

Research paper

An Ensemble Convolutional Neural Networks for Detection of Growth Anomalies in Children with X-ray Images

Homeyra Sarabi Sarvarani and Fardin Abdali-Mohammadi*

Department of Computer Engineering and Information Technology, Razi University, Kermanshah, Iran.

Article Info

Article History:

Received 14 March 2022

Revised 08 May 2022

Accepted 01 August 2022

DOI:10.22044/jadm.2022.11752.2323

Keywords:

Growth Anomalies Detection, X-ray Images, Ensemble Learning, Convolutional Neural Networks.

*Corresponding author:
fardin.abdali@razi.ac.ir (F. Abdali-Mohammadi)

Abstract

Bone age assessment is a method that is constantly used for investigating growth abnormalities, endocrine gland treatment, and pediatric syndromes. Since the advent of digital imaging, for several decades, the bone age assessment has been performed by visually examining the ossification of the left hand, usually using the G&P reference method. However, the subjective nature of hand-craft methods, the large number of ossification centers in the hand, and the huge changes in the ossification stages lead to some difficulties in the evaluation of the bone age. Therefore, many efforts were made to develop the image processing methods. These methods automatically extract the main features of the bone formation stages to effectively and more accurately assess the bone age. In this paper, a new fully automatic method is proposed in order to reduce the errors of subjective methods and improve the automatic methods of age estimation. This model is applied to 1400 radiographs of healthy children from 0 to 18 years of age and gathered from 4 continents. This method starts with the extraction of all regions of the hand, the five fingers and the wrist, and independently calculates the age of each region through examination of the joints and growth regions associated with these regions by CNNs. It ends with the final age assessment through an ensemble of CNNs. The results obtained indicate that the proposed method has an average assessment accuracy of 81%, and has a better performance in comparison to the commercial system that is currently in use.

1. Introduction

This Developing new technologies in the field of radiology have brought improvements in the diagnosis and treatment of many diseases [1]. However, on a yearly basis, the number of patients and radiology images is 7.5 times more than the number of specialist radiologists [2]. Considering that, it is necessary to examine all these images; it should be possible to carry out the diagnosis process in a short period of time. It is a highly specialized, challenging, and error-prone task to perform a medical image analysis, which requires the radiologist's expertise. However, the excessive workload of the radiologist per day can lead to the misinterpretation of images.

Oftentimes because of the absence of orthopedic specialists working in emergency departments, these kinds of errors tend to increase by approximately 41% to 80%. Clearly, misdiagnosis can have a devastating impact on ensuing functionalities within the patient's treatment period [3]. In order to diminish such problems, it has been suggested to use computers (CAD) to improve the diagnosis. Physicians may employ them as a backup opinion for supporting and assisting their decision-making [4]. Many studies have been made to develop computer-aided diagnosis (CAD) systems to be used in different medical fields including diagnosis of cancer,

diagnosis of bone fractures, brain tumors, and growth abnormalities. Hereditary, hormonal, nutritional, disease-related, and psychosocial factors all influence a child's growth. For this reason, it is possible for children to undergo impairment in the process of normal growth and development [5]. The difference between the so-called ID age and bone age is indicative of a growth abnormality. Age estimation is also applicable for serious crimes such as child pornography, human trafficking, issues concerning asylum and immigration, perpetrators and victims of crimes, adoption of children across national borders, and competitive sports [6]. Therefore, we have seen a significant increase in the demand for age estimation in the last decades. It is possible to evaluate the bone age through examining radiographs of various skeletal limbs but the most frequent examinations are those of the left hand and the clavicle bone. The reason for recommending the use of hand radiographs is simplicity, minimal exposure to x-ray radiation, and the availability of multiple ossification centers [7].

The most common clinical methods of estimating the bone age from left hand radiographs are GP (Greulich & Pyle) [8] and TW (Tanner-Whitehouse) [9]. GP is a straightforward, atlas-based method. In this method, the X-ray image of the client is compared with a set of standard and reference images taken from healthy girls and boys of different ages. The image that most closely matches the image of interest is considered to be the bone age. The TW method is more complex, more accurate, and better than the GP method. This method is based on separately classifying (rating) 20 hand bones. The developmental stages of each bone will be divided into 9 stages (from early childhood to adolescence, or from A to I), and the sum of the scores will be used to estimate the age. Next versions of this method are TW2 and TW3 [10, 11], which similarly are scoring-based. The above-mentioned clinical methods are heavily dependent on the observer. Therefore, a great deal of effort has been made to automatically estimate bone age with the help of CAD systems. The CAD systems employ the features that are used in clinical methods to estimate the age. With regard to extracting features, the CAD systems can be divided into two distinct categories: 1. Feature extraction in hand-craft way, 2- Feature extraction with deep learning. In the first one, before a feature is produced, an expert has to identify a suitable feature and then extract it by hand-crafting. The most important of these features

include Gaussian mixture model [12], canny edge detector, mean, variance, Particle Swarm Optimization (PSO)[13-15], Gabor filters, and Histogram [15]. There are two main problems in methods of the first category. 1- These systems cannot wholly cover the problems of diagnosis and classification in a single framework. 2- Producing features by hand-craft necessitates a field knowledge of the expert [14]. The researchers have shifted to the second category in the recent times, which is the creation of features with the help of deep learning. Deep Learning approach that needs no hand-crafts or knowledge of the expert and lessens the distance between human vision and computer vision for pattern recognition achieved much better results in classification than the other conventional methods in computer vision as well. There have been many studies in the field of medicine that have been conducted using deep learning, and have yielded remarkable outcomes. For example, Convolutional Neural Networks (CNNs) have been applied in predicting the division of brain tumors [16, 17] and the diagnosis of lung nodules [18]. CNNs have also been utilized in some cases for diagnosing bone age, and have provided acceptable results.

Despite the existence of deep learning methods for estimating bone age, there are the following challenges associated with them: 1. Neither of the existing systems is truly based on the TW method. 2. None of the existing methods extracted all fingers of the hand, and did not employ the combination of scores of different parts of the hand. 3. In each of these methods, a CNNs constructing architecture is used but one should note that the construction of a proper architecture for CNNs depends on the knowledge level of the designer.

In this paper, we propose a fully automatic system on the basis of TW and radiographs of the left hand using deep neural CNNs for the process of bone age estimation. The proposed system is composed of three steps: 1. preprocessing, 2. feature extraction, 3. classification. In the preprocessing stage, the hand is divided into six regions in the first step, as shown in Figure 1; region 1 is assumed to be for the wrist bones (carpal bones), and regions 2 to 6 are for the fingers. Then these processes are carried out to extract the regions. Firstly, the extra sections of the images were removed, and the hand mask region was extracted from the image. Then. an algorithm was developed to locate the coordinates of the points on the two edges of the wrist and the midpoint of the hand and with its help the region 1

was extracted correctly. Then using the convex-hull function as the edge finding technique and calculating the slope of the line between the boundary points of the image, the coordinates of the finger points were obtained, and through its help, the finger-related region was also extracted. In the next step, in order to extract the feature, the following deep learning technique was used: each region was assigned to its own CNN, and each network, when extracting features of each region, also estimated the probability of age in that region. The final classification was done by an ensemble of CNN neural networks of the six regions of the hand.

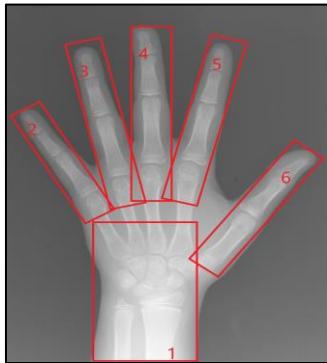


Figure 1: Segmentation of hand to 6 regions: region 1 for wrist, regions 2-6 for fingers.

Finally, the main innovations of this paper include:

1. Developing a comprehensive system for the integrated estimation of bone age across all growth ages, from zero to 18 years of age, and across all races.
2. The method proposed in this work correctly identifies and extracts all parts of the hand (joints and fingers), and determines the final scoring based on the scores of these regions.
3. Another feature that makes this method superior is that it uses an ensemble between CNN models, which turns it into a majority algorithm.

The automated methods presented in the estimation of bone age are discussed below in Section 2, and then the proposed method of this article will be described in Section 3. Section 4 also addresses the evaluation of the performance of the proposed method compared to other methods. Finally, in Section 5, conclusions and future orientations are outlined.

