International Journal of Computer Science and Information Technology
ISSN: 3005-9682 (Print), ISSN: 3005-7140 (Online) | Volume 5, Number 3, Year 2025
DOI: https://doi.org/10.62051/ijcsit.v5n3.06
Journal homepage: https://wepub.org/index.php/IJCSIT/index

Bone Age Assessment Method Based on Improved ResNet

Xiaoxun Ma

Southwest Minzu University, Chengdu, China
lajiaos39@gmail.com

ABSTRACT

To extract the features of hand X-ray images in more detail, a bone age assessment method based on an improved residual network is proposed. Based on the RUS-CHN method, a lightweight and efficient attention module is combined with the residual network to improve the accuracy of extracting fine-grained features. Experimental results show that on the dataset provided by a certain tertiary hospital in Xi'an, the average absolute errors (MAE) for males and females are 0.4228 years and 0.4341 years respectively. Within a 1-year error range, the accuracy rates for males and females reach 94.6% and 93.7% respectively, significantly improving the accuracy of bone age assessment.

KEYWORDS

Bone age assessment; Improved residual network; Lightweight and efficient attention module; RUS-CHN

1. INTRODUCTION

With the continuous development of economy and society, the growth and development of adolescents has been widely concerned [1]. Bone age is an important indicator of human growth and development [2]. In hospital, bone age detection is generally carried out by spectrogram method or scoring method. Doctors will compare the development status of hand bone to be measured with the standard spectrogram to give the evaluation results [3]. Because different doctors have different assessment experience and different assessment methods, the subjectivity of bone age detection is larger. At the same time, due to the complexity of testing standards and testing processes, the general evaluation time is long and the error is high, which brings greater difficulties to the accurate evaluation [4].

The commonly used bone age assessment methods in clinical fields in China include counting method, GP method [5-6] and TW2 and TW3 scoring method [7]. and RUS-CHN method. As the name suggests, counting is the first method to evaluate bone age by calculating the number of ossification centers. Bone age counting method showed that the age of ossification center was positively correlated with the speed of bone maturation, and abnormal ossification center number was associated with abnormal development. The calculation method of wrist ossification points is that the number of ossification points equals the actual age plus 1, that is, the difference between the number of ossification centers and the actual age is 1, indicating that the physical development of the child is close to the normal level.

A series of X-ray charts of wrist bone maturity were developed from the location of ossification center, the proportion of epiphysis to metaphyseal, the appearance of certain indusses and the relationship between epiphysis and metaphyseal in adolescents from newborn to 19 years old. The interpretation was based on the time, size, shape and density of each ossification center. The mean absolute error of calculating bone age by counting method is large, so this method is rarely used to judge the degree of epiphyseal development at home and abroad. GP Atlas method relies on the rich experience of doctors and is highly subjective. Different doctors may have different results when evaluating the same X-ray image. The idea of RUS-CHN scoring method is to grade some Regions of Interest (ROIs) on the left hand, obtain the corresponding grade score for each epiphyseal region, and convert the total grade score into the corresponding bone age [8]. Rus-chn method is more accurate than GP atlas method in bone age assessment, but it also has the problem of relying on the subjective experience of doctors and taking a long time [9]. In recent years, due to the rapid development of deep learning technology, the innovation and improvement of deep learning architecture has become a hot issue in industry and academia. As a new deep learning architecture, cyclic convolutional neural network has shown great potential in image recognition, image classification and image detection.

In the research field of bone age assessment, with the development and rise of artificial intelligence, Convolutional Neural Networks (CNNs) [10-11] are also used to evaluate the development degree of epiphysis instead of traditional manual reading of X images. In 2018, Hu Tinghong et al. [12] used AlexNet network to automate bone age assessment based on left wrist X-ray images of Uyghur teenagers. Within the 1-year error range, the accuracy of male and female test sets was 79.5% and 71.2%, respectively. In 2019, Zhan Mengjun et al. [13] conducted an automatic detection and analysis of the left wrist bone age of Han teenagers in Sichuan. Using the improved AlexNet network to evaluate bone age, the accuracy of male and female bone age was 81.08% and 87.56%, respectively, within 1-year error range. In 2020, Zhang Shijie et al. [14] proposed a method based on deep learning convolutional neural network to evaluate epiphyseal maturity. In this study, the Second Hospital of West China, Sichuan University, was selected as the research object, and the accuracy rates were 94.25% for females and 94% for males within the error range of ±1 year. The mean absolute error is 0.512 5 years for women and 0.557 5 years for men. In 2021, Tang Zhihao et al. [8] proposed a residual network combined with high-efficiency channel attention module to evaluate the degree of bone growth based on RSNA and DHA data samples. When the error is 1 year, the average absolute error is 4.69 months and 5.98 months respectively, and the accuracy rate of bone age assessment can reach 98.36% and 94.88%.

