. Lee, et al., Clinical validation of a deep learning-based hybrid (Greulich-Pyle and modified Tanner-Whitehouse) method for bone age assessment, *Korean J. Radiol.*, **22** (2021), 2017–2025. <https://doi.org/10.3348/kjr.2020.1468>
9. P. Lv, C. Zhang, Tanner–Whitehouse skeletal maturity score derived from ultrasound images to evaluate bone age, *Eur. Radiol.*, **2022** (2022), 1–8.
10. S. Zhang, L. Liu, The skeletal development standards of hand and wrist for Chinese ChildrenChina 05 I. TW\_3-C RUS, TW\_3-C Carpal, and RUS-CHN methods, *Chin. J. Sports Med.*, **2023** (2023), 6–13.
11. J. R. Kim, Y. S. Lee, J. Yu, Assessment of bone age in prepubertal healthy Korean children: comparison among the Korean standard bone age chart, Greulich-Pyle method, and Tanner-Whitehouse method, *Korean J. Radiol.*, **16** (2015), 201–205. <https://doi.org/10.3348/kjr.2015.16.1.201>

12. C. Gao, Q. Qian, Y. Li, X. Xing, X. He, M. Lin, et al., A comparative study of three bone age assessment methods on Chinese preschool-aged children, *Front. Pediatr.*, **10** (2022), 1–12. <https://doi.org/10.3389/fped.2022.976565>
13. H. Liu, M. Liu, D. Li, W. Zheng, L. Yin, R. Wang, Recent advances in pulse-coupled neural networks with applications in image processing, *Electronics*, **11** (2022), 3264. <https://doi.org/10.3390/electronics11203264>
14. Y. Ban, Y. Wang, S. Liu, B. Yang, M. Liu, L. Yin, et al., 2d/3d multimode medical image alignment based on spatial histograms, *Appl. Sci.*, **12** (2022), 8261. <https://doi.org/10.3390/app12168261>
15. S. Xiong, B. Li, S. Zhu, DCGNN: A single-stage 3D object detection network based on density clustering and graph neural network, *Complex Intell. Syst.*, **2022** (2022), 1–10.
16. C. Spampinato, S. Palazzo, D. Giordano, M. Aldinucci, R. Leonardi, Deep learning for automated skeletal bone age assessment in X-ray images, *Med. Image Anal.*, **36** (2017), 41–51. <https://doi.org/10.1016/j.media.2016.10.010>
17. D. B. Larson, M. C. Chen, M. P. Lungren, S. S. Halabi, N. V. Stence, C. P. Langlotz, Performance of a deep-learning neural network model in assessing skeletal maturity on pediatric hand radiographs, *Radiology*, **287** (2018), 313–322. <https://doi.org/10.1148/radiol.2017170236>
18. Y. Liu, C. Zhang, J. Cheng, X. Chen, Z. J. Wang, A multi-scale data fusion framework for bone age assessment with convolutional neural networks, *Comput. Biol. Med.*, **108** (2019), 161–173. <https://doi.org/10.1016/j.compbiomed.2019.03.015>
19. Q. H. Nguyen, B. P. Nguyen, M. T. Nguyen, M. C. Chua, T. T. Do, N. Nghiem, Bone age assessment and sex determination using transfer learning, *Expert Syst. Appl.*, **200** (2022), 116926. <https://doi.org/10.1016/j.eswa.2022.116926>
20. X. Ren, T. Li, X. Yang, S. Wang, S. Ahmad, L. Xiang, et al., Regression convolutional neural network for automated pediatric bone age assessment from hand radiograph, *IEEE J. Biomed. Health Inf.*, **23** (2018), 2030–2038. <https://doi.org/10.1109/JBHI.2018.2876916>
21. C. Liu, H. Xie, Y. Zhang, Self-supervised attention mechanism for pediatric bone age assessment with efficient weak annotation, *IEEE Trans. Med. Imaging*, **40** (2020), 2685–2697. <https://doi.org/10.1109/TMI.2020.3046672>
22. N. E. Reddy, J. C. Rayan, A. V. Annapragada, N. F. Mahmood, A. E. Scheslinger, W. Zhang, et al., Bone age determination using only the index finger: a novel approach using a convolutional neural network compared with human radiologists, *Pediatr. Radiol.*, **50** (2020), 516–523. <https://doi.org/10.1007/s00247-019-04587-y>
23. S. Mutasa, P. D. Chang, C. Ruzal-Shapiro, R. Ayyala, MABAL: A novel deep-learning architecture for machine-assisted bone age labeling, *J. Dig. Imaging*, **31** (2018), 513–519. <https://doi.org/10.1007/s10278-018-0053-3>
24. X. Li, Y. Jiang, Y. Liu, J. Zhang, S. Yin, H. Luo, RAGCN: Region aggregation graph convolutional network for bone age assessment from X-ray images, *IEEE Trans. Instru. Measure.*, **71** (2022), 1–12.

