Computerized Skeletal Bone Age Assessment from Radius and Ulna bones

1P.Thangam, 2Dr.K. Thanushkodi
1Assistant Professor, CSE (PG) Department, 2Director, 
1Sri Ramakrishna Engineering College, Coimbatore -641022, Tamilnadu, India. 
2Akshaya College of Engineering and Technology, Coimbatore -642109, Tamilnadu, India. 
email: saithangam@gmail.com1, thanush12@gmail.com2

Abstract - The work aims at computerizing the skeletal Bone Age Assessment (BAA) process according to the Tanner Whitehouse (TW) method, based on features from Radius and Ulna bones. The system ensures accurate BAA for the age range of 0 – 10 years for males and 0 – 8 years for females. The two most dominant wrist bones, radius and ulna are considered here for estimating the bone age. The input radiograph is initially preprocessed by thresholding and superimposing the thresholded image onto the input image to remove hand borders. Then the radius and ulna Region of Interest (ROI) are cropped and enhanced using anisotropic diffusion filter. 11 morphological features are extracted from the radius ROI (RROI) and ulna ROI (UROI). These features are fed into a decision tree classifier (different for male and female cases), which outputs the age class to which the radiograph is categorized (Class A – Class J). This class is finally mapped on to the final bone age. The system is very robust and does not reject any radiographs. The system utilized a data set of 220 radiographs (110 males & 110 females) which was organized in three types of partitions and the performance of the system for each partition was analyzed. The results were evaluated with the help of diagnosis results obtained from two skilled radiologists.

Keywords: Bone Age Assessment (BAA), Tanner Whitehouse (TW), radiograph, radius, ulna, Decision tree, Classification.

I. INTRODUCTION

Bone age assessment using a hand radiograph is an important clinical tool in the area of pediatrics, especially in relation to endocrinological problems and growth disorders [1]. Based on a radiological examination of skeletal development of the left-hand wrist, bone age is assessed and compared with the chronological age. A discrepancy between these two values indicates abnormalities in skeletal development. The procedure is often used in the management and diagnosis of endocrine disorders and also serves as an indication of the therapeutic effect of treatment. Bone age indicates whether the growth of a patient is accelerating or decreasing, based on which the patient can be treated with growth hormones. BAA is universally used due to its simplicity, minimal radiation exposure, and the availability of multiple ossification centers for evaluation of maturity. The main clinical methods for skeletal bone age estimation are the Greulich & Pyle (GP) method and the Tanner & Whitehouse (TW) method. GP is an atlas matching method while TW is a score assigning method [2]. GP method is faster and easier to use than the TW method. Bull et. al. compared the GP and TW method and concluded TW method to be more accurate [3]. TW method uses a detailed analysis of each individual bone, assigning it to one of eight classes reflecting its developmental stage. Each bone is illustrated in terms of scores. The sum of all scores assesses the bone age. The development of each ROI is divided into discrete stages, as shown in Figure 1, and each stage is given a letter (A,B,C,D,…,I), reflecting the development stage as:

- Stage A – absent
- Stage B – single deposit of calcium
- Stage C – center is distinct in appearance
- Stage D – maximum diameter is half or more the width of metaphysis
- Stage E – border of the epiphysis is concave
- Stage F – epiphysis is as wide as metaphysis
- Stage G – epiphysis caps the metaphysis
- Stage H – fusion of epiphysis and metaphysis has begun
- Stage I – epiphyseal fusion completed.

By adding the scores of all ROIs, an overall maturity score is obtained. This score is correlated with the bone age differently for males and females [4].

