

Article

A Cascade Model with Prior Knowledge for Bone Age Assessment

Nanxin Li ^{1,†} , Bochao Cheng ^{2,†}  and Junran Zhang ^{1,*,†}

¹ College of Electrical Engineering, Sichuan University, Chengdu 610065, China; linanxin@stu.scu.edu.cn
² West China Second University Hospital, Sichuan University, Chengdu 610041, China; wonder9527@163.com
* Correspondence: zhangjunran@126.com; Tel.: +86-028-8540-5614
† These authors contributed equally to this work.

Abstract: Bone age is commonly used to reflect growth and development trends in children, predict adult heights, and diagnose endocrine disorders. Nevertheless, the existing automated bone age assessment (BAA) models do not consider the nonlinearity and continuity of hand bone development simultaneously. In addition, most existing BAA models are based on datasets from European and American children and may not be applicable to the developmental characteristics of Chinese children. Thus, this work proposes a cascade model that fuses prior knowledge. Specifically, a novel bone age representation is defined, which incorporates nonlinear and continuous features of skeletal development and is implemented by a cascade model. Moreover, corresponding regions of interest (RoIs) based on RUS-CHN were extracted by YOLO v5 as prior knowledge inputs to the model. In addition, based on MobileNet v2, an improved feature extractor was proposed by introducing the Convolutional Block Attention Module and increasing the receptive field to improve the accuracy of the evaluation. The experimental results show that the mean absolute error (MAE) is 4.44 months and significant correlations with the reference bone age is ($r = 0.994, p < 0.01$); accuracy is 94.04% for ground truth within ± 1 year. Overall, the model design adequately considers hand bone development features and has high accuracy and consistency, and it also has some applicability on public datasets, showing potential for practical and clinical applications.

Keywords: multi-point distribution label; cascade model; prior knowledge; bone age assessment; deep learning



Citation: Li, N.; Cheng, B.; Zhang, J. A Cascade Model with Prior Knowledge for Bone Age Assessment. *Appl. Sci.* **2022**, *12*, 7371. <https://doi.org/10.3390/app12157371>

Academic Editor: Rossella Bedini

Received: 5 May 2022

Accepted: 6 July 2022

Published: 22 July 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Bone age is one of the independent indicators used to measure the biological maturity of the human body [1,2]. It is often used clinically to diagnose a series of endocrine diseases such as short stature, obesity, and metabolic disorders [3,4]. Physicians can make accurate judgments about developmental abnormalities in children based on the difference between their bone age and chronological age.

Currently, bone age is mainly assessed by an experienced physician utilizing certain assessment methods based on the X-rays of a child's infrequently used hand (usually the left hand). The most common assessment methods are the Greulich–Pyle (GP) method, proposed in 1999 [5], and the Tanner–Whitehouse (TW3) method, proposed in 2002 [6]. Since the development trends of the hand bones between eastern and western children are different due to the influences of race, nutrition, genetics, and other factors [7–9], the methods mentioned above, which use a sample of western children, might be inaccurate in Asian children. Therefore, Chinese 05 assessment criteria were developed, which used samples of contemporary Chinese children [10]. This criterion incorporates TW3-C RUS, TW3-C Carpal, and RUS-CHN methods. Compared to TW3, Chinese 05 assessment criteria have added more areas of RoIs and established more detailed rating rules. Since Chinese 05 criteria were proposed, a multitude of studies have evaluated their feasibility and have

shown that the RUS-CHN method is more suitable for assessing the bone ages of Chinese children [11].

Automatic bone age assessment (BAA) has been one of the research focuses in radiology and other related fields due to the time-consuming and subjective factors of manual assessment, which lead to inconsistent results between the two assessments. For example, in the field of intelligence-assisted diagnosis and treatment, Toddberg et al. proposed the first commercially available automated BAA system, called BoneXpert [12]. Although this method shortens assessment time, it does not use the carpal bone, which is necessary for infant age assessment [13]. In addition, the above method also requires the input images with extremely high quality, which poses an obstacle to its wide application.

Many machine learning (ML) methods are often applied to BAA in previous studies. For example, Tong et al. [14] used Multiple Kernel Learning (MKL) to incorporate heterogeneous features of the dataset into the network and then used Support Vector Regression (SVR) for classification, resulting in a mean absolute error (MAE) of 0.428 years. Pietkaet et al. [15] performed a series of image preprocessing such as background denoising, RoIs extraction, and distance length determination for hand bone images and developed BAA of semantic features. In addition, many researchers have solved BAA tasks using the decision tree method, K-Nearest-Neighbors (KNN) classification [16], and Support Vector Machine (SVM) [17]; however, the above methods utilize manual prior knowledge extraction, which is time-consuming and less accurate.

With the development of deep learning (DL), DL models have been gradually applied to the medical field and have achieved certain goals. Compared with traditional manual methods that extract features using ML methods, DL models can automatically extract features quickly; thus, a multitude of BAA models based on DL have developed. For example, Spampinato et al. [18] proposed the first BAA model based on convolutional neural networks (CNNs), called BoNet. Son S. J. et al. [19] proposed a BAA model based on prior knowledge provided by the TW method. The RoIs of the TW method were extracted as prior knowledge, and then prior knowledge was input into the classification network to obtain bone age prediction values. In addition, some researchers have used DL models instead of ML methods for background denoising, improving accuracy while eliminating manual operations. Among them, Salem et al. [20] used the Visual Geometry Group network (VGG) and the mask region-based convolutional neural network (Mask R-CNN) to extract hand regions for background denoising. Ren et al. [21] proposed a multichannel evaluation model that preprocesses images to extract coarse and fine attention maps with a multichannel input regression network. Han et al. [22] extracted 17 RoIs of the hand using Residual Network (ResNet), and the extracted RoIs were input into the Spatial Transformer Network (STN) for bone age assessment. Overall, using a deep learning model to solve the BAA problem can improve the accuracy of the evaluation process, shorten the evaluation time, and avoid the inconsistency between evaluation results caused by subjective factors. However, the above model is based on children from the same countries/regions. Due to the different characteristics of hand bone development in different countries/regions, models need to be experimented on different regional datasets to demonstrate the generalisability of the model.

Most DL models treat BAA as a regression or multiclass classification problem at present, and most of the models were built based on data from populations different from Chinese, which may be less applicable to contemporary Chinese children. Clinically, the development of hand bones exhibits the characteristics of continuity and nonlinearity [13]. A regression network excessively simplifies the entire developmental process and does not consider the nonlinearity caused by the different developmental rates experienced by hand bones at each stage, rendering the network vulnerable to the influence of outliers and prone to cause overfitting in training [23]. A classification network ignores the continuity caused by the similarity among the forms of hand bones at adjacent ages; however, the use of ordered labels allows the model to achieve a good assessment [24]. In order to improve

the above-mentioned shortcomings and limitations, this paper proposes a cascade model fusing of prior knowledge for BAA. The main contributions of this paper are as follows:

- The bone age label is redefined by transforming it into a multi-point distribution label with semantic information. The proposed label, which considers the nonlinear and continuous characteristics, is embedded in the network through a cascade model;
- Base on a data sample of contemporary Chinese adolescents, the RoIs of the RUS-CHN method is applied to the model as prior knowledge to improve accuracy;
- A new feature extractor is proposed based on MobileNet v2, which can capture more detailed features of hand bone;
- The network is trained with an improved Focal loss function to prevent model’s overfitting and to address the impact of uneven data distribution on the network.

This paper proposes a new representation of bone age labels and embeds it into the network through a cascade model to address the problem that the existing models did not consider nonlinearity and continuity simultaneously. In addition, the model fuses RoIs of RUS-CHN as prior knowledge, improving the adaptability of the model to children in the Chinese region.

The rest of this paper is organized as follows. Section 2 presents details of our proposed model and the dataset. Section 3 demonstrates the model’s assessment results. Section 4 provides relevant discussion and analysis of the strengths and weaknesses of our model, followed by conclusions in Section 5.