2. Related Works

A comprehensive review of computational methods for BAA is provided in [19], which shows that although the amount of automated methods for BAA has increased, these methods are still in their early phases of development.

Furthermore, noise in images and incomplete data or poor contrast are major problems in automated age estimation. In [20], the boundaries of the bones are manually labeled using tagged points; in [21], the top and bottom PROI boundaries (regions include finger bones or phalanges and secondary growth regions or epiphyses) are also identified using the scanning of two horizontal lines. Estimation of bone age in these methods was based on the extraction of such features as bone length, width, and region using a kernel that is based on a linear search in the phalanx length reference table and the Gaussian mixture model, respectively. The main drawback associated with these methods was that user intervention was required. In the method [22], the bone age was estimated using the bone regeneration algorithm (AAMs-for Active Appearance Model) and the TW and G&P methods. This system that is called BoneXpert dismisses low quality images or bones with abnormal structure. The analysis must be done manually in these cases. Another method based on neural networks is that of [23]; in this method, the bones are reconstructed using hand-craft points and adaptive clustering is used for segmentation.

Methods [24] and [25] also estimated bone age through analysis of the joints of the hand (phalanx of the third finger, wrist, and radius-ulna bone). They have used the description of change modes, covariance matrices, and PSO algorithms for the statistical measurement. For teaching the models, they used a neural network (that begins with an unknown state), and the BP algorithm, respectively. Method [26] is a web-based method, and employs the histogram technique for the assessment. Despite its good performance and its ability to overcome the segmentation problem, it still rejects poor quality images. Its evaluation was also carried out on a small basis with 32 random images. In a fully automated method based on the GP (Greulich & Pyle), the reference model is presented for the age range of 5 to 18 years. This model proposed a new pre-processing method that includes identifying and isolating the hand and wrist, and then standardizing image features using CNNs. In [27], the first stage, following the normalization of the color and size of the image, CNN (with LeNet-5 architecture) performed the identification and separation of the required parts. Afterwards, a CNN (with GoogLeNet architecture) was used for teaching the model and performing the classification. Although this system has a good potential, it is not generalizable to all age groups because it removed the age group from 0 to 4 years old. In the work [28], all

growth ages have been taken into consideration, and various deep learning methods have been proposed and tested for assessing the bone age. The general architecture in this model is as follows: a convolutional neural network for extracting features, a regression network with fully connected layers, and an output layer for providing bone age assessments. In this method, a number of pre-trained neural networks (OxfordNet, GoogLeNet and OverFeat) are initially being tested on the dataset. GoogLeNet shows a better performance than the other networks. In addition, a custom CNN is developed from the beginning consisting: five convolutional layers for extracting low and medium level visual features, a deformation layer preceding the last convolutional layer for face non-rigid object

deformation, and a completely connected layer for regression of bone age, which has performed better than GoogLeNet and achieved an accuracy of 79%. This system has not performed any pre-processing on the images. The model [29] that was introduced in 2019 is based on the reference model of TW, and uses a private database of 3,300 radiographs from the left hands of Korean children under the age of 18 years. This model conducts the age identification process in three steps: A. Extraction of fingers (thumb, third finger, and fifth finger) and wrist with image processing techniques. B. Extraction of joints from the aforementioned segments (using the Faster R-CNN architecture). C. Classification and assessment of the age. The accuracy of this method is 79.6%.

Table 1. A comparison of automated approaches in BAA.

Reference	Dataset	Method	Advantage	Limitation
M.Niemeijer et al. [24]	Private-243 Image	Phalanx of the third finger, correlation coefficients, active shape model	Accuracy 73 to 80%	Age range is 7-16 years old, for Belgian children, test with 71 images
J. Liu et al. [25]	Private-1046 Image	Based on intelligent algorithms (PSO and BP algorithms), Bones of radius, ulna and carpal	High compatibility rate in comparison to hand-craft method	Having a high loading time in image processing
A. Tristán et al. [23]	Private-158 Image digital and analogical	Segmentation By Adaptive Clustering And Neural Network	Improving the pattern recognition procedure for bones	Limited to four levels of TW3. Use hand-craft points.
H. Thodberg et al. [22]	Private-1559 Image	Base On Bone Reconstruction Algorithm(AAMs,Shape Driven), And GP & TW Model	High accuracy, Based On Two Clinical Methods (GP and TW)	Rejecting low quality images, age range 2-17
M. Mansourvar et al. [26]	Public-1100 Image	Based on the Histogram And Content-Based Image Retrieval (CBIR) Technique, Web-based system	Overcomes The Segmentation Problem	Unreliable for low quality images or bones of abnormal structure, evaluation and test on a small scale.
D. Giordano et al. [30]	Private-360 Image	Gaussians (DoG) filter, Hidden Markov Model (HMM)	High Accuracy 95%	Limited Areas Are Evaluated, Only In The Range (0-6)Years
H. Lee et al. [27]	Private-8334 Image	CNN-LetNET	Standardization of all kinds of hand radiographs in different formats.	Not Applicable For Children Below 4 Years, Usage of Integer-Based BAA
C. Spampinato et al. [28]	Public-1391 Image	CNN-BoNet	Tested Several Existing Pre-Trained Convolutional Neural Networks On A Dataset	No Preprocessing On Images
J. Son et al. [29]	Private-3300 Image	CNN-VGGNet	Introducing the ROIs concept in order to decrease the region, high accuracy.	The Model Database Was For Korean Children Only.
X. Chen et al. [31]	Private-12536 Image	CNN-ResNet - SVM	Introducing a new deep neural network (ST-Res) model based on Res-Net and Spatial Transformer.	Only for Chinese children
H. Lee et al. [2]	Private-3000 Image	CNN-GoogleNet ResNet CaffeNet	Using a set of deep learning architectures	Manually marking the points on the image in order to find the hand region and remove the background
A. Senel et al. [32]	Private-150 Image	CNN- GoogleNet AlexNet VGGNet	Achieving high accuracy	Small amount of data - not confirming the growth health of the individuals to whom the images belong.

In method [30], the bone age is calculated in the age range 0-6 years through examination of the epiphyseal, metaphyseal, and diaphyseal regions (in the knuckle and palm bone regions-metacarpal), of the thumb, 3rd finger and 5th finger (little finger). The system has been implemented on a set of 360 radiographs (180

images for males and 180 images for females). MAE for this model amounts to 0.37% at best. The pre-processing part of this model consists of noise removal (for increasing the image resolution), background removal with the flood-fill algorithm, and hand alignment correction. Then the wedge functions are applied to separate

each finger and find the fingertip, midpoint, and base of the finger. Then the soft tissue of the image is removed using a Gaussian (DoG) filter. The final classification has been done by Hidden Markov Model (HMM) [33]. In this model, the parameters that are evaluated are only the identification of joints in three fingers (thumb, third, and fifth) but in order to obtain a more accurate evaluation, we should also consider the carpal bones and the wrist joints (radius and ulna bones) since they are the most important parameters for evaluation at young ages. The evaluation is also limited to a specific age group (0 to 6 years old). These models have been implemented on private and inaccessible datasets, and we do not have access to their source code, therefore, it is not possible to reproduce the results and generalize them to all races and age groups. Method [31] is carried out on 12536 images that were collected in China. In this method, sensitive regions in the pre-processing stage have been extracted by RCNN and then has tested three networks namely SVM, ResNet and the proposed network of itself named ST-Res (based on ResNet and Spatial Transformer) on its dataset. The best accuracy was obtained by ST-Res (78.4%). In methods [2] and [32], bone age assessment has been also carried out by applying a set of deep learning architectures. In [2], the pre-processing stage has been carried out by marking points on the image and extracting the wrist region and a part of the finger.

3. Proposed Method

As we can see in Figure 3, the proposed system for bone age identification is composed of three steps. In the first step (pre-processing), composes the extraction of the areas marked in Figure 1 (wrist and fingers). We call these regions as regions of interest (ROI), and they are extracted using image processing techniques. In the second step, feature extraction and determination of the maturity level of each ROI is measured with the help of a deep neural network with a pre-designed architecture. In the third step, the bone age of the person is estimated using the scores that have been determined in the previous step. This is achieved using an ensemble between the CNNs that were used in the previous step, making the final decision on the basis of the majority across the networks. In the next section, each of these steps is described.