II. SURVEY OF LITERATURE

In 1980s, Pal and King proposed the theory of fuzzy sets for edge detection in X-ray images [5]. Kwabwe et. al. in 1986, proposed certain algorithms to recognize the bones in an X-ray image of the hand and wrist [6]. Pathak and Pal [7] developed a fuzzy classifier for syntactic recognition of different stages of maturity of bones from X-rays of hand and wrist. Michael and Nelson [8] developed a model-based system for automatic segmentation of bones from digital hand radiographs named as HANDX, in 1989. This computer vision system, offered a solution to automatically find, isolate and measure bones from digital X-rays. In 1991, Pietka et. al. described a method [9] based on independent analysis of the phalangeal regions. Phalangeal analysis was performed in several stages by measuring the lengths of the distal, middle and proximal phalanx which were converted into skeletal age by using the standard phalangeal length table proposed by Garn et.al [10]. Tanner and Gibbons introduced the Computer- Assisted Skeletal Age Scores (CASAS) system in 1992 [11]. In 1993, Pietka et. al. performed phalangeal and
carpal bone analysis using standard and dynamic thresholding methods to assess skeletal age [12]. Cheng et. al. [13] proposed the methods to extract a region of interest (ROI) for texture analysis in 1994, with particular attention to patients with hyperparathyroidism. The techniques included multiresolution sensing, automatic adaptive thresholding, detection of orientation angle, and projection taken perpendicular to the line of least second moment. In the same year, Drayer and Cox [14] designed a computer aided system to estimate bone age based on Fourier analysis on radiographs to produce TW2 standards for radius, ulna and short finger bones. In 1996, Al-Taani et. al. classified the bones of the hand-wrist images into pediatric stages of maturity using Point Distribution Models (PDM) [15]. Wastl and Dickhaus proposed a pattern recognition based BAA approach, in the same year [16]. The approach consisted of four major steps: digitization of the hand radiograph, segmentation of ROI, prototype matching and BAA. Mahmoodi et. al. (1997) used Knowledge-based Active Shape Models (ASM) in an automated vision system to assess the bone age [17]. Pietka et. al. conducted a computer assisted BAA procedure [18] by extracting and using the epiphyseal/metaphyseal ROI (EMROI), in 2001. From each phalanx 3 EMROIs were extracted which include: metaphysis, epiphysis and diaphysis of the distal and middle phalanges and for the proximal phalanges it includes metaphysis, epiphysis and upper part of metacarpals of proximal phalanges. The diameters of metaphysis, epiphysis and diaphysis of each EMROI were measured. The extracted features described the stage of skeletal development more objectively than visual comparison. Niemeijer et. al. [19] automated the TW method by constructing a mean image and using a query image to assess age. M.Fernandez et. al. [20] described a method for registering human hand radiographs for automatic BAA using the GP method. A.Fernandez et. al. proposed a fuzzy logic based neural architecture for BAA [21]. The system employed a computing with words paradigm, wherein the TW3 statements were directly used to build the computational classifier. Luis Garcia et. al. presented a fully automatic algorithm [22] to detect bone contours from hand radiographs using GVF model, to extract a variety of carpal bone features [23]. In 2005, Tristan and Arribas [24] designed an end-to-end system to partially automate the TW3 bone age assessment procedure, using a modified K-means adaptive clustering algorithm for segmentation, extracting up to 89 features and employing LDA for feature selection and finally estimating bone age using a Generalized Softmax Perceptron (GSP) NN, whose optimal complexity was estimated via the Posterior Probability Model Selection (PPMS) algorithm. Zhang et. al. developed a knowledge based carpal ROI analysis method [25] for fully automatic carpal bone segmentation and feature analysis for bone age assessment by fuzzy classification. Thodberg et. al. proposed a 100% automated approach called the Bone Xpert method [26]. The architecture of Bone Xpert divided the processing into three layers: Layer A to reconstruct the bone borders, Layer B to compute an intrinsic bone age value for each bone and Layer C to transform the intrinsic bone age value using a relatively simple post-processing. Giordano et. al. [27] designed an automated system for skeletal bone age evaluation using DoG filtering and a novel adaptive thresholding. Hsieh et. al. [28] proposed an automatic bone age estimation system based on the phalanx geometric characteristics and carpal fuzzy information. Zhao Liu and Jian Liu proposed an automatic BAA method with image template matching based on PSO [29]. Giordano et. al. [30] presented an automatic system for BAA using TW2 method by combining Gradient Vector Flow (GVF) Snakes and derivative difference of Gaussian filter. We have presented a thorough survey of literature on BAA methods in our previous work [31], explaining in detail the various work done.
in the area and providing directions for future research. We have also presented an analysis of features extracted from the carpal and phalangeal bones, and their contribution in BAA [32]. We have developed a PSO based segmentation algorithm to segment wrist bones from radiographs [33]. In addition, we have also done a comparative study on the performance of four different classifiers, namely Artificial Neural Networks (ANN), Support Vector Machines (SVM), Naïves Bayesian (NB) and Decision Tree (DT) in estimating the bone age [34]. Our previous work [35] describes a computerized BAA method for carpal bones, by extracting features from the convex hull of each carpal bone. We have also developed a BAA system [36] to estimate bone age from carpal and radius wrist bones.