In order to enable the network to extract more epiphyseal information and achieve a more accurate and efficient bone age assessment, an improved residual network combined with attention model was proposed to evaluate bone age. The residual network ResNet [15-16] was double-pooled and combined with the lightweight and efficient attention module SA [17], so that the network could learn more fine particle epiphyseal features.

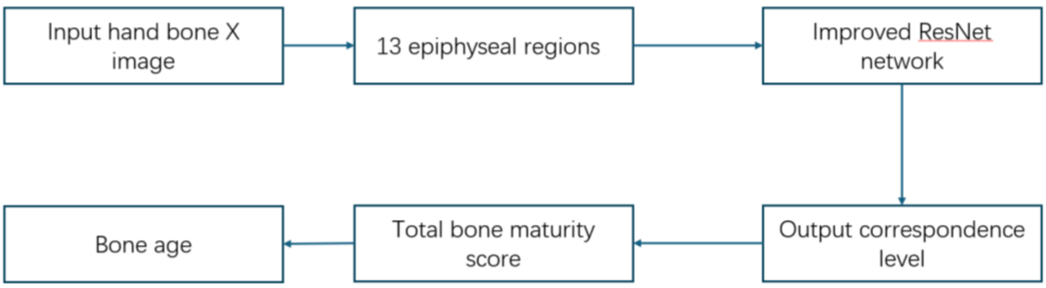

Figure 1. The process of bone age assessment

2. IMPROVED BONE AGE ASSESSMENT NETWORK

2.1. Combining Lightweight and Efficient Attention Modules

In order to make the convolutional neural network accurately focus on the features of input epiphysis, suppress the features of disinterest, and improve the accuracy of bone age assessment, a new efficient attention mechanism (SA) model was added to the residual network. This model integrates spatial attention and channel attention, and takes feature grouping and channel replacement as a lightweight and effective attention mechanism. The model is shown in Figure 2.

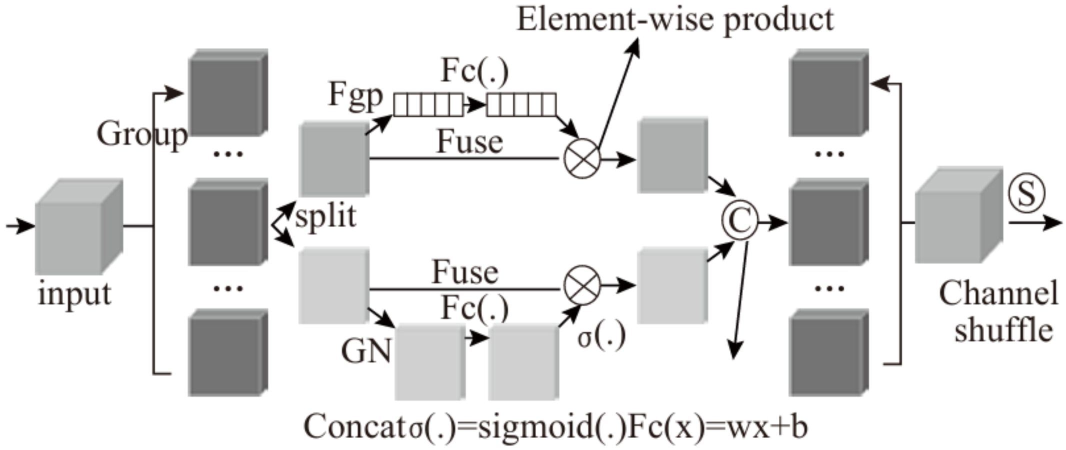

Figure 2. SA structure diagram of the module

Lightweight and efficient attention mechanism module uses Shuffle unit (effectively construct multiple branch structure to realize parallel processing of multiple branches, which can reduce the computation amount) to effectively combine two types of attention mechanism, and the function of the module can be summarized as split-transform-integration. Specifically, it is divided into four parts.