3.1. Step 1: Extracting ROIs

In order to locate the finger and wrist regions, the first step is to extract the main region of the hand

from the whole image. To do this, the input image (Figure 3.1), which is in grayscale form, has been converted to binary (Figure 3.2). In this section, a thresholding algorithm has been used to turn the image into binary. This algorithm compares the intensity of each pixel with the average intensity of the entire image; if this intensity is smaller than the average, it assigns zero to that pixel, and otherwise, it assigns one to that pixel. In some images, besides the main part of the hand, there are other regions in the image like labels and extraneous parts that are also indicated in the binary image. These extraneous parts affect the process of extracting the regions. Therefore, it is necessary to remove these parts and extract only the hand area. For this purpose, the Regionprops function has been used; this function calculates the size of the different regions of the input image and returns the region that has the largest area. As we can clearly see in the image, the main part of the hand has the largest area. The output image of this function can be seen in Figure 3.2. It can be seen that only the largest region (main part of the hand) has been extracted at the output. Next, to correct the direction of the images using the coordinates obtained from the image edge-finding, the middle coordinates of the wrist are obtained and the points are marked on the image, then a Hough algorithm is employed to draw a straight line from the middle of the wrist Figure 2.a. The direction of the hand is determined using this straight line Figure 2.b, and to correct the direction of the image, the original image is rotated in the direction of the resulting line Figure 2.c.

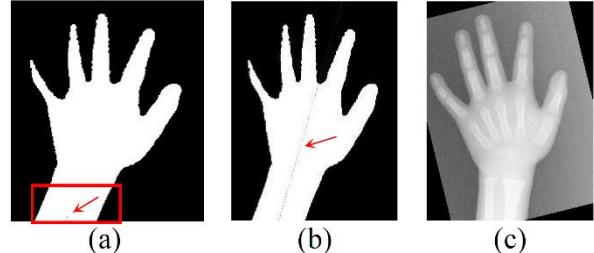


Figure 2. Adjust angle of rotated images.

3.1.1. Text Font of Entire Document

After extracting the main part of the hand, in order to find the coordinates of the wrist (carpal bones), the boundary pixels (edges) of the hand are indicated on the binary image using the edge detection technique (Figure 3.3). Then an algorithm has been created to record the coordinates of the edge pixels of the hand. Using these coordinates, two lines are drawn vertically, parallel to the edges of the wrist (Figure 3.4). A horizontal line is then drawn to the center of the hand using the coordinates of the center of the

hand. These lines are used to divide the main area of the hand. Finally, we can see the extracted part of the carpal section in Figure 3.5).

3.1.2. Extraction of finger ROIs

In order to extract finger ROIs, the fingertip coordinates are initially determined, and then the finger region is identified and extracted from it. For solving the challenge of finding the fingertip coordinates, the idea used is to create a convex region and find the angles of this region Figure 3.6. Drawing lines between the fingers connects each finger to its neighboring fingers, resulting in a geometric shape. Since the shape of interest resembles a convex, a convex computation algorithm was used to draw it, which is a subset of computational geometry. The details of the convex implementation algorithm are as follow (Algorithm 1). The input of this algorithm is the coordinate matrix of the points around the hand image obtained from the edge-detected image. The algorithm adds the coordinates of the points to the Lupper list from the very beginning. It calculates the slope of the lines passing through all three consecutive points (three points in the Lupper list), and in the case they are in the same direction, it removes the middle point and adds the next point to the Lupper list. This procedure will continue until it reaches a break (where the points are not in the same direction). After the break, the points in the Lupper are regarded as a line and are added to the Llower. Then the next

points are added to Lupper. The same procedure also takes place by the Llower list and in the opposite direction to the list of points. The output of the algorithm is a list of interconnected lines (L). As it can be seen in Figure 3.6 and Figure 3.7, the points associated with the fingertips are located at the vertices (the break between the lines) of the convex region, and to determine their coordinates, the coordinates of the vertices of the convex region must be determined. For this purpose, an algorithm is used to calculate the chord angle between the points (Algorithm 2). The input of this algorithm is the matrix of coordinates of the boundary points of the convex region obtained from the edge-detected image. This algorithm first draws a line between two consecutive points, and then calculates the slope of this line. The output of this algorithm consists of five points with the largest angles that have exactly the same fingertip coordinates (Figure 3.7). After the coordinates of the fingertips have been determined, two points are defined as marker points (from the fingertip to the center of the hand or from the fingertip to the center of the wrist) and a line is drawn between these two points. Then the image is rotated on the basis of this line (Figure 2.8 and Figure 3.9) until the intended finger is in a straight position. The region associated with each finger is identified according to the fingertip's coordinate (Figure 3.10 and Figure 3.11) and after each ROI is determined, then the cropping has happened (Figure 3.12).

Algorithm 1. (Convex Hull) – Pseudo code –

Input. A set of points in the plan.

Output. A list containing the vertices of $\text{conv}(P)$.

1. Sort the points by (x,y)-coordinate, resulting in a sequence P_1, P_2, \dots, P_n .
 2. Put the points P_1 and P_2 in the list L_{upper} with P_1 in the first position.
 3. For $i \leftarrow 3$ to n
 4. **do** Append P_i to L_{upper}
 5. **While** L_{upper} contains more than two points **and** the last three points in L_{upper} do not make a right turn
 6. **do** delete the middle of the last three points from L_{upper} .
 7. the same in the opposite direction: P_n to P_1 Make L_{lower} .
 8. Remove the first and the last point from L_{upper} .
 9. Append L_{upper} and L_{lower} to make L .
 10. **Return** L
-

Algorithm 2. (slope of chord) – Pseudo code –

Input. a sequence P_1, P_2, \dots, P_n .

Output. A list points of L_{result}

1. Sort the points by (x,y)-coordinate, resulting in a sequence P_1, P_2, \dots, P_n .
 2. Put the tan points P_i and P_j in the list L_{deg} .
 3. **For** $i \leftarrow 1$ to n
 4. Put the point $i(x,y)$ to X_1 and Y_1 and Put the point $i+1(x,y)$ to X_2 and Y_2
 5. Put the tan $((Y_1 - Y_2)/(X_1 - X_2))$ to $L_{\text{deg}}(i)$
 6. **For** $i \leftarrow 1$ to n
 7. if $L_{\text{deg}}(i) > \mu$ put the $L_{\text{deg}}(i)$ in the L_{result}
 8. **Return** L_{result}
-