2. Materials and Methods

As shown in Figure 1, our model consists of two phases, i.e., RoIs extraction and BAA, and the details of the two phases will be elaborated as follows.

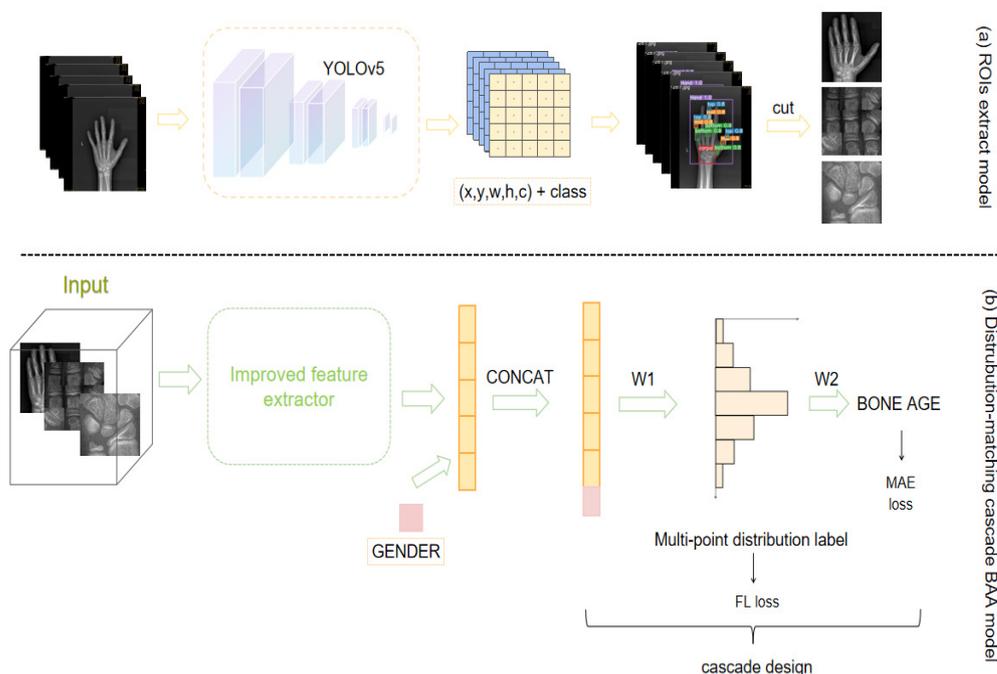


Figure 1. Our method consists of two stages. In the first stage, an object detector is trained to locate the coordinates of RoIs for the RUS-CHN method and image after background denoising, and the localized coordinates are used to crop high-resolution RoIs from the original image. All RoIs is aggregated and used to train a cascade module in the second stage.

2.1. Dataset

As there is no publicly available database based on the RUS-CHN method, our dataset was constructed in this paper. From January 2020 to June 2022, the left hand X-ray images of 753 subjects (382 males and 371 females) aging from 1 to 18 years who went to the clinic

of West China women’s and children’s Hospital for health screening, height prediction, etc., were collected, and those of developmental anomalies and endocrine dysfunction disease were excluded. Additionally, the ages of the hand bones shown in the X-ray images were assessed based on the RUS-CHN method by two experienced radiologists. the dataset mentioned above is divided randomly into two parts, namely, a training set and a test set, at a ratio of 8:2 during the bone age assessment stage (i.e., the second stage). The bone age and gender distributions of the entire dataset are shown in Figure 2.

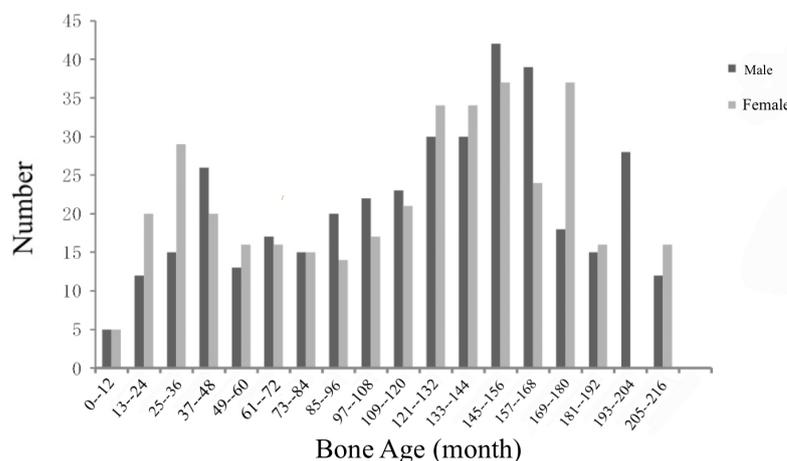


Figure 2. Statistical chart of the distribution of men and women by age in the dataset.

2.2. RoIs Extraction

For the model to focus on the area of interest for physician assessment, the RoIs of the RUS-CHN method were fused into the BAA model as prior knowledge to learn more significant hand bone features. At the same time, considering the detrimental effect of background noise on network learning, the entire hand region was extracted to reduce interference from the background. The total number of RoIs to be extracted is shown in the red box part of Figure 3.

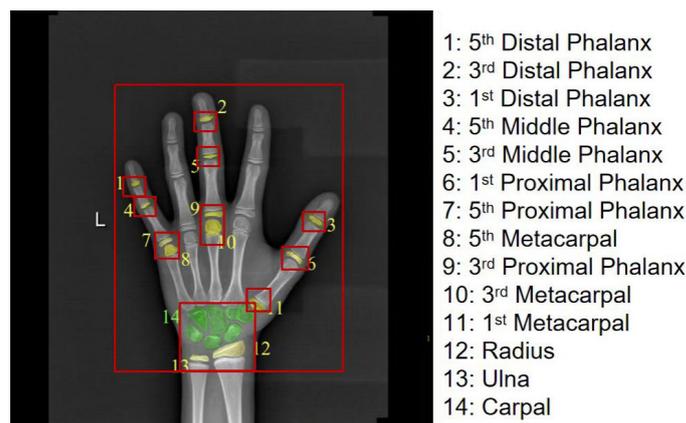


Figure 3. The markers 1–14 in the image are the RoIs of the RUS-CHN method (including the metacarpals and phalanges, i.e., the yellow part in Figure 1, and the carpal bones, i.e., the green part in Figure 1). The red boxes in the images indicate all RoIs to be extracted, including the hand bone image after the background denoising and the RoIs of the RUS-CHN method.

This paper applies the You Only Look Once version 5 (YOLOv5) [25] algorithm to extract RoIs. The workflow of this task is shown in Figure 4.

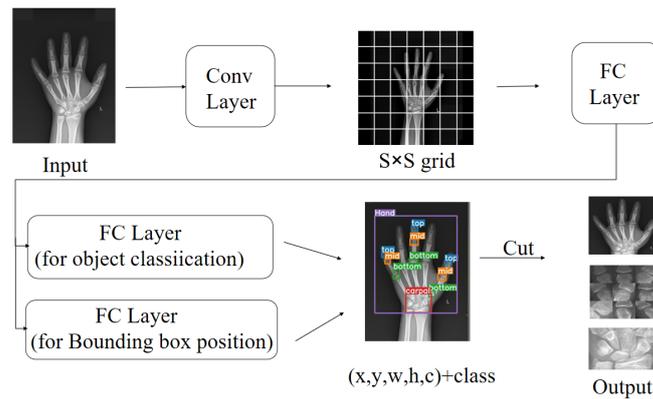


Figure 4. Workflow of faster R-CNN (Conv Layer: convolutional layers; FC Layer: fully connected layers).

First, the network divides the input image into $S \times S$ grids, and each grid is responsible for detecting objects for which their centers fall in the grid and then outputs the coordinates of the object boxes and the predicted class of objects in the object boxes. The probability formula is described as follows:

$$P_r(class_i|object) * P_r(object) * IoU_{pre}^{truth} = P_r(class_i) * IoU_{pre}^{truth} \tag{1}$$

where $P_r(object)$ is the probability that the bounding box contains an object, IoU_{pre}^{truth} (intersection over union) is the intersection ratio between the predicted box and the ground truth, $P_r(class_i|object)$ is the conditional probability at the confidence level of each bounding box, and $P_r(class_i)$ is the probability to which the class of the object box belongs.