25. S. S. Halabi, L. M. Prevedello, J. Kalpathy-Cramer, A. B. Mamonov, A. Bilbily, M. Cicero, et al., The RSNA pediatric bone age machine learning challenge, *Radiology*, **290** (2019), 498–503. <https://doi.org/10.1148/radiol.2018180736>
26. B. C. Russell, A. Torralba, K. P. Murphy, W. T. Freeman, LabelMe: a database and web-based tool for image annotation, *Int. J. Comput. Vision*, **77** (2008), 157–173. <https://doi.org/10.1007/s11263-007-0090-8>
27. K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, 770–778.
28. J. Hu, L. Shen, G. Sun, Squeeze-and-excitation networks, in *Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR)*, 2018, 7132–7141.
29. G. Huang, Z. Liu, L. Van Der Maaten, K. Q. Weinberger, Densely connected convolutional networks, in *Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR)*, 2017, 4700–4708.
30. M. Tan, Q. Le, Efficientnet: Rethinking model scaling for convolutional neural networks, in *Proceedings of the International conference on machine learning (ICLR)*, 2019, 6105–6114.
31. C. Y. Wang, H. Y. M. Liao, Y. H. Wu, P. Y. Chen, J. W. Hsieh, I. H. Yeh, CSPNet: A new backbone that can enhance learning capability of CNN, in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops (CVPR)*, 2020, 390–391.
32. D. Giordano, C. Spampinato, G. Scarciofalo, R. Leonardi, An automatic system for skeletal bone age measurement by robust processing of carpal and epiphysal/metaphysal bones, *IEEE Trans. Instru. Measure.*, **59** (2010), 2539–2553. <https://doi.org/10.1109/TIM.2010.2058210>
33. D. Knapik, J. Sanders, A. Gilmore, D. Weber, D. Cooperman, R. Liu, A quantitative method for the radiological assessment of skeletal maturity using the distal femur, *Bone Joint J.*, **100** (2018), 1106–1111.
34. A. K. Bhat, A. M. Acharya, M. Pai, Radiology of the wrist and hand, in *Clinical Examination of the Hand*, CRC Press, (2022), 275–299.
35. K. S. Ahn, B. Bae, W. Y. Jang, J. H. Lee, J. H. Lee, S. Oh, B. H. Kim, et al., Assessment of rapidly advancing bone age during puberty on elbow radiographs using a deep neural network model, *Eur. Radiol.*, **31** (2021), 8947–8955. <https://doi.org/10.1007/s00330-021-08096-1>
36. U. Nemeč, S. F. Nemeč, M. Weber, et al., Human long bone development in vivo: analysis of the distal femoral epimetaphysis on MR images of fetuses, *Radiology*, **267** (2013), 570–580. <https://doi.org/10.1148/radiol.13112441>
37. A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, et al., Attention is all you need, in *Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS)*, 2017, 5998–6008.
38. Z. Q. Zhao, P. Zheng, S. Xu, X. Wu, Object detection with deep learning: A review, *IEEE Trans. Neural Networks Learn. syst.*, **30** (2019), 3212–3232. <https://doi.org/10.1109/TNNLS.2018.2876865>

39. S. Ren, K. He, R. Girshick, J. Sun, Faster R-CNN: Towards real-time object detection with region proposal networks, *IEEE Trans. Pattern Anal. Mach. Intell.*, **39** (2017), 1137–1149. <https://doi.org/10.1109/TPAMI.2016.2577031>
40. A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, et al., An image is worth 16x16 words: Transformers for image recognition at scale, in *Proceedings of the International Conference on Learning Representations (ICLR)*, 2020, 1–21.
41. M. Roberts, D. Driggs, M. Thorpe, J. Gilbey, M. Yeung, S. Ursprung, et al., Common pitfalls and recommendations for using machine learning to detect and prognosticate for COVID-19 using chest radiographs and CT scans, *Nat. Mach. Intell.*, **3** (2021), 199–217.
42. Z. Liu, Y. Lin, Y. Cao, H. Hu, Y. Wei, Z. Zhang, et al., Swin transformer: Hierarchical vision transformer using shifted windows, in *Proceedings of the IEEE/CVF international conference on computer vision (CVPR)*, **2021** (2021), 10012–10022.
43. C. Y. Wang, A. Bochkovskiy, H. Y. M. Liao, Scaled-yolov4: Scaling cross stage partial network, in *Proceedings of the IEEE/cvf conference on computer vision and pattern recognition (CVPR)*, (2021), 13029–13038.
44. A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, et al., Pytorch: An imperative style, high-performance deep learning library, in *Proceedings of the Advances in Neural Information Processing Systems (NIPS)*, (2019), 8026–8037.
45. L. Bottou, F. E. Curtis, J. Nocedal, Optimization methods for large-scale machine learning, *Siam Rev.*, **60** (2018), 223–311. <https://doi.org/10.1137/16M1080173>
46. C. Gonzalez, M. Escobar, L. Daza, F. Torres, G. Triana, P. Arbelaez, SIMBA: Specific identity markers for bone age assessment, in *Proceeding of Medical Image Computing and Computer Assisted Intervention (MICCAI)*, (2020), 753–763.
47. C. Chen, Z. Chen, X. Jin, L. Li, W. Speier, C. W. Arnold, Attention-guided discriminative region localization and label distribution learning for bone age assessment, *IEEE J. Biomed. Health Inf.*, **26** (2021), 1208–1218.
48. B. D. Lee, M. S. Lee, Automated bone age assessment using artificial intelligence: the future of bone age assessment, *Korean J. Radiol.*, **22** (2021), 792–800. <https://doi.org/10.3348/kjr.2020.0941>



AIMS Press

©2023 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>)