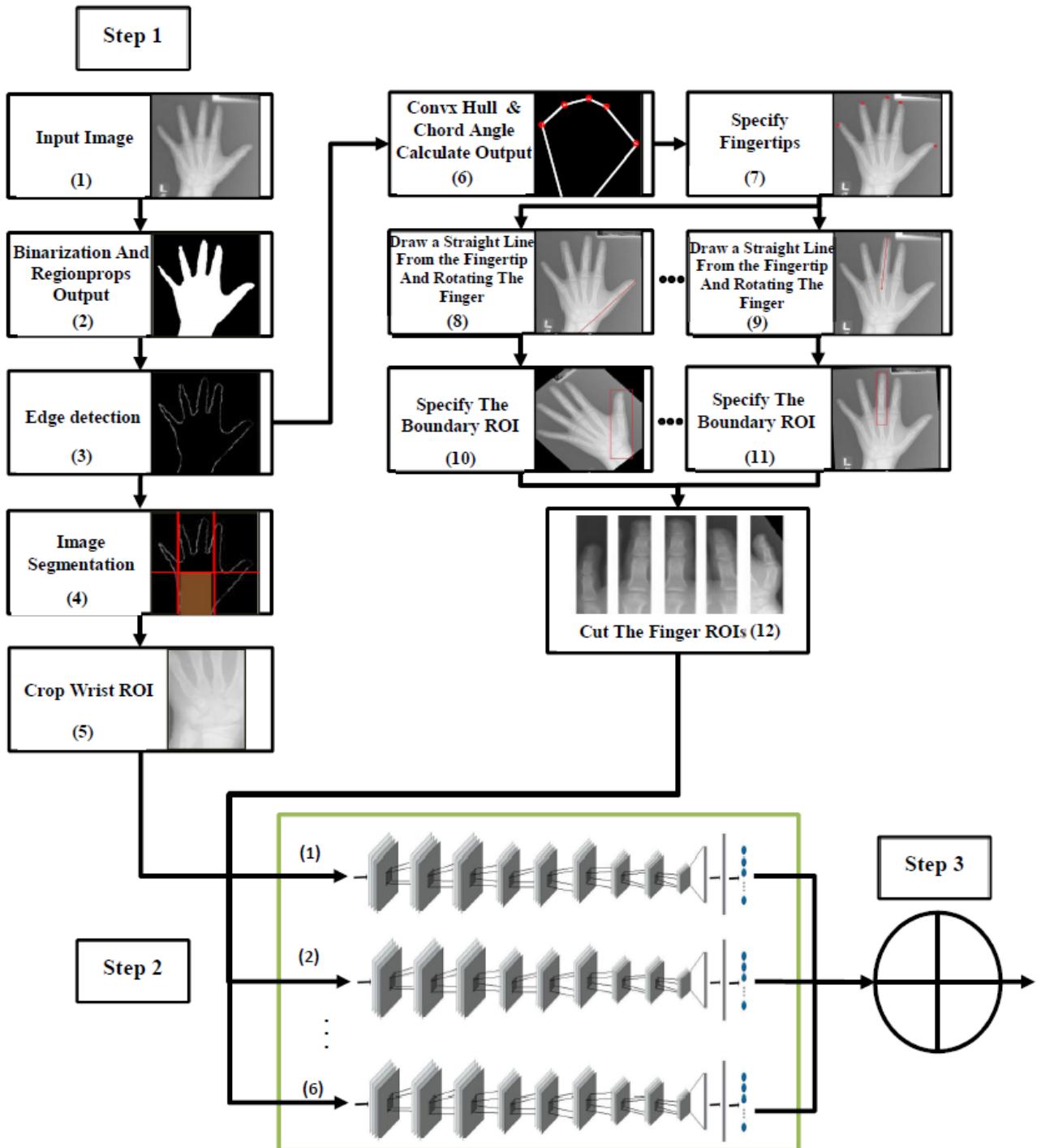


Figure 3. Flowchart of proposed method. Step 1: Pre-processing (extraction of ROIs); Step 2: Automated extraction of proposed neural network features and architecture; Step 3: Final age assessment by majority voting ensemble under CNNs.

3.2. Step 2: Architecture of proposed CNN

Each ROI has its specific and unique features that show the degree of bone maturity. For this reason, feature extraction and bone age assessment in each ROI is performed by an independent CNN. The architecture specified in Figure 4 has been used as the architecture of the main CNN. The first convolution layer has 32 filters of size 3x3. Two more convolution layers follow, each with 64 filters of size 3x3 [34, 35]. Next is the first pooling layer, which uses a max-pooling

performance of 2x2 size. Then we have the fourth and the fifth convolutional layers with 64 filters of size 3x3 and a max-pooling layer with a core of size 2x2. The sixth convolutional layer also consists of 64 layers with a core similar to the previous convolutional layers. The third max-pooling layer is also implemented after the sixth convolutional layer. The details of the CNN architecture are shown in Table 2.

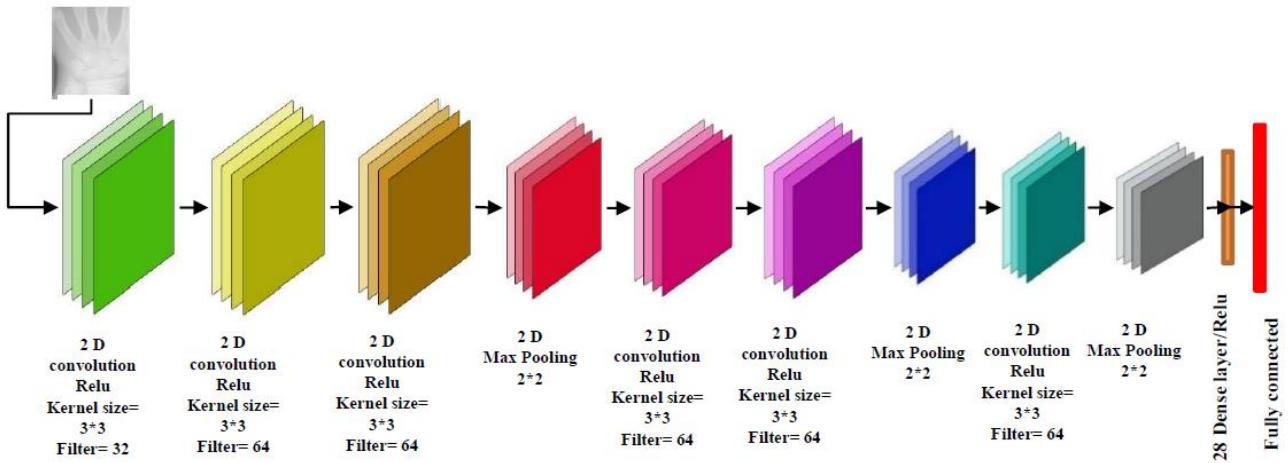


Figure 4. Architecture of proposed neural network.

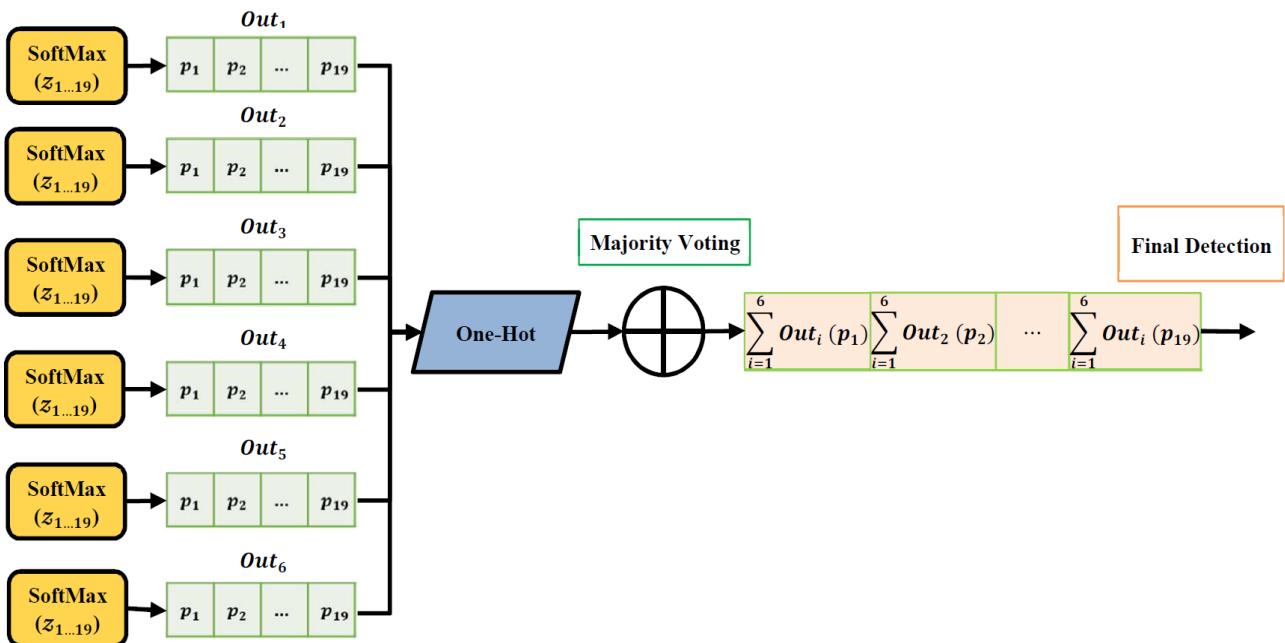


Figure 5. Ensemble algorithm based on Majority vote.

Table 2. Convolutional and pooling parameters in proposed CNN.

Layer	Kernel sizes	Number of filters	Parameters
Convolutional Layer 1	3×3	32	Stride = 1, Padding = 1 Activation = Relu Batch Normalization with $\epsilon=1.001\times10^{-5}$
Convolutional Layer 2	3×3	64	Stride = 1, Padding = 2 Activation = 'relu'
Convolutional Layer 3	3×3	64	Stride = 1, Padding = 2 Activation = 'relu'
Pooling Layer 1	2×2	-	Stride = 2
Convolutional Layer 4	3×3	64	Stride = 1, Padding = 2 Activation = 'relu'
Convolutional Layer 5	3×3	64	Stride = 1, Padding = 2 Activation = 'relu'
Pooling Layer 2	2×2	-	Stride = 2
Convolutional Layer 6	3×3	64	Stride = 1, Padding = 2 Activation = 'relu'
Pooling Layer 3	2×2	-	Stride = 2
Dense 1	128	-	Activation = 'relu'
Dense (output)	19	-	Activation = 'softmax'

3.3. Step 3: Identification of BAA

An ensemble of CNN models was used for the ultimate assessment of the bone age (Figure 5). The ensemble uses the majority algorithm (based on majority vote) [36, 37]. The general classification consists of 19 classes (from 0 to 18 years of age). The output of each CNN in the previous section represents the probability of belonging to one of the 19 classes (z_1, z_2, \dots, z_{19}). The output of each CNN is given to a softmax function. For each image, this function determines the normalized probability function of all parts of the hand (out 1, ..., out 6). Afterwards, for each ROI, the one-hot algorithm is used to assign 1 to the class with the highest probability and 0 to the others. Then the probability of belonging to a class is determined through summation of the probabilities of that class in each of the 6 regions, and the class with the highest probability is considered as the final assessment.