According to the object box coordinates positioned by YOLOv5, high-quality RoIs images are cut out from the original images, and the process of extracting hand bone RoIs in the middle is shown in Figure 5.

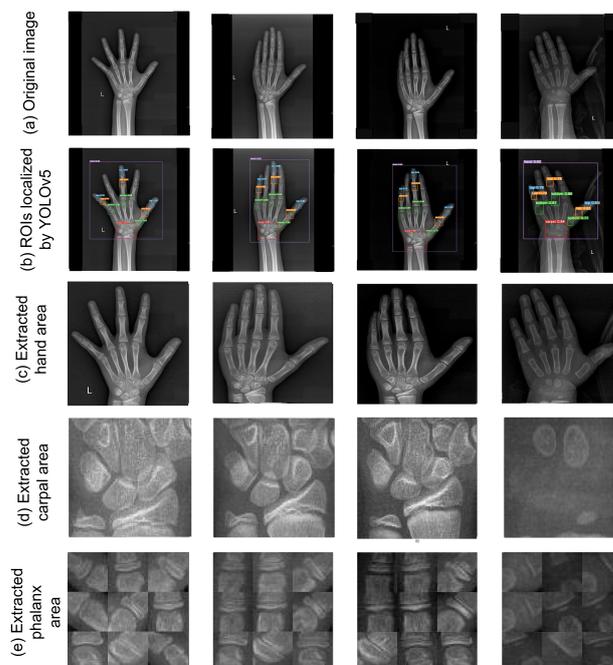


Figure 5. The change process of the RoIs extraction module in the first stage of image input. (a) is the original image. (b) is a picture of RoIs localized by YOLOv5. (c–e) are the RoIs outputted by YOLOv5 of 3 parts, i.e., hand (hand bone image by background denoising), carpal bone, and phalanx bone (RoIs of RUS-CHN method).

2.3. Bone Age Assessment

This paper proposes a cascade model based on MobileNet v2 [26]. The model’s overall structure is shown in Figure 1b, and the specific improvement method and the construction process of the model are described in the following.

2.3.1. Improved Feature Extractor Structure

In order to maximize the model’s ability to extract hand bone features, we optimized the feature extractor in MobileNet v2.

Firstly, expanding the size of the depth-separable convolution kernel from 3×3 to 5×5 expands the receptive field (RF) of the feature extractor. If the RF is larger, more features can be obtained [27]. The relationship between the size of the convolution kernel and RF is as follows:

$$RF_k = RF_{k-1} + (f_k - 1) * \prod_{i=0}^{k-1} s_i \quad n \geq 2 \tag{2}$$

where RF_{k-1} denotes the size of the RF corresponding to the k -th layer, f_k is the k -th layer convolution kernel size, and s denotes the step size of the convolution operation.

Secondly, a convolutional block attention module (CBAM) [28] is added to the feature extractor to adjust the spatial and channel weights separately. CBAM consists of a channel attention module (CAM) and a spatial attention module (SAM) and is calculated as follows.

$$\begin{aligned} M_c(F) &= \sigma(MLP(AvgPool(F)) + MLP(MaxPool(F))) \\ M_s(F) &= \sigma(f^{7 \times 7}([AvgPool(F); MLP(MaxPool(F))])) \end{aligned} \tag{3}$$

Here, F denotes the input feature map, $M_c(F)$ is CAM features, $M_s(F)$ denotes SAM features, $AvgPool$ and $MaxPool$ denote the average pooling and maximum pooling, σ denotes the sigmoid activation function, and MLP denotes the multilayer perceptron structure.

The baseline module of the improved feature extractor is shown in Figure 6, which is used to replace the depth separable convolutional block of MobileNet v2, thus enhancing the feature extraction capability of the model and permitting the extraction of more detailed features of the hand bone.

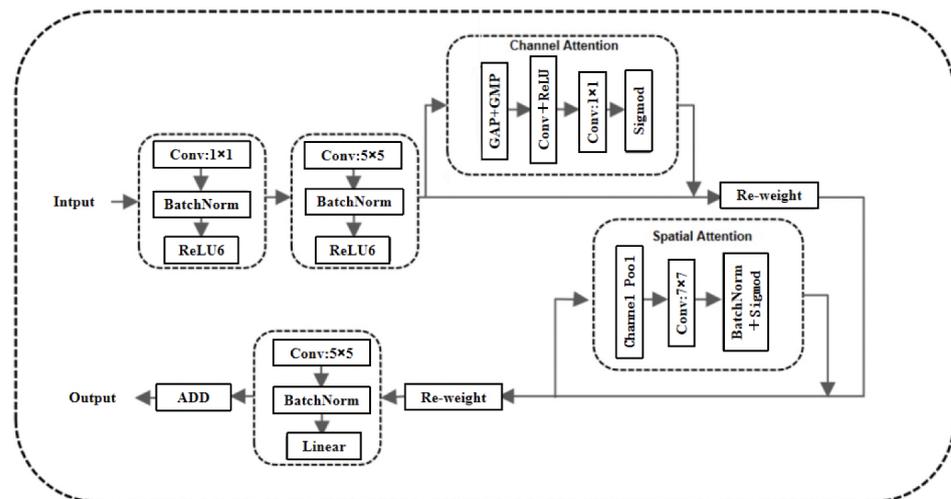


Figure 6. The baseline module of the improved feature extractor(Conv Layer: convolutional layers; GAP: global average pooling; GMP: global max pooling).

2.3.2. Multi-Point Distribution Representation of Bone Age

Considering that a single bone age regression label does not well represent the nonlinear and continuous characteristics of skeletal development of the hand. Inspired by the

facial age assessment task [29], we proposed a novel method for representing bone age. Given any bone age label y_n , it can be represented as a convex combination of a set of point sequences with the formula shown below:

$$\begin{aligned}
 y'_n &= \frac{1}{\min} \left[\sum_{i=0}^{\min} \left(1 - \frac{y_n - x_i}{2i - 1}\right) x_i + \sum_{j=0}^{\min} \left(1 - \frac{z_j - y_n}{2j + 1}\right) z_j \right] \\
 \text{s.t. } &\sum_{i=0}^{\min} \lambda_i^1 + \sum_{j=0}^{\min} \lambda_j^2 = 1 \\
 &0 \leq \lambda_i^1 \leq 1 \\
 &0 \leq \lambda_j^2 \leq 1
 \end{aligned}
 \tag{4}$$

where $\min = \text{Min}\{\lfloor y_n \rfloor - 0, \lceil y_n - k \rceil\}$, $z_j = \lfloor y_n \rfloor - i$, and $x_i = \lceil y_n \rceil + i$; here, $\lfloor \cdot \rfloor$ and $\lceil \cdot \rceil$ denote the ceiling function and the floor function, respectively. The expressions for λ_i^1 and λ_j^2 in Equation (5) are shown below.

$$\begin{aligned}
 \lambda_i^1 &= \frac{1}{\min} \left(1 - \frac{y_n - x_i}{2i - 1}\right) \\
 \lambda_j^2 &= \frac{1}{\min} \left(1 - \frac{z_j - y_n}{2j + 1}\right)
 \end{aligned}
 \tag{5}$$

For instance, as shown in Figure 7, if $k = 10$ when $y_n = 3.6$ and $y_n = 7.8$, the corresponding λ is shown in the Figure 7.