Feature Grouping, grouping each channel as a group, grouping the input characteristic Xk1, and dividing the input characteristics into G groups along the channel dimension. According to each group of characteristics, each channel and space requires different importance coefficients to enhance. These Channel and Spatial importance coefficients are derived from the channel and spatial attention modules, known as channel attention and spatial attention, respectively.

Channel Attention extracts Xk1 features using a simple single-layer transformation of GAP+Scale+Sigmoid. First, the two-dimensional channels are compressed into one-dimensional channels along the spatial dimension (X, Y) by global average pooling (GAP), and each two-dimensional channel is converted to a real number. Then, in order to make the features more obvious, the feature map of each channel is enhanced proportionally and achieved through sigmoid activation to obtain the Xk1 enhanced feature map.

Spatial Attention extracts Xk2 features from different viewing angles, uses grouping standardization in deep networks, divides channels into G groups, calculates the mean and variance for each group, and then carries out image enhancement. Finally, The characteristic diagram of Xk2 is obtained through the module of spatial attention mechanism.

Aggregation features. After extracting Xk1 features through channel attention and Xk2 features through spatial attention, the extracted features are fused to get an enhanced feature map. The channel permutation operation is used for inter-group communication, and finally the features extracted by the SA attention mechanism module have the same size as the input features to continue the learning of the following network.

The SA attention mechanism module imposes an "attention" convolutional channel on the residual unit structure of the residual network to form a lightweight and efficient attention residual unit, as shown in FIG. 3, which only shows the fusion relationship between the convolutional layer, the secure connection layer and the average pooling layer.

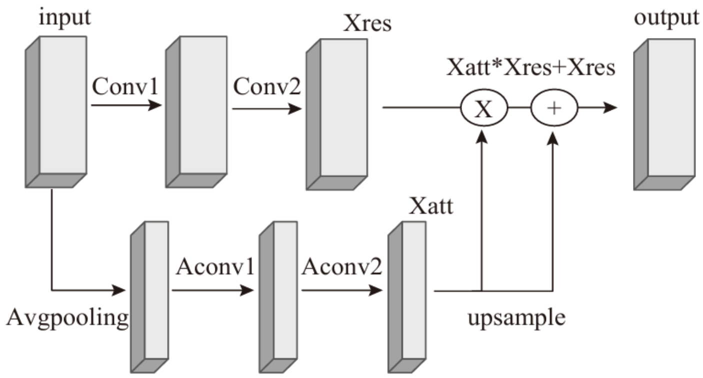

Figure 3. Lightweight and efficient attention residual structure

As shown in Figure 3, an additional path for learning weights is added to the original residual unit to correct the output characteristic graph. First of all, Avgpooling is used to reduce the sample feature map, and then the attention convolution channel (Aconv) is used to obtain the feature map, and then the attention feature map is enlarged by up-sampling to obtain the feature map of the input image.

2.2. Improved Residual Network Combined With Lightweight and Efficient Attention Module

Since epiphyseal of hand bone is a dynamic growth process rather than a discrete change process, and the classification of each grade belongs to the extraction of similar features of fine particles, it is very challenging to determine the grade. In order to better enable convolutional neural network to learn the local features of key epiphyses and improve the accuracy of each epiphyses grade, the residual network is further improved on the basis of combining the lightweight and efficient attention module. The improved residual network is to replace the last two layers of the residual network ResNet with double-pooling layers, divide two branches behind the last residual block of the network ResNet, one branch is average pooling, the other branch is maximum pooling, and then merge the channels of the two branches, and reduce the dimension of the data through convolution 1×1. Then the recognition and classification results are obtained through full connection operation. Combining the efficient attention module is to integrate the attention module into the residual network module on the basis of improving the residual network. This idea can further extract more epiphyseal feature information of fine particles. The structure diagram of the improved residual network combined with the attention module is shown in Figure 4.

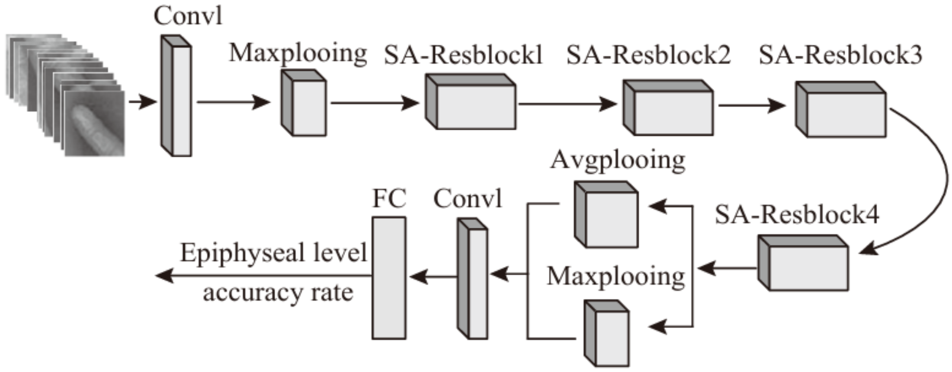

Figure 4. Improved residual network structure diagram combined with attention module