4. Experimental Results

4.1. Dataset

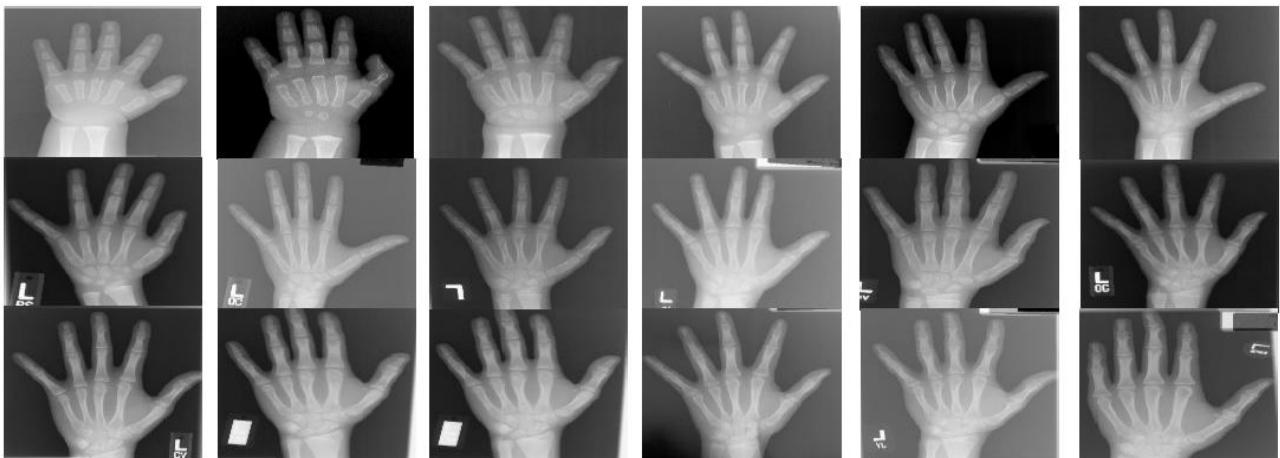


Figure 6. A sample of images from database (growth process of ossification centers).

Table 3. Distribution of images in dataset by race and age.

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Asian	3	10	10	10	10	17	12	14	14	14	29	27	29	30	25	20	20	20	20
Black	9	10	10	10	10	18	16	18	21	19	27	25	30	30	26	20	20	20	20
Caucasian	6	10	10	10	10	17	15	17	19	15	23	27	28	25	21	20	20	20	20
Hispanic	5	10	10	10	10	19	19	20	19	20	26	29	30	30	28	20	20	20	20
All																			
															1400				

4.2. Primary parameters of CNN

The primary parameters of CNN are as follows: 1. Learning rate. Determining the learning rate is very difficult and demanding. If a small value is chosen for the learning rate, the algorithm may remain in local minima, and the network may not be trained properly; if a large value is chosen for the learning rate, the network may enter an oscillatory and unstable state, and as a result, may

The Digital Atlas (DHA) Database [38] system is available at <http://www.ipilab>. This dataset is a public and comprehensive database for the evaluation of automated bone age assessment methods. It consists of 1400 radiographs taken from the left hand of children aged 0 to 18 years old, which have been classified according to gender and race. Each image also provides two values for bone age suggested by two radiologists. A sample of the images is shown in Figure 6, and the distribution of the images is shown in Table 3. The data division is carried out as follows: 70% for train, 15% for validation and 15% for testing. Also due to the small number of images for training ‘cnn’ model, image data augmentation is applied. Each image is rotated, transformed, and resized, resulting in total samples to be upgraded to 5600 images for each cnn. All implementations of the proposed method were performed in the Python (TensorFlow and the Cross Library). The implementations and evaluations were performed on a lap top with the following features: Asus Core i7, 12 Gigabyte RAM, NVIDIA MX130 graphics.

not converge and not be trained-batch size. This parameter is suitable for preventing the model from overfitting. 3. Number of training epochs: The number of epochs is used for training the network. This parameter is very sensitive [39], and if the number of epochs is excessively high, the network may over-fit. In the model proposed in this work, these parameters were quantified by several experimental tests (see Table 4).

Table 4. Optimal configuration for parameters of convolutional neural network.

Parameter	Best value
Learning rate	0.0050
Learning rate decay	0.96
Batch size	400
Number of training epochs	1000

4.3. Assessment criteria

The criteria used to evaluate the proposed method include F-measure, precision, accuracy, recall, and MAE [40, 41]. MAE is calculated as the sum of absolute errors divided by the sample size (the difference between the predicted value and the actual value). The formula for each one of these criteria is as follows, where x_i is the label and y_i is the estimated bone age.

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + FN + TN} \quad (1)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (2)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (3)$$

$$F1 = \frac{2(\text{Recall} * \text{Precision})}{(\text{Recall} + \text{Precision})} \quad (4)$$

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - x_i| \quad (5)$$

4.4. Evaluation of proposed model

The loss function diagram of the proposed model is depicted in Figure 7. This figure illustrates the degree of loss function during the training and validation phases based on MAE. As illustrated in the figure, the MAE value decreases during training and validation, and the network eventually reaches convergence.

In the following, the accuracy level of the networks with different structures (from 1 to 12 layers) were tested on images for different races using multiple experiments. As it could be seen in Figure 8, the networks with a smaller number of layers did not produce acceptable results. Adding more convolutional layers also increased the accuracy of the configuration. Finally, a network with 9 convolutional layers achieved an acceptable level of accuracy (81%). Adding more convolutional layers (12 layers) not only did not increase the accuracy of the configuration but the accuracy level was even lower than in the case with 9 layers. Next, the accuracy of the proposed method was investigated for different epoch numbers from 100 to 1100. These examinations were carried out separately for different races. Taking a close look at the diagram in Figure 9, we can see that as the number of epochs increases, the accuracy also increases, and at the end, the network with 1000 epochs provides the highest

accuracy and the training model reaches convergence at this point. Therefore, the number of training epochs was set to 1000 for all experiments and scales.

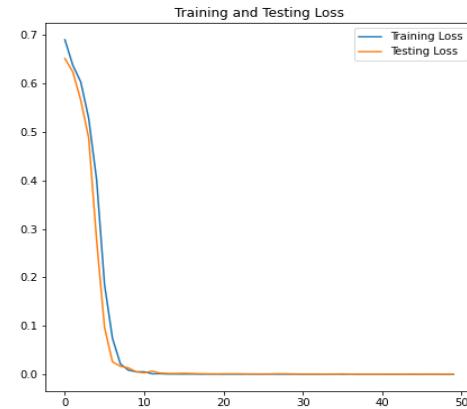


Figure 7. Loss function diagram of training and validation phases according to MAE

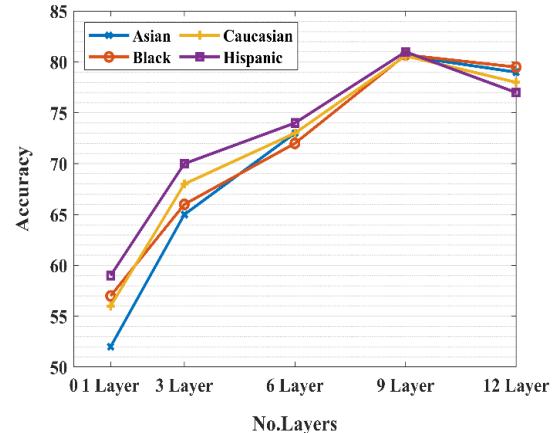


Figure 8. Diagram of evaluating number of layers in proposed model. This diagram shows results of networks with different structures on a correct basis for images of different races (Asian, African, European, and American).