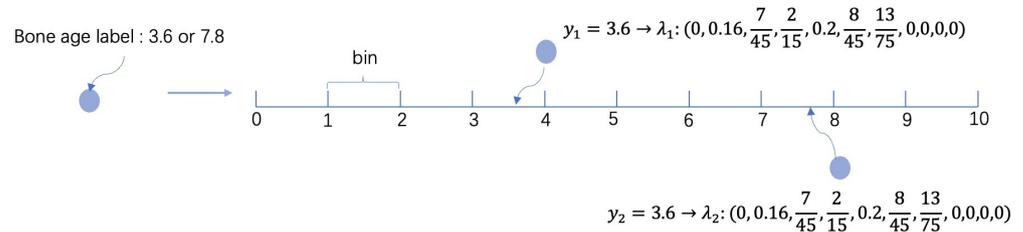


Figure 7. Multi-point distribution of bone age labels with semantic information.

The bone age label is translated into adjacent “min” bins. We let $k = 216$, i.e., one bin represents one month, and a sequence of k bins simulates each stage of skeletal development. Since each bin represents a different developmental stage, translating a single label into multi-point distribution labels that disperse in multiple different bins considering the non-linear features of different stages. The adjacency between bins can explain the relative continuity of hand bone morphology between adjacent age stages. Thus, this new representation of bone age uses distribution vectors containing rich semantic information to incorporate clinical features of skeletal development into the sequence.

2.3.3. Model Design for Cascade Training

Cascade operations are used to embed this vector of bone age labels containing semantic information into the BAA model, allowing the network to utilize information from classification, regression, and distribution learning simultaneously.

First, the hand bone X-ray image is input into the improved convolutional feature extractor to output feature map X . Then, considering the difference in hand bone development speed due to gender [30], we incorporate it into the network as a trainable neuron. Specifically, X is compressed into a vector by flattening layers and then feature vector Z is output by concatenating with gender (G).

$$Z = \text{Concat}(\text{flatten}(X), G)
 \tag{6}$$

Second, feature vector Z is passed through the FC layer and SoftMax activation function σ to output the probability that the bone age label belongs to each month (i.e., each bin), and we set the number of neurons in the FC to 216, corresponding to each month of the age range 0–18:

$$y'_n = \sigma(W_1 * Z + b) \quad (7)$$

where y'_n denotes the output of the FC layer, i.e., the probability that the network makes predictions each month, W_1 denotes the weight between z and FC, and b is the bias in network learning. The SoftMax activation function is the key to embedding the semantic distribution information labels into the network, which is represented as follows.

$$\sigma_i(y'_n) = \frac{e^{y_i}}{\sum_{j=1}^k e^{y_j}} \quad (8)$$

Finally, the bone age probabilities obtained in the FC layer for each month are inputted into a regression layer with only one neuron to obtain the final bone age assessment results. This is performed as follows: y'_n is multiplied with the weight W_2 from FC to regression layer plus bias to obtain the final bone age prediction output.

$$y_n = W_2 * y'_n + b \quad (9)$$

We call the design of the FC layer to the regression layer a cascade module, which allows the model to utilize classification, regression, and distribution learning simultaneously to maximize the use of the bone age label information.

2.4. Loss Function Optimization

In this paper, the focal loss (FL) function [31] and MAE function are chosen to train the cascade model together, where FL is used as the loss function for label distribution learning in the FC layer, and MAE is the loss function for predicting bone age values in the regression layer.

Due to the small and unevenly distributed characteristics of the dataset of medicine, the training network is prone to overfitting, so the loss function is improved in this paper to prevent overfitting and low accuracy. Specifically, for the uneven distribution of samples in the dataset, FL is chosen as the loss function in the label distribution learning phase. FL can automatically reduce the weight of simple samples during the training process by adding scale and quickly focusing the model on the hard-to-score samples to solve the category imbalance problem with the following expressions:

$$FL(p, y) = - \sum_{i=1}^N (1 - p_i)^\gamma \log(p_i), \quad p_i = \begin{cases} p & \text{if } y = 1 \\ 1 - p & \text{otherwise} \end{cases} \quad (10)$$

where y is the ground truth, and $p \in [0, 1]$ is the estimated probability of the model with label of $y = 1$. Then, optimization is performed at FL to prevent the overfitting problem; i.e., label smoothing regularization (LSR) [32] is added. LSR is a label smoothing process, which turns the label into a soft label with probability values. The probability value at the ground truth label is the maximum, and the probability value at other locations is a minimal number, and its expression is as follows.

$$y' = (1 - \epsilon)y + \epsilon u \quad (11)$$

Here, u is a prior overview distribution, and ϵ is the hyperparameter that controls the weight, and we set it to 0.2. The optimized label distribution learning loss function is shown below.

$$FL' = -(1 - \epsilon)FL(p, y) - \epsilon FL(u, y) \quad (12)$$

MAE is used to measure the evaluation results. The concrete form is expressed as follows, where y_i is the ground truth, and \hat{y}_i is the predicted bone age value, and it represents the total amount of data.

$$MAE = \frac{1}{m} \sum_{i=1}^m |(y_i - \hat{y}_i)| \quad (13)$$

The total loss function of the our model is written as follows. λ is a hyperparameter used to control the weight of the two loss functions, and we set it to 0.25.

$$Loss = MAE + \lambda FL' \quad (14)$$

3. Results

3.1. Implementation Details

The deep learning framework used in this paper is PyTorch, and the hardware configuration includes an Intel (R) core (TM) i7-8700k CPU and an NVIDIA GeForce GTX 1080ti GPU. The software configuration contains the Ubuntu 16.04 operating system, CUDA 9.0, cuDNN 7.0, PyTorch 1.7, and Python 3.7. The hyperparameters of the first stage module are set as follows: Epoch = 600; input image size = 640×640 ; initial learning rate = 0.01 and One Cycle Policy was used to adjust the learning rate; optimizer is SGD; batch size = 16; and the activation function is Swish. The hyperparameters of the second stage module are set as follows: batch size = 16; epoch = 600; input image size = 640×640 ; learning rate = 0.001, 10% reduction in learning rate for every 10 epochs without reduction in loss; optimizer is Adam; and activation functions is ReLU. Moreover, to further prevent overfitting, we performed data enhancement operations such as flipping, rotating, and cropping on the training samples.

3.2. Evaluation of the RoIs Extraction

To validate the localization performance of the RoIs extraction model, we randomly selected 200 images to test YOLOv5. Table 1 shows the mean Average Precision (mAP) of different RoIs. The expression of mAP is as follows.

$$mAP = \frac{\sum_{i=1}^k AP_i}{k} \quad (15)$$

Here, $AP = \frac{IoU > \gamma}{N}$, $IoU > \gamma$ denotes the number of Intersection over Union (IoU) between the predicted object boxes, and the ground truth is greater than threshold value γ . N denotes the total number of test samples, mAP is the average of all category APs, and k is the number of different categories contained in the test set. In Table 1, mAP@0.5 denotes the mean average precision for $IoU > 0.5$, and mAP@0.5:0.95 denotes the mean average precision for increasing the IoU threshold from 0.5 to 0.95 (in steps of 0.05).

Table 1. mAP for different RoIs positioning.

RoIs	mAP@0.5	mAP@0.5:0.95
Hand (H)	0.99	0.91
Carpal (C)	0.99	0.80
Phalanx (P)	0.99	0.60
H+C+P	0.99	0.70

It can be seen from Table 1 that the localization accuracy of “H”, “P”, “C”, and “H+P+C” RoIs are close to 1 when the threshold value of IoU is 0.5. When the threshold of IoU is limited to 0.5, an IoU of 0.5 between the ground truth and the predicted object box is considered as a successful localization, resulting in a high value of mAP. In addition, since the hand bone background is single and the pose and position of the hand are required when taking X-rays, the position of the hand bone in the image is relatively fixed and is conducive to the improvement of localization accuracy. Second, the mAP@0.5:0.95 of “H”,

“P”, and “C” in the table gradually decreases due to the relatively small RoIs objects of “C” and “P”, resulting in the inability of the network to locate them accurately. In particular, the RoIs of “P” also has interference from the same parts of other fingers. To ensure that the model can smoothly cut the corresponding RoIs object boxes, the threshold value of IoU in the training network is set to 0.5 in this paper.