3. EXPERIMENTAL ANALYSIS

3.1. Data Set Collection and Enhancement

The sample data in this paper is RSNA data set, and then multiple images of healthy left hands are obtained by using web crawler and other methods to ensure the average distribution of sample data at different age levels. The left hand bone images of Chinese underage children, including 3315 males and 2910 females, were collected in 14 groups ranging in age from 3 to 16 years. All subjects were healthy and had no congenital disease.

In order to enrich the data set, the data volume of all age groups was evenly distributed. In order to avoid over-fitting during the evaluation, image rotation was adopted to expand the data volume of the samples, so that the samples of all ages were evenly distributed and the important epiphyses of all levels were covered more completely. This study evaluated the degree of epiphyseal development based on TW3 evaluation method. First, 13 key epiphyseal regions in the X-ray images were divided, which were radius, ulna, metacarpal I, metacarpal III, metacarpal V, proximal finger I, proximal finger III, proximal finger V, middle finger III and middle finger V, and distal finger I, III and middle finger V[18]. Then, these epiphyseal regions were graded. The total score of the obtained grade is converted to the corresponding bone age. In order to evenly distribute the number of grades of key epiphyseal, make epiphyseal information clearer and obtain higher accuracy, image enhancement algorithm is used to enhance the contrast and clarity of epiphyseal information, and the enhanced image generated is shown in Figure 5.

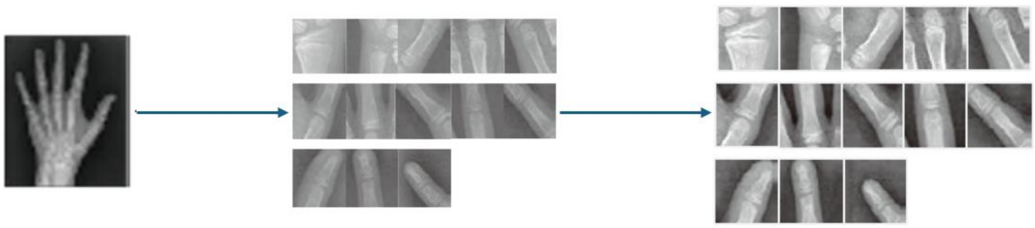

Figure 5. Comparison of epiphyseal images

3.2. Test Environment

This paper takes Windows 10 as the platform, uses Intel Corei5-9400F CPU and GeForce GTX 1660 Ti two Gpus as the core, uses PyTorch framework based on deep learning and Python 3.7. In each training session, the number of training samples is set to 16 and the initial learning rate is set to 0.000 1 to prevent the failure of gradient burst training during training. During the execution of the program, the training sample will be

The samples were randomly separated at a ratio of 8:2.

3.3. Evaluation Index

The results of bone age assessment were analyzed using a new assessment method - mean absolute error. According to age, gender and other factors, the evaluation results were grouped according to age. Within the error range of ±1 year, the assessed bone age of each test sample was compared with the real age to obtain the accuracy and mean absolute error value of epiphyseal age.

3.4. Test Results and Analysis

In order to make the bone age assessment network better learn the information of epiphysis fine particles, the residual network was improved based on the Shuffee Attention (SA) module. In order to verify the feasibility of the improved residual network, the improved ResNet network and ResNet network are compared based on TW3 method, and the results are shown in Table 1.

Table 1. Evaluation result

method	meal		female	
	±One year accurac/%	MAE/ year	±One year accurac/%	MAE/ year
ResNet	92.11	0.42	93.45	0.45
Improved ResNet	94.25	0.40	94.16	0.42

As can be seen from Table 1, improved ResNet network can better evaluate performance than ResNet. Within the 1-year error range, the mean absolute error (MAE) of male and female was 0.422 8 years old and 0.434 1 years old, respectively, and the accuracy of male and female were 93.82% and 93.16%, respectively. Compared with the residual network, the improved residual network significantly improved the accuracy of bone age assessment and reduced the mean absolute error. In addition, in order to verify the feasibility of the bone age assessment method proposed in this paper based on the improved residual network, that is, the improved convolutional neural network -- improved ResNet was compared with the bone age assessment methods proposed by other researchers, such as Hu Tinghong et al. [12], Zhan Mengjun et al. [13], Zhang Shijie et al. [14], Tang Zhihao et al. [8]. Within the 1-year error range, the proposed bone age assessment method achieved higher accuracy and lower mean absolute error.