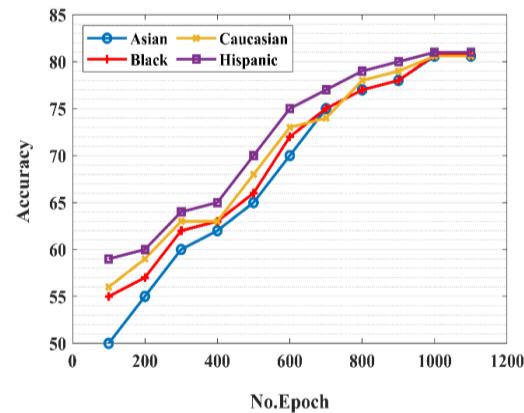


Figure 9. Diagram of evaluating number of training epoch. This diagram shows results of networks with different structures on a correct basis for images of different races (Asian, African, European, and American).

Table 5. Performance evaluation of proposed method by age and race. 19 age group between 0 and 18 years, and 4 races: Asian, African, European, and American.

Age	Asian (%)	Black (%)	Caucasian (%)	Hispanic (%)
Age 0	80.1	80.6	80.4	79.5
Age 1	80	78.8	79.1	78.4
Age 2	81.2	82	81.6	80.6
Age 3	80.9	80.5	80.4	81.9
Age 4	80.7	80.9	81.7	82.8
Age 5	81	81	82.1	83.6
Age 6	80.9	80.9	82.5	81.5
Age 7	79.9	78.2	78.6	80.4
Age 8	78.9	77.7	79.4	79.8
Age 9	83.4	82.9	84.2	82.9
Age 10	81.5	80.6	80.5	82.7
Age 11	80.6	80.8	79.2	81.3
Age 12	78.9	78.1	79.9	79.2
Age 13	82.3	82.9	81.2	81.1
Age 14	81.6	81.5	82.4	79.9
Age 15	82.9	81.9	81.6	79.8
Age 16	83.4	83.9	82.2	82.3
Age 17	77.9	77.4	79.5	79.6
Age 18	78.6	78.2	79.7	80.4
AVG	80.77	80.46	80.85	81

In the following, the performance of the proposed method is evaluated against the data according to the accuracy criteria. There are 19 age groups from 0 to 19 years. The images of each age group correspond to four different races. In Table 5, performance of the model has been computed and reported for all groups. As it can be seen in the table, the average accuracy for the performance of the proposed model is calculated for each race, and the average value of these average accuracies is 81%. This shows the efficiency of the model on X-ray images for different races.

Table 6. Evaluating performance of proposed method for different age groups according to criteria of MAE, accuracy, precision, recall, and F1.

Age	MAE (%)	Accuracy (%)	Recall (%)	Precision (%)	F1 (%)
Age 0	0/113	80.1	81.3	81.4	81.3
Age 1	0/094	79	79.8	80.3	80.04
Age 2	0/083	81.3	80.6	80.5	80.5
Age 3	0/098	80.9	81.8	81.3	81.5
Age 4	0/124	81.5	80.4	80.4	80.4
Age 5	0/107	81.9	80.6	81.8	81.1
Age 6	0/087	81.4	82.8	83.6	83.1
Age 7	0/096	79.2	78.9	79.9	79.3
Age 8	0/12	78.9	79.5	80.5	79.9
Age 9	0/094	83.3	82.2	82.1	82.1
Age 10	0/13	81.3	80.1	81.7	80.8
Age 11	0/099	80.4	81.6	80.3	80.9
Age 12	0/109	79.2	80.7	80.8	80.7
Age 13	0/094	81.8	82.9	80.2	81.5
Age 14	0/098	81.3	82.6	81.4	81.9
Age 15	0/091	81.5	82.5	82.5	82.5
Age 16	0/131	82.9	81.6	83.3	82.4
Age 17	0/089	78.3	89.2	78.4	83.4
Age 18	0/103	78.9	79.7	80.2	79.9
AVG	0/103	80.68	81.51	81.08	81.22

Then the performance of the proposed CNN model is presented for all images by age (regardless of race) according to the criteria of

accuracy, precision, recall, and F1. As Table 6 shows, the average values of the different criteria for all age groups are as follows: 80.68% for the accuracy criterion, 81.08 for the precision criterion, 81.51% for the recall criterion, and 81.05% for the F1 criterion. This shows the efficiency of the model on radiographs for all age groups from 0 to 18 years old. Therefore, the advantage of the proposed system is its good performance for all images with different races and different age groups, and it is not dependent on a specific age group or a specific race.

4.5. Performance evaluation of proposed method with standard deep learning architectures

In Table 7, the performance comparison of the proposed method with different types of standard deep learning architectures is included. As shown in this table, there is not much difference between the ‘cnn’ model and the standard methods. Also the advantage of using the simple ‘cnn’ model is that compared to other models that have different complexities and parameters, it has less computational complexity

Table 7. Performance comparison of proposed method with different types of standard architecture.

	Accuracy(%)	Precision (%)
VGG Net	80%	79%
ResNet	78.12%	78.89%
GoogleNet	79.1%	78.16%
Proposed method	81%	80.18%

4.6. Comparison of proposed method with other advanced methods

In the last two decades, many automated methods for bone age assessment have been proposed. However, only a limited number of these methods were implemented on the digital hand atlas database. The method proposed in this paper was performed on all 1400 radiographs in this database, which in itself demonstrates the comprehensiveness of the proposed method.

Among the methods that use the digital hand atlas database, methods [42] and [26] use 1100 radiographs from this database. These two methods use SVM and histogram techniques, respectively.

Another proposed method is made by [43], which was proposed in 2007. This method, applied to all 1400 radiographs of the digital hand atlas database and using a fuzzy classifier, achieved an accuracy of 71%. The method [28] was introduced in 2017. It implemented two networks, GoogLeNet and OxfordNet, as well as its own architecture called BoNet on all radiographs of the

digital hand atlas database and the best accuracy was achieved by its own BoNet network with 79.0%.

Among the recently introduced methods, The classification performed by Reference [44] using the three pre-trained models of VGG-19, GoogleNet, and AlexNet has achieved 67% accuracy in the best case. Combining a convolutional network and an SVR network, [45] investigated the growth status of six ROI regions. This method has, in the best case, achieved an average MAE of 0.59%.

The method proposed in this article stands next to all these methods, while it has used all the radiographs of the digital hand atlas database. This method is based on the convolutional neural networks, and has an average accuracy of 81% accuracy of 81% and MAE 0.103%. Table 8 shows the comparison of the proposed method with other state of the art methods in the DHA dataset.

Table 8. Comparison of proposed method with other advanced methods on DHA dataset.

Reference	Method	No.Image	Age	MAE (%)	Accuracy (%)
M. Kashif et al. [42]	SVM	1100-DHA	0-18	0.605	-
A. Gertych et al.[43]	Fuzzy classifier s	1400-DHA	0-18	-	79
M. Mansour var et al. [26]	HistogramTechnique	1100-DHA	0-18	0.170	-
C. Spampinato et al. [28]	CNN	1391-DHA	0-18	-	79
A. Ding et al.[44]	CNN	1400-DHA	0-18	0.59	-
D. Bui et al. [45]	CNN+S VR	1400-DHA	0-18	-	67
Propose d method	CNN	1400-DHA	0-18	0.103	81

5. Conclusion

In this work, an effective algorithm for a fully automated bone age assessment was proposed. The proposed method is based on the study of growth and ossification stages in all regions of interest (ROIs) of the hand. Feature extraction and age assessment in each ROI was conducted by an independent CNN. The final conclusion was also drawn from an ensemble of CNN models that used the majority algorithm. Classification of test and training data was also based on a cross-validation algorithm. Several cases were

examined in the results section of this work. At first, the key parameters of this method were investigated. For this purpose, the appropriate number of layers for neural networks was investigated. It appeared in these results that the number of 9 layers was the best status. Then the proper number of epochs for a convolutional neural network was investigated. In this analysis, it was determined that 1000 epochs were sufficient for training. A higher number of epochs had no effect on improving the method. Then in the section on evaluation of the proposed method, the results were evaluated based on the 5 criteria of accuracy, precision, recall, F1 and MAE. For example, the proposed method resulted in an of 81% for the accuracy and average value 0.103% for the MAE criteria. Finally, the proposed method was examined along with the existing methods in the field. Considering the fact that the proposed method studied all the regions on the hand, comparing the obtained results, we saw that the best accuracy among all the works belonged to the proposed method of this work. Another feature that distinguished this method was the use of an ensemble of CNN models that used the majority algorithm in an active way. There is no doubt that using an ensemble of CNN models is better than using a single CNN model, as we have seen in similar previous work. Finally, due to the computational complexity of the genetic algorithm, in the future we want to reduce the number of features extracted by using optimization algorithms [46, 47].