To validate the classification performance of the model for RoIs, the normalized confusion matrix is also analyzed in this paper. The confusion matrix is a matrix with the number of the predicted category as the vertical axis and the number of the ground truth as the horizontal axis. The diagonal of its normalized matrix represents the probability of accurately predicting the category as its ground truth category.

As shown in Figure 8, the confusion matrix’s high diagonal values indicate the model’s high classification accuracy. We believe that this is due to the large morphological gap between the “H”, “C”, and “P” RoIs. Therefore, the model classifies them correctly with high accuracy.

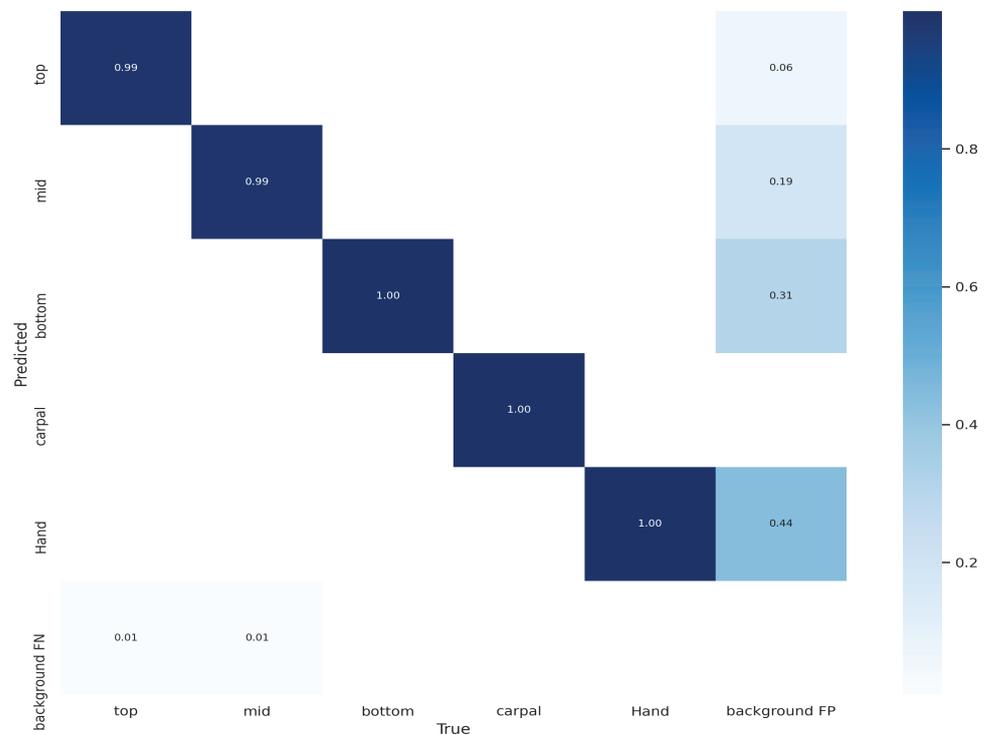


Figure 8. Confusion matrix of RoIs.

3.3. Evaluation of Bone Age Assessment

Baseline networks We pre-trained on five baseline models to select a better performing baseline network for the BAA task, including ResNet50, MobileNet v3, AlexNet, VGG, and MobileNet v2. All baseline networks were trained from scratch, no pre-trained weights were loaded, and the networks were initialized randomly. The MAE assessed by the baseline model and the parametric model size are shown in Table 2.

Table 2. Test results for different baseline models.

Baseline	MobileNet v2	MobileNet v3	ResNet50	VGG	AlexNet
MAE (month)	7.21	7.73	8.49	8.77	10.23
Para-size (M)	3.4	3.2	25	138	61

As observed from Table 2, the MobileNet v2 baseline model has the best evaluation results with the least number of model parameters and, thus, runs efficiently. This is attributed to the fact that the feature extractor of MobileNet v2 is composed of deep-separable convolution, which makes the network light, effectively prevents overfitting of the model, and is suitable for small datasets such as medical datasets. Therefore, MobileNet v2 is chosen as the baseline network for the BAA task in this paper.

Fusion of RoIs after background denoising and RoIs of prior knowledge. In order to verify the effectiveness of background denoising and fusion of prior knowledge for the BAA task, we compare the original image (O) and the background denoising hand image (Hand, H) input model. In clinical practice, if the error between the chronological age and bone age is within one year of the ground truth [33], it is considered that the development of this bone is within the normal range. This paper sets three accuracy standards for different ranges of time (that is, 6 months, 12 months, and 24 months) to evaluate the accuracy of the proposed model in a multiscale manner. Above this, the RoIs of RUS-CHN (i.e., Carpal and Phalanx and C and P) multi-channel input networks are compared, and the experimental results are shown in Table 3.

Table 3. Test results for different inputs base on baseline models.

O	H	C	P	MAE	± 0.5	± 1	± 2
✓				7.21	56.82%	79.23%	96.86%
	✓			6.45	58.86%	83.33%	97.22%
	✓	✓		6.31	58.75%	86.71%	97.92%
	✓	✓	✓	6.17	59.01%	88.89%	98.31%

As shown in Table 3, compared to using “O” as input, the effect of “H” as the input network was improved and MAE decreased by 0.76 months, and accuracy is improved by 2.04%, 4.1%, and 0.36% for ground truth within ± 0.5 year, ± 1 year, and ± 2 year, respectively. This illustrates that background denoising can improve the accuracy of model evaluation and has a positive effect on BAA model evaluation. In addition, Table 3 shows that the MAE of the model decreases by 1.04 when “H+C+P” is used as input compared to “H” as input, and accuracy is improved by 2.19%, 9.66%, and 1.45% for ground truth within ± 0.5 year, ± 0.5 year, and ± 0.5 year, respectively. The experimental results in Table 3 illustrate that background denoising and fused prior knowledge can effectively improve the evaluation accuracy of the BAA model and bring a positive impact to the BAA task.

BAA model ablation experiments. To evaluate the effectiveness of the improved feature extractor and the proposed bone age assessment label representation for the BAA task, ablation experiments were conducted. Where the improved feature extractor contains two modules for ablation experiments, i.e., CBAM and 5×5 convolution kernel size (k: 5×5), “MP” in Table 4 represents the proposed Multi-point distribution of bone age labels, and “G” denotes the gender information vector.

Table 4. BAA model ablation experiments.

Improved Feature Extractor		MP	G	MAE	± 0.5	± 1	± 2
CBAM	k: 5×5						
✓				5.05	66.11%	92.23%	98.81%
	✓			5.53	62.51%	90.79%	98.61%
✓	✓			5.01	66.67%	92.25%	98.65%
		✓		4.86	71.11%	93.33%	98.67%
		✓	✓	4.72	74.81%	93.60%	98.67%
✓	✓	✓	✓	4.53	78.15%	94.04%	99.34%

From Table 4, we can see that, compared to the baseline model, applying the improved feature extractor decreases MAE by 1.16 and accuracy is improved by 7.66%, 3.36%, and 0.34% for ground truth within ± 0.5 year, ± 1 year, and ± 2 year, respectively. We believe

this increase in the size of the convolutional kernel can expand the range of the RF, thus enhancing the local feature extraction capability of the model, which can extract more local features of the hand bone and improve the performance of the model. Meanwhile, the adaptive assignment of channel weights by CAM allows the model to automatically place larger weights on the more important RoIs channels, thus improving the evaluation accuracy, and the SAM module enables the model to focus attention on the spatial region of the hand bone, further improving evaluation accuracy. After adding Multi-point distributed bone age labels containing rich semantic information to the model, compared to the baseline model, MAE decreased by 1.31 and accuracy is improved by 12.1%, 4.44%, and 0.36% for ground truth within ± 0.5 year, ± 1 year, and ± 2 year, respectively. It is shown that the proposed multi-point distribution bone age label, which well interprets the characteristics of skeletal development, contains the semantic information of classification, regression, and distribution learning within the label such that the model fully utilizes the label information and achieves a better assessment effect. In addition, we add the gender information to the network, and MAE decreased by 0.08 and accuracy is improved by 3.70% and 0.27% for ground truth within ± 0.5 year and ± 1 year, respectively. This is because adding gender information to the bone age label can further enrich the semantic information of the label, allowing the model to learn more useful information to help bone age assessment.