4. CONCLUSION

In order to extract the characteristic information of epiphyseal fine particles and improve the accuracy of bone age assessment, lightweight and efficient attention module SA was introduced into ResNet network in this paper, and dual pooling fusion was performed to improve the residual network for bone age assessment. The experimental results show that the modified residual network can effectively extract the epiphyseal information of fine particles, improve the accuracy of bone age assessment and reduce the average absolute error. Since the age distribution of the data set is 3 to 16 years old, it is assumed that the data set of 0 to 3 years old is added, so that the samples of 0 to 16 years old are evenly distributed, and the accuracy rate and average absolute error coverage of males and females are more comprehensive.

REFERENCES

[1] Shen Jiyun Evaluation method of adolescent bone age based on NTS deep neural network [j]. digital technology and applications, 2022, 40 (12): 123-128

[2] Li Rui, zhangshijie, Huang Aoyun, et al Evaluation of wrist bone age in adolescents based on deep learning [j]. computer technology and development, 2020, 30 (1): 124-128+134

[3] Wang Lin Design and implementation of bone age calculation and height prediction system based on data model [d]. Hangzhou: Zhejiang University of technology, 2019

[4] BIAN Z, ZHANG R. Bone Age Assessment Method Based on Deep Convolutional Neural Network [C]//2018 8th International Conference on Electronics Information and Emergency Communication (ICEIEC).Beijing: IEEE, 2018: 194-197.

[5] GARN S M. Radiographic Atlas of Skeletal Development of the Hand and Wrist [J]. American Journal of Human Genetics, 1959, 11(3):282-283.

[6] BAYER L M. Radiographic Atlas of Skeletal Development of the Hand and Wrist [J]. California Medicine, 1959, 91(1):53.

[7] MALINA R M, BEUNEN G P. Assessment of Skeletal Maturity and Prediction of Adult Height (TW3 Method) [J]. American Journal of Human Biology, 2002, 14(6):788-789.

[8] Tangzhihao, liulijun, fengxupeng, et al Residual network bone age assessment combined with efficient channel attention module [j]. optoelectronics · laser, 2021, 32 (3): 331-338

[9] Zhanmengjun, zhangshijie, Chen Hu, et al Research progress in automated assessment of bone age [j]. Journal of forensic medicine, 2020, 36 (2): 249-255

[10] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet Classification with Deep Convolutional Neural Networks [C]//Advances in Neural Information Processing Systems 25 (NIPS 2012). CURRAN ASSOCIATION INC., 2012:1-9.

[11] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Imagenet Classification with Deep Convolutional Neural Networks [J]. Communications of the ACM, 2017, 60(6):84-90.

[12] Hutinghong, huozhong, liutaiang, et al Automatic assessment of left wrist bone age of Uygur adolescents based on deep learning [j]. Journal of forensic medicine, 2018, 34 (1): 27-32

[13] Zhanmengjun, zhangshijie, Liuli, et al Automatic assessment of left wrist joint bone age of Han adolescents in Sichuan Based on deep learning [j]. Chinese Journal of forensic medicine, 2019, 34 (5): 427-432

[14] Zhangshijie, Li Rui, zhanmengjun, etc Research on bone age evaluation based on deep learning region fusion [j]. modern computer, 2020 (9): 54-59+68

[15] HE K, ZHANG X, REN S, et al. Deep Residual Learning for Image Recognition [C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR).Las Vegas: IEEE, 2016:770-778.

[16] Pang Sisi, Huang Chengcheng Research on image classification based on convolutional neural network [j]. modern computer, 2019 (23): 40-44

[17] ZHANG Q L, YANG Y B. SA-NetP: Shuffle Attention for Deep Convolutional Neural Networks [C]//ICASSP 2021— 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Toronto: IEEE, 2021:2235 2239.

[18] Gu Jing, Ma Ruiqi, zhuhengan Evaluation method of X-ray image bone age based on convolutional neural network [j]. Chinese Journal of medical physics, 2022, 39 (3): 305-310.