References

- [1] S. A. Abdelaziz Ismael, A. Mohammed, and H. Hefny, "An enhanced deep learning approach for brain cancer MRI images classification using residual networks," *Artificial intelligence in medicine*, vol. 102, pp. 101779-101779, 2019. [Online]. Available: https://neuro.unboundmedicine.com/medline/citation/31980109/An_enhanced_deep_learning_approach_for_brain_cancer_MRI_images_classification_using_residual_networks_ [https://linkinghub.elsevier.com/retrieve/pii/S0933-3657\(19\)30617-7](https://linkinghub.elsevier.com/retrieve/pii/S0933-3657(19)30617-7).
- [2] J. H. Lee, Y. J. Kim, and K. G. Kim, "Bone age estimation using deep learning and hand X-ray images," *Biomedical engineering letters*, vol. 10, no. 3, pp. 323-331doi: 10.1007/s13534-020-00151-y.
- [3] R. Lindsey *et al.*, "Deep neural network improves fracture detection by clinicians," *Proceedings of the National Academy of Sciences*, vol. 115, no. 45, pp. 11591-11596, 2018, doi: doi:10.1073/pnas.1806905115.

- [4] B. Guan, J. Yao, G. Zhang, and X. Wang, "Thigh fracture detection using deep learning method based on new dilated convolutional feature pyramid network," *Pattern Recogn. Lett.*, vol. 125, no. C, pp. 521–526, 2019, doi: 10.1016/j.patrec.2019.06.015.
- [5] V. Gilsanz and O. Ratib, *Hand bone age: a digital atlas of skeletal maturity*. Springer, 2005.
- [6] H. Crawley, *When is a child not a child?: Asylum, age disputes and the process of age assessment*. Immigration Law Practitioners' Association (ILPA), 2007.
- [7] A. Mateen, H. K. AFRIDI, and A. R. MALIK, "Age estimation from medial end of clavicle by X-ray examination," *Pak J Med Health Sci*, vol. 7, no. 4, pp. 1106-1108, 2010.
- [8] M. K. Crocker *et al.*, "Sexual dimorphisms in the associations of BMI and body fat with indices of pubertal development in girls and boys," (in eng), *J Clin Endocrinol Metab*, vol. 99, no. 8, pp. E1519-29, Aug 2014, doi: 10.1210/jc.2014-1384.
- [9] J. Tanner, R. Whitehouse, W. Marshall, and B. Carter, "Prediction of adult height from height, bone age, and occurrence of menarche, at ages 4 to 16 with allowance for midparent height," *Archives of disease in childhood*, vol. 50, no. 1, pp. 14-26, 1975.
- [10] J. M. Tanner, "Assessemnt of Skeletal Maturity and Predicting of Adult Height(TW2 Method)," *Prediction of adult height*, pp. 22-37, 1983 1983. [Online]. Available: <https://cir.nii.ac.jp/crid/1573387449513088640>.
- [11] H. M. L. Carty, "Assessment of skeletal maturity and prediction of adult height (TW3 method).: 3rd edition. Edited by J. M. Tanner, M. J. R. Healy, H. Goldstein and N. Cameron. Pp 110. London, etc: W. B. Saunders, 2001. ISBN: 0-7020-2511-9. £69.95," *Journal of Bone and Joint Surgery-british Volume*, pp. 310-311, 2002.
- [12] Z. Zivkovic, "Improved adaptive Gaussian mixture model for background subtraction," in *Proceedings of the 17th International Conference on Pattern Recognition, 2004. ICPR 2004.*, 26-26 Aug. 2004 2004, vol. 2, pp. 28-31 Vol.2, doi: 10.1109/ICPR.2004.1333992.
- [13] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proceedings of ICNN'95 - International Conference on Neural Networks*, 27 Nov.-1 Dec. 1995 1995, vol. 4, pp. 1942-1948 vol.4, doi: 10.1109/ICNN.1995.488968.
- [14] M. Sepahvand and F. Abdali-Mohammadi, "A Meta-heuristic Model for Human Micro Movements Recognition Based on Inertial Sensors," *TABRIZ JOURNAL OF ELECTRICAL ENGINEERING*, vol. 49, no. 1, pp. 221-234, 2019.
- [15] M. K. Mandal, T. Aboulnasr, and S. Panchanathan, "Fast wavelet histogram techniques for image indexing," *Computer Vision and Image Understanding*, vol. 75, no. 1-2, pp. 99-110, 1999.
- [16] J. Seetha and S. S. Raja, "Brain tumor classification using convolutional neural networks," *Biomedical & Pharmacology Journal*, vol. 11, no. 3, p. 1457, 2018.
- [17] M. Sepahvand and F. Abdali-Mohammadi, "Overcoming limitation of dissociation between MD and MI classifications of breast cancer histopathological images through a novel decomposed feature-based knowledge distillation method," *Computers in Biology and Medicine*, vol. 145, p. 105413, 2022/06/01/ 2022, doi: <https://doi.org/10.1016/j.combiomed.2022.105413>.
- [18] S. Hussein, K. Cao, Q. Song, and U. Bagci, "Risk stratification of lung nodules using 3D CNN-based multi-task learning," in *International conference on information processing in medical imaging*, 2017: Springer, pp. 249-260.
- [19] M. Mansourvar, M. A. Ismail, T. Herawan, R. Gopal Raj, S. Abdul Kareem, and F. H. Nasaruddin, "Automated Bone Age Assessment: Motivation, Taxonomies, and Challenges," *Computational and Mathematical Methods in Medicine*, vol. 2013, p. 391626, 2013/12/16 2013, doi: 10.1155/2013/391626.
- [20] G. W. Gross, J. M. Boone, and D. M. Bishop, "Pediatric skeletal age: determination with neural networks," *Radiology*, vol. 195, no. 3, pp. 689-695, 1995.
- [21] E. Pietka, M. F. McNitt-Gray, M. L. Kuo, and H. K. Huang, "Computer-assisted phalangeal analysis in skeletal age assessment," *IEEE Transactions on Medical Imaging*, vol. 10, no. 4, pp. 616-620, 1991, doi: 10.1109/42.108597.
- [22] H. H. Thodberg, S. Kreiborg, A. Juul, and K. D. Pedersen, "The BoneXpert Method for Automated Determination of Skeletal Maturity," *IEEE Transactions on Medical Imaging*, vol. 28, no. 1, pp. 52-66, 2009, doi: 10.1109/TMI.2008.926067.
- [23] A. Tristán-Vega and J. I. Arribas, "A Radius and Ulna TW3 Bone Age Assessment System," *IEEE Transactions on Biomedical Engineering*, vol. 55, no. 5, pp. 1463-1476, 2008, doi: 10.1109/TBME.2008.918554.
- [24] M. Niemeijer, B. van Ginneken, C. Maas, F. Beek, and M. Viergever, *Assessing the skeletal age from a hand radiograph: automating the Tanner-Whitehouse method* (Medical Imaging 2003). SPIE, 2003.
- [25] J. Liu, J. Qi, Z. Liu, Q. Ning, and X. Luo, "Automatic bone age assessment based on intelligent algorithms and comparison with TW3 method," *Computerized Medical Imaging and Graphics*, vol. 32, no. 8, pp. 678-684, 2008/12/01/ 2008, doi: <https://doi.org/10.1016/j.compmedimag.2008.08.005>.
- [26] M. Mansourvar *et al.*, "Automated web based system for bone age assessment using histogram