Loss function comparison experiment. In order to verify the effectiveness of the improved label distribution learning loss function, we conducted ablation experiments. At the same time, we compared it with the commonly used probability distribution learning loss function Kullback–Leibler Divergence (KL), and the experimental results are shown in Table 5. Note that the loss functions of the regression layers are all using the MAE.

Table 5. Loss function comparison experiment.

	MAE	± 0.5	± 1	± 2
KL	4.69	75.32%	93.71%	98.80%
FL	4.53	78.15%	94.04%	99.34%
FL + LS	4.44	79.47%	94.70%	99.34%

From the Table 5, compared to KL, using the FL cascade to train the model, MAE decreased by 0.16 and the accuracy is improved by 2.83%, 0.33%, and 0.54% for ground truth within ± 0.5 year, ± 1 year, and ± 2 year, respectively. We believe this is because FL is more applicable to data with uneven sample distribution such as medical datasets, which enables the model to adaptively adjust the weights of difficult samples, thus focusing the model's attention on difficult sample discovery and improving the model recognition performance. In addition, evaluation performance is further enhanced by the combined loss function of FL + LS, MAE decreased by 0.09, and accuracy is improved by 1.32% and 0.66% for ground truth within ± 0.5 year and ± 1 year, respectively. That is, the proposed loss function can improve the uneven distribution of samples in medical data sets and the characteristics of overfitting caused by the small amount of data to a certain extent.

Assessment performance at different ages. In order to detect the performance of the model in each age group, the paper divides 0–18 years into 18 age intervals to verify the performance of the model. A new evaluation metric, RMSE, is also introduced, which will provide a higher weight to larger errors than with MAE will, and for larger errors, RMSE is the most useful evaluation metric with the following expression.

$$RMSE = \sqrt{\frac{1}{m} \sum_{i=1}^m (y_i - \hat{y}_i)^2} \quad (16)$$

From Table 6, we can find that the model assessment results are more stable at (0, 15] years of age, with low MAE and RMSE, while the error in the assessment is larger at (15, 18] years of age, especially between (17, 18] years of age. We speculate is due to the fact

that from 15 to 18 years of age, the skeletal development of the hand gradually matures, and its morphology no longer changes significantly, making it difficult for the model to accurately assess bone age.

Table 6. MAE and RMSE for different age groups.

Age Groups	MAE	RMSE
(0, 1]	3.30	3.92
(1, 2]	4.29	4.69
(2, 3]	3.86	4.64
(3, 4]	3.57	4.26
(4, 5]	4.20	4.82
(5, 6]	3.80	4.27
(6, 7]	3.67	4.40
(7, 8]	4.13	4.65
(8, 9]	3.63	4.70
(9, 10]	3.10	4.63
(10, 11]	2.11	2.53
(11, 12]	3.10	3.81
(12, 13]	2.69	3.83
(13, 14]	3.62	4.29
(14, 15]	4.09	4.69
(15, 16]	5.20	5.94
(16, 17]	6.00	7.84
(17, 18]	11.00	11.59

Comparison experiments with advanced models To verify the performance as well as the applicability of the proposed model, it is validated on the RNSA2017 public dataset, and since this dataset is not evaluated using the RUS-CHN method, only background denoising RoIs are used in the comparative experiments on the public dataset, and the results of the comparative experiments are shown in Table 7.

Table 7. Performance comparison experiments.

Dataset	Improved Feature Extractor		MAE	± 0.5	± 1	± 2
	Background Denoise	Prior Knowledge				
Miao, J. et al. [34]	RNSA2017		10.41	-	68.80	93.20
Salim, I. et al. [20]	RNSA2017	✓	6.38	-	-	-
Hao, P. et al. [35]	RNSA2017	✓	6.29	54.45%	90.80%	98.50%
Iglovikov, V.I. et al. [23]	RNSA2017	✓	4.97	-	-	-
Our	RNSA2017	✓	4.95	70.00%	95.20%	99.40%
Son, S.J. et al. [19]	Privacy		5.52	-	-	-
Mutasa, S. et al. [2]	Privacy	✓	7.64	-	-	-
Our	Privacy	✓	4.79	72.45%	93.80%	98.81%
Our	Privacy	✓	4.44	79.47%	94.70%	99.81%

Table 7 shows that when the model uses only background denoised RoIs, it achieves good levels of accuracy on both private and public datasets, with model evaluation accuracies of 93.80% and 95.20% for ground truth within ± 1 year. Compared with other BAA models using background denoising on public datasets, MAE is lower, and the evaluation accuracy is higher. We believe that this is due to the cascade model proposed in this paper being more consistent with the characteristics of skeletal development and makes full use of the bone age labeling information, thus improving the evaluation of the BAA model. Meanwhile, Table 7 shows that the cascade model proposed in this paper also has some applicability for the public dataset.

3.4. Consistency Analysis

The intraclass consistency coefficient (ICC) and Pearson's correlation coefficient (r) are used to analyze the consistency between the assessment results of the model and experts. The experimental results show that $ICC = 0.997$ and $r = 0.994$, $p < 0.01$, demonstrating a significant positive correlation between the two assessment results. i.e., the results of the model evaluation and those of the physicians' evaluation are highly consistent.

In addition, regression analysis plots as well as the Bland–Altman method are used to visualize the consistency between the physician assessment results and the model assessment results, as shown in Figures 9 and 10.

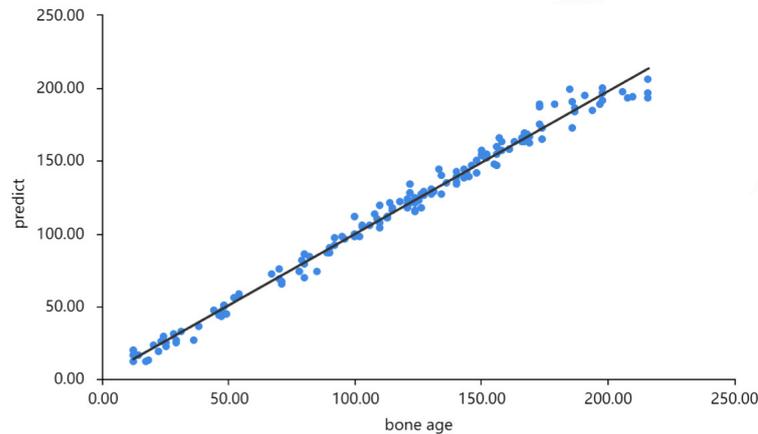


Figure 9. Regression Fitting Plot. A straight line represents ground truth, and discrete points represent predicted bone age values.

In the Bland–Altman plot, the upper and lower horizontal solid lines represent the limits of the 95% Confidence Interval (CI), i.e., 1.96 times the standard deviation, and the middle horizontal dashed line represents the mean of the differences. The Bland–Altman plot shows that most of the results predicted by the model are within the 95% CI, and only a few results are outside the 95% CI. Moreover, the regression fit plots show that the fit between the model evaluation results and the physician evaluation results is close. In conclusion, Figure 10 fully demonstrates that the above evaluation results are highly positively correlated.

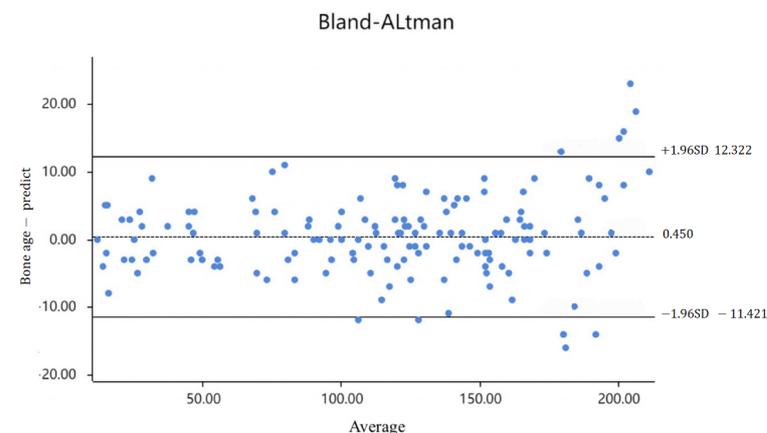


Figure 10. Bland–Altman plot for the predicted bone age.

4. Discussion

Based on the skeletal growth and development characteristics of Chinese adolescents and children since the start of the century, this paper proposes a cascade model that fuses prior knowledge. First, the model is characterized by the simultaneous use of classification,

regression, and distribution learning. Compared to traditional classification or regression models, the proposed model combines the nonlinearity and continuity of hand bone development. Second, the feature extractor of the model is improved based on the baseline network, which allows the model to extract more hand bone features and, thus, improves the evaluation accuracy. Third, the RoIs of the RUS-CHN method are considered prior knowledge in the network, which also improve the accuracy of the model. The final MAE of the model is 4.4 months, and the accuracy of the assessment within one year of the ground truth is 94.04%.

On the one hand, aiming at the problem that the traditional network does not take the nonlinearity and continuity of bone growth and development into account, a novel bone age representation is proposed, where the semantic information is embedded in the BAA task using a cascade model. In detail, a fully connected layer with a semantic distribution of 216 neurons (corresponding to each month in the age range of 0–18 years) is inserted between the feature and regression layers. A multi-point distribution bone age label is output between the fully connected layer and the regression layer, and this label subsequently passes through the regression layer to obtain the corresponding predicted bone age value. A single digital bone age label is converted into 216 discrete and adjacent bins. Because the bins are discrete, each month of bone development between the ages of 0 and 18 is regarded as an independent individual, thereby taking the phased characteristics into account. At the same time, the adjacent bins between bins can be regarded as relatively continuous, as in the development ages of hand bones in adjacent months, which are independent of each other but have relative continuity. In brief, this design more closely reflects the reality of skeletal development, thus improving the consistency of the model assessment.

On the other hand, the feature extractor is improved in three aspects, namely channel, spatial, and RF. Among them, the weight sizes of the channels where the three RoIs are located are newly assigned by CAM so that the model places a larger weight on top of the channel where evaluation is more meaningful. Moreover, the model's attention is spatially integrated by SAM to focus the model learning region on the hand bone region. In addition, increasing the RF size of the feature extractor allows the model to enhance the extraction of global information and obtain more global features. The three improvements enhance the feature extractor's ability to obtain valuable features, thus improving the accuracy of the evaluation.

As to comparison with the most recent works [36,37], our model considers the nonlinearity and continuity simultaneously of skeletal development by converting the bone age label into a multi-point distribution label. In addition, we extracted the RoIs input model using YOLO v5, allowing the model to catch more detailed features and reach better assessment, with a 1.6 months reduction in MAE for our model compared to the work of [36].

In addition, the proposed network adopts target detection algorithms to localize the RoIs present in the RUS-CHN method and the region of the entire hand, and the corresponding regions of the three types of bones (hand, carpus, and phalanges) are extracted as prior knowledge, which can be subsequently learned by the model in a multichannel manner. Among them, the carpal bone is very important for bone age evaluations in infants and young children, and after the development and maturity of the carpal bone, observing the morphology of the phalanges can help doctors diagnose bone ages. The employed prior knowledge is based on the skeletal RoIs that are used by physicians in clinical evaluations, so the evaluation accuracy of the model is further improved.

The weakness of this paper is found in the experimental results: A relatively poor level of consistency was observed between the results of the model and those of physicians' assessments after 15 years of age (180 months), as shown in Figure 9 and Table 6. It is speculated that poorer assessment results may be obtained since the form of the hand bone is approximately stable and no longer changes significantly in the later stages of development. For this reason, some specific bones, such as the knee and vertebrae, are generally used to replace the hand bone in evaluations conducted during later developmental stages or

in adulthood [38]. Moreover, the lack of data in some age groups may be another reason why the characteristics learned by the model are insufficient. Further studies are needed to determine whether the satisfactory results produced by the proposed model can be maintained after the age of 15.

In future studies, we will consider improving the performance of the model in terms of three aspects. First, we will enter knee X-rays that are used for assessing bone age in adulthood into the model to enhance the accuracy of the model during later stages of development. Second, the sample data will be expanded in each stage within 15–18 years, thereby enabling the model to fully learn the characteristics of each age group in this range and achieve better assessment accuracy. The expansion of the data volume and the operation of integrating multisite joints into the evaluation model may improve the evaluation accuracy of our model. Third, based on the RoIs of different evaluation methods, such as TW method and GP method, the priori knowledge can be extracted and further integrated into the model, which can expand the application scope.

5. Conclusions

In this paper, a cascade model incorporating prior knowledge is proposed. Compared with the conventional model, the model is made to consider both nonlinearity and continuity of bone development using the bone age label representation proposed in this paper, thus improving the consistency of the model. In addition, the accuracy of the model is improved by using an improved feature extractor and incorporating RoIs. Experimental results show that the model proposed in this paper exhibits good performance on both private and public datasets, which indicates that the model is adaptable and can be applied to other fields.

Author Contributions: Drafting articles, building deep learning models, experimental analysis, and experimental design, N.L.; data collation and experimental analysis, B.C.; revision of the article and methodological suggestions, J.Z. All authors have read and agreed to the published version of the manuscript.

Funding: Sichuan Science and Technology Program (No. 2022YFS0178); Chengdu Science and Technology Program (No. 2021-YF05-00916-SN); Smart Grid Sichuan Provincial Key Laboratory Emergency Key Project (No. 2020IEPG-KL-20YJ01).

Institutional Review Board Statement: The West China Medical Board of Sichuan University concluded that the study conducted was not characterized as a medical experiment and, therefore, agreed to proceed with the work.

Informed Consent Statement: Not applicable.

Data Availability Statement: The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Wang, F.; Gu, X.; Chen, S.; Liu, Y.; Shen, Q.; Pan, H.; Shi, L.; Jin, Z. Artificial intelligence system can achieve comparable results to experts for bone age assessment of Chinese children with abnormal growth and development. *PeerJ* **2020**, *8*, e8854.
2. Mutasa, S.; Chang, P.D.; Ruzal-Shapiro, C.; Ayyala, R. MABAL: A Novel Deep-Learning Architecture for Machine-Assisted Bone Age Labeling. *J. Digit. Imaging* **2018**, *31*, 513–519.
3. Nejedly, N. Normal and abnormal growth in the pediatric patient. *Curr. Probl. Pediatr. Adolesc. Health Care* **2020**, *50*, 100771.
4. Lee, H.; Tajmir, S.; Lee, J.; Zissen, M.; Yeshiwas, B.A.; Alkasab, T.K.; Choy, G.; Do, S. Fully Automated Deep Learning System for Bone Age Assessment. *J. Digit. Imaging* **2017**, *30*, 427–441.
5. Grulich, W.; Pylc, S. *Radiographic Atlas of Skeletal Development of the Hands and Wrists*; Stanford University Press: Redwood City, CA, USA, 1959.
6. Liu, J.; Qi, J.; Liu, Z.; Ning, Q.; Luo, X. Automatic bone age assessment based on intelligent algorithms and comparison with TW3 method. *Comput. Med Imaging Graph.* **2008**, *32*, 678–684.
7. Nadeem, M.W.; Goh, H.G.; Ali, A.; Hussain, M.; Khan, M.A.; Ponnusamy, V.a. Bone age assessment empowered with deep learning: A survey, open research challenges and future directions. *Diagnostics* **2020**, *10*, 781.