- technique," *Malaysian Journal of Computer Science*, vol. 25, no. 3, pp. 107-121, 2012.
- [27] H. Lee *et al.*, "Fully Automated Deep Learning System for Bone Age Assessment," *Journal of Digital Imaging*, vol. 30, no. 4, pp. 427-441, 2017/08/01 2017, doi: 10.1007/s10278-017-9955-8.
- [28] C. Spampinato, S. Palazzo, D. Giordano, M. Aldinucci, and R. Leonardi, "Deep learning for automated skeletal bone age assessment in X-ray images," *Medical Image Analysis*, vol. 36, pp. 41-51, 2017/02/01/ 2017, doi: https://doi.org/10.1016/j.media.2016.10.010.
- [29] S. J. Son *et al.*, "TW3-Based Fully Automated Bone Age Assessment System Using Deep Neural Networks," *IEEE Access*, vol. 7, pp. 33346-33358, 2019, doi: 10.1109/ACCESS.2019.2903131.
- [30] D. Giordano, I. Kavasidis, and C. Spampinato, "Modeling skeletal bone development with hidden Markov models," *Computer Methods and Programs in Biomedicine*, vol. 124, pp. 138-147, 2016/02/01/ 2016, doi: https://doi.org/10.1016/j.cmpb.2015.10.012.
- [31] X. Chen, J. Li, Y. Zhang, Y. Lu, and S. Liu, "Automatic feature extraction in X-ray image based on deep learning approach for determination of bone age," *Future Generation Computer Systems*, vol. 110, pp. 795-801, 2020/09/01/ 2020, doi: https://doi.org/10.1016/j.future.2019.10.032.
- [32] F. A. Senel, A. Dursun, K. Ozturk, and V. A. Ayyildiz, "Determination of bone age using deep convolutional neural networks," 2021.
- [33] L. Rabiner and B. Juang, "An introduction to hidden Markov models," *ieee assp magazine*, vol. 3, no. 1, pp. 4-16, 1986.
- [34] M. Sepahvand and F. Abdali-Mohammadi, "A novel method for reducing arrhythmia classification from 12-lead ECG signals to single-lead ECG with minimal loss of accuracy through teacher-student knowledge distillation," *Information Sciences*, vol. 593, pp. 64-77, 2022/05/01/ 2022, doi: https://doi.org/10.1016/j.ins.2022.01.030.
- [35] M. Sepahvand, F. Abdali-Mohammadi, and A. Taherkordi, "Teacher-student knowledge distillation based on decomposed deep feature representation for intelligent mobile applications," *Expert Systems with Applications*, vol. 202, p. 117474, 2022/09/15/ 2022, doi: https://doi.org/10.1016/j.eswa.2022.117474.
- [36] M. Sepahvand and F. Abdali-Mohammadi, "A novel representation in genetic programming for ensemble classification of human motions based on inertial signals," *Expert Systems with Applications*, vol. 185, p. 115624, 2021/12/15/ 2021, doi: https://doi.org/10.1016/j.eswa.2021.115624.
- [37] M. Sepahvand and F. Abdali-Mohammadi, "A novel multi-lead ECG personal recognition based on signals functional and structural dependencies using time-frequency representation and evolutionary morphological CNN," *Biomedical Signal Processing and Control*, vol. 68, p. 102766, 2021/07/01/ 2021, doi: https://doi.org/10.1016/j.bspc.2021.102766.
- [38] F. Cao *et al.*, *An image database for digital hand atlas*. 2003, pp. 461-470.
- [39] M. Sepahvand and F. Abdali-Mohammadi, "A New Learning-based Spatiotemporal Descriptor for Online Symbol Recognition," *Journal of AI and Data Mining*, vol. 10, no. 1, pp. 75-86, 2022.
- [40] M. Sepahvand and F. Abdali-Mohammadi, "A Deep Learning-Based Compression Algorithm for 9-DOF Inertial Measurement Unit Signals Along With an Error Compensating Mechanism," *IEEE Sensors Journal*, vol. 19, no. 2, pp. 632-640, 2019, doi: 10.1109/JSEN.2018.2877360.
- [41] M. Sepahvand, F. Abdali-Mohammadi, and F. Mardukhi, "Evolutionary Metric-Learning-Based Recognition Algorithm for Online Isolated Persian/Arabic Characters, Reconstructed Using Inertial Pen Signals," *IEEE Transactions on Cybernetics*, vol. 47, no. 9, pp. 2872-2884, 2017, doi: 10.1109/TCYB.2016.2633318.
- [42] M. Kashif, T. M. Deserno, D. Haak, and S. Jonas, "Feature description with SIFT, SURF, BRIEF, BRISK, or FREAK? A general question answered for bone age assessment," *Computers in Biology and Medicine*, vol. 68, pp. 67-75, 2016/01/01/ 2016, doi: https://doi.org/10.1016/j.combiomed.2015.11.006.
- [43] A. Gertych, A. Zhang, J. Sayre, S. Pospiech-Kurkowska, and H. K. Huang, "Bone age assessment of children using a digital hand atlas," *Computerized Medical Imaging and Graphics*, vol. 31, no. 4, pp. 322-331, 2007/06/01/ 2007, doi: https://doi.org/10.1016/j.compmedimag.2007.02.012.
- [44] Y. A. Ding, F. Mutz, K. F. Côco, L. A. Pinto, and K. S. Komati, "Bone age estimation from carpal radiography images using deep learning," *Expert Systems*, vol. 37, no. 6, p. e12584, 2020.
- [45] T. D. Bui, J.-J. Lee, and J. Shin, "Incorporated region detection and classification using deep convolutional networks for bone age assessment," *Artificial Intelligence in Medicine*, vol. 97, pp. 1-8, 2019/06/01/ 2019, doi: https://doi.org/10.1016/j.artmed.2019.04.005.
- [46] F. Abdali-Mohammadi, V. Bajalan, and A. Fathi, "Toward a Fault Tolerant Architecture for Vital Medical-Based Wearable Computing," *Journal of Medical Systems*, vol. 39, no. 12, p. 149, 2015/09/12 2015, doi: 10.1007/s10916-015-0347-7.
- [47] H. Lobabi-Mirghavami, F. Abdali-Mohammadi, and A. Fathi, "A novel grammar-based approach to atrial fibrillation arrhythmia detection for pervasive healthcare environments," *Journal of Computing and Security*, vol. 2, no. 2, pp. 155-163, 2015.

ترکیب شبکه های عصبی پیچشی برای تشخیص ناهنجاری های رشد در کودکان با تصاویر اشعه ایکس

حمیرا سرابی سرورانی و فردین ابدالی محمدی*

گروه مهندسی کامپیوتر و فناوری اطلاعات، دانشکده فنی و مهندسی، دانشگاه رازی، کرمانشاه، ایران.

ارسال ۲۰۲۲/۰۳/۱۴؛ بازنگری ۰۵/۰۸/۲۰۲۲؛ پذیرش ۰۱/۰۸/۲۰۲۲

چکیده:

تشخیص سن استخوان روشی است که به طور مکرر برای ارزیابی ناهنجاری رشد، درمان اختلالات غدد درون‌ریز و سندرم‌های کودکان بیمار انجام می‌شود. با پیدایش تصویربرداری دیجیتال، چندین دهه است که تعیین سن استخوانی با ارزیابی بصری از میزان رشد اسکلت دست چپ انجام می‌شود و معمولاً از روش مرجع G&P استفاده می‌شود. باین حال ماهیت ذهنی روش‌های دستی، تعداد زیاد مراکز استخوان در دست و تغییرات گسترده در مراحل استخوان‌سازی سبب پیچیدگی ارزیابی سن استخوانی شده است. بنابراین تلاش‌های زیادی برای توسعه روش‌های پردازش تصویر صورت گرفته است. این روش‌ها به طور خودکار و بیزگی‌های اصلی مراحل تشکیل استخوان را استخراج می‌کند تا به طور موثر و دقیق تر سن استخوان را ارزیابی کنند. در این مقاله یک روش کاملاً اتوماتیک جدید به منظور کاهش خطاهای روش‌های ذهنی و بهبود روش‌های خودکار تخمین سن پیشنهاد شده است. این مدل برای ۱۴۰۰ رادیوگرافی از کودکان سالم صفر تا هجده ساله و جمع آوری شده از چهار قاره استفاده می‌شود. این روش با استخراج تمام نواحی دست، پنج انگشت و مج شروع می‌شود، به طور مستقل سن هر ناحیه را از طریق بررسی مفاصل و نواحی رشد مرتبط با این نواحی توسط CNN محاسبه می‌کند و با ارزیابی سن نهایی از طریق مجموعه ای از CNN ها به پایان می‌رسد. نتایج بدست آمده نشان می‌دهد که روش پیشنهادی دارای دقت ارزیابی متوسط ۸۱ درصد است و در مقایسه با سیستم تجاری که در حال حاضر استفاده می‌شود، عملکرد بهتری دارد..

کلمات کلیدی: تشخیص ناهنجاری های رشد، تصاویر اشعه ایکس، یادگیری مجموعه، شبکه های عصبی کانولوشن.