8. Zhao, H.; Lazarenko, O.P.; Chen, J. Hippuric acid and 3-(3-hydroxyphenyl) propionic acid inhibit murine osteoclastogenesis through RANKL-RANK independent pathway. *J. Cell. Physiol.* **2020**, *235*, 599–610.
9. Tamme, R.; Jürime, J.; Mestu, E.; Rimmel, L.; Purge, P.; Mengel, E.; Tillmann, V. Physical Activity in Puberty Is Associated with Total Body and Femoral Neck Bone Mineral Characteristics in Males at 18 Years of Age. *Medicina* **2019**, *55*, 203.
10. Zhang, S.; Liu, L. The Skeletal Development Standards of Hand and Wrist for Chinese Children China 05 1. TW3-C RUS, TW3-C Carpal, and RUS-CHN Methods. *Chin. J. Sport. Med.* **2003**. Available online: <https://pesquisa.bvsalud.org/portal/resource/pt/wpr-587292> (accessed on 4 May 2022).
11. Guang, C.; Qile, P.; Rongxin, Z. A comparative study of the consistency between the Greulich-Pyle method and the China 05 method in normal children aged 6–18 years. *Chin. J. Evid. Based Pediatr.* **2020**, *15*, 441.
12. Thodberg, H.H.; Kreiborg, S.; Juul, A.; Pedersen, K.D. The BoneXpert method for automated determination of skeletal maturity. *IEEE Trans. Med Imaging* **2008**, *28*, 52–66.
13. Liu, B.; Zhang, Y.; Chu, M.; Bai, X.; Zhou, F. Bone Age Assessment Based on Rank-Monotonicity Enhanced Ranking CNN. *IEEE Access* **2019**, *7*, 120976–120983.
14. Tong, C.; Liang, B.; Li, J.; Zheng, Z. A deep automated skeletal bone age assessment model with heterogeneous features learning. *J. Med. Syst.* **2018**, *42*, 1–8.
15. Pietka, E.; Gertych, A.; Pospiech, S.; Cao, F.; Huang, H.K.; Gilsanz, V. Computer-assisted bone age assessment: Image preprocessing and epiphyseal/metaphyseal ROI extraction. *IEEE Trans. Med. Imaging* **2001**, *20*, 715–729.
16. Dallora, A.L.; Anderberg, P.; Kvist, O.; Mendes, E.; Diaz Ruiz, S.; Sanmartin Berglund, J. Bone age assessment with various machine learning techniques: A systematic literature review and meta-analysis. *PLoS ONE* **2019**, *14*, e0220242.
17. Dehghani, F.; Karimian, A.; Sirous, M. Assessing the bone age of children in an automatic manner newborn to 18 years range. *J. Digit. Imaging* **2020**, *33*, 399–407.
18. Spampinato, C.; Palazzo, S.; Giordano, D.; Aldinucci, M.; Leonardi, R. Deep learning for automated skeletal bone age assessment in X-ray images. *Med. Image Anal.* **2017**, *36*, 41–51.
19. Son, S.J.; Song, Y.; Kim, N.; Do, Y.; Kwak, N.; Lee, M.S.; Lee, B.D. TW3-based fully automated bone age assessment system using deep neural networks. *IEEE Access* **2019**, *7*, 33346–33358.
20. Salim, I.; Hamza, A.B. Ridge regression neural network for pediatric bone age assessment. *Multimed. Tools Appl.* **2021**, *80*, 30461–30478.
21. Ren, X.; Li, T.; Yang, X.; Wang, S.; Ahmad, S.; Xiang, L.; Stone, S.R.; Li, L.; Zhan, Y.; Shen, D.; et al. Regression convolutional neural network for automated pediatric bone age assessment from hand radiograph. *IEEE J. Biomed. Health Inform.* **2018**, *23*, 2030–2038.
22. Han, Y.; Wang, G. Skeletal bone age prediction based on a deep residual network with spatial transformer. *Comput. Methods Programs Biomed.* **2020**, *197*, 105754.
23. Igloukov, V.I.; Rakhlin, A.; Kalinin, A.A.; Shvets, A.A. Paediatric bone age assessment using deep convolutional neural networks. In *Deep Learning in mEdical Image Analysis and Multimodal Learning for Clinical Decision Support*; Springer: Berlin/Heidelberg, Germany, 2018; pp. 300–308.
24. Chen, S.; Zhang, C.; Dong, M. Deep age estimation: From classification to ranking. *IEEE Trans. Multimed.* **2017**, *20*, 2209–2222.
25. Jubayer, F.; Soeb, J.A.; Mojumder, A.N.; Paul, M.K.; Barua, P.; Kayshar, S.; Akter, S.S.; Rahman, M.; Islam, A. Detection of mold on the food surface using YOLOv5. *Curr. Res. Food Sci.* **2021**, *4*, 724–728.
26. Sandler, M.; Howard, A.; Zhu, M.; Zhmoginov, A.; Chen, L.C. Mobilenetv2: Inverted residuals and linear bottlenecks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 4510–4520.
27. Xueli, X.; Chuanxiang, L.; Xiaogang, Y.; Jianxiang, X.; Tong, C. Dynamic Receptive Field-Based Object Detection in Aerial Imaging. *Acta Opt. Sin.* **2020**, *40*, 0415001.
28. Woo, S.; Park, J.; Lee, J.Y.; Kweon, I.S. Cbam: Convolutional block attention module. In Proceedings of the European conference on computer vision (ECCV), Munich, Germany, 8–24 September 2018; pp. 3–19.
29. Zhang, C.; Liu, S.; Xu, X.; Zhu, C. C3AE: Exploring the limits of compact model for age estimation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 12587–12596.
30. Su, L.; Fu, X.; Hu, Q. Generative adversarial network based data augmentation and gender-last training strategy with application to bone age assessment. *Comput. Methods Programs Biomed.* **2021**, *212*, 106456.
31. Tran, G.S.; Nghiem, T.P.; Nguyen, V.T.; Luong, C.M.; Burie, J.C. Improving accuracy of lung nodule classification using deep learning with focal loss. *J. Healthc. Eng.* **2019**, *2019*, 5156416.
32. Mo, X.; Wei, T.; Zhang, H.; Huang, Q.; Luo, W. Label-Smooth Learning for Fine-Grained Visual Categorization. In Proceedings of the Asian Conference on Pattern Recognition, Auckland, New Zealand, 26–29 November 2019; Springer: Berlin/Heidelberg, Germany, 2019; pp. 17–31.
33. Zulkifley, M.A.; Abdani, S.R.; Zulkifley, N.H. Automated bone age assessment with image registration using hand X-ray images. *Appl. Sci.* **2020**, *10*, 7233.
34. Miao, J.; Yue, H.; Wu, X.; Xu, D.; Chen, W. Bone age assessment based on SuperPoint features. In Proceedings of the 2020 Chinese Automation Congress (CAC), Shanghai, China, 6–8 November 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 3921–3926.
35. Hao, P.; Ye, T.; Xie, X.; Wu, F.; Ding, W.; Zuo, W.; Chen, W.; Wu, J.; Luo, X. Radiographs and texts fusion learning based deep networks for skeletal bone age assessment. *Multimed. Tools Appl.* **2021**, *80*, 16347–16366.

36. Nguyen, Q.H.; Nguyen, B.P.; Nguyen, M.T.; Chua, M.C.; Do, T.T.; Nghiem, N. Bone age assessment and sex determination using transfer learning. *Expert Syst. Appl.* **2022**, *200*, 116926.
37. Palaniswamy, T. Hyperparameter optimization based deep convolution neural network model for automated bone age assessment and classification. *Displays* **2022**, *73*, 102206.
38. Dallora, A.L.; Berglund, J.S.; Brogren, M.; Kvist, O.; Ruiz, S.D.; Dübbel, A.; Anderberg, P. Age Assessment of Youth and Young Adults Using Magnetic Resonance Imaging of the Knee: A Deep Learning Approach. *JMIR Med. Inform.* **2019**, *7*, e16291.