III. MATERIALS AND METHODS

a) Data Set
The database consists of totally 220 images, of which 110 are male images and 110 female. The system was initially trained with 6 male and 6 female radiographs for each age group, thus with a total of 120 radiographs, 60 male and 60 female cases and was tested with 50 male images and 50 female images.

b) Pre-Processing
The input radiograph is initially thresholded to detect the foreground and background pixels. The thresholded image is superimposed on the input image to remove the hand borders in the image. Then the epiphyseal ROI of radius and ulna bones are cropped and isolated. The cropped ROI are pre-processed for noise reduction using anisotropic diffusion filter. Figure 2 shows an input, thresholded and superimposed image during preprocessing. Figure 3 depicts samples for cropped and filtered radius ROI (RROI) and ulna ROI (UROI) for feature extraction.

c) Feature set
For both radius and ulna bones, the degree of maturity is based on the extent of ossification of their epiphysis. So, for both the categories, we consider the same set of epiphyseal features to be extracted, namely:
1. Presence – whether the epiphysis of the bone is present or absent
2. Circularity – whether the epiphysis is circular in shape or not
3. Roughness – whether the epiphysis surface is smooth or irregular
4. Capping – whether the epiphysis capping has begun or not
5. Fusion – whether the epiphysis fusion has begun or not.

1. Presence
The presence of the radius or ulna epiphysis is marked by a TRUE value for Presence and returning a FALSE value if absent.

2. Circularity
Circularity is defined as a simple shape factor based on the projected area of the sample and the overall perimeter of the projection given by,

\[
\text{Circularity} = 4\pi A/P^2
\]

where \(A\) is the area and \(P\) is the perimeter of the epiphysis sample. Values for Circularity range from 1 for a perfect circle to 0 for a line. We check for the Circularity value of the bone. If it is greater than or equal to 0.5, it returns a TRUE

Figure 4 Decision Tree for (a) Males (b) Females.
### Table 1: Criteria Selection for Male

<table>
<thead>
<tr>
<th>S.No.</th>
<th>Age Class</th>
<th>Age Value v (Years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>A</td>
<td>1 month &lt; v &lt; 1 year</td>
</tr>
<tr>
<td>2.</td>
<td>B</td>
<td>1 year &lt; v &lt; 2 year</td>
</tr>
<tr>
<td>3.</td>
<td>C</td>
<td>2 year &lt; v &lt; 3 year</td>
</tr>
<tr>
<td>4.</td>
<td>D</td>
<td>3 year &lt; v &lt; 4 year</td>
</tr>
<tr>
<td>5.</td>
<td>E</td>
<td>4 year &lt; v &lt; 5 year</td>
</tr>
<tr>
<td>6.</td>
<td>F</td>
<td>5 year &lt; v &lt; 6 year</td>
</tr>
<tr>
<td>7.</td>
<td>G</td>
<td>6 year &lt; v &lt; 7 year</td>
</tr>
<tr>
<td>8.</td>
<td>H</td>
<td>7 year &lt; v &lt; 8 year</td>
</tr>
<tr>
<td>9.</td>
<td>I</td>
<td>8 year &lt; v &lt; 9 year</td>
</tr>
<tr>
<td>10.</td>
<td>J</td>
<td>9 year &lt; v &lt; 10 year</td>
</tr>
</tbody>
</table>

### Table 2: Criteria Selection for Female

<table>
<thead>
<tr>
<th>S.No.</th>
<th>Age Class</th>
<th>Age Value v (Years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>A</td>
<td>1 month &lt; v &lt; 3 months</td>
</tr>
<tr>
<td>2.</td>
<td>B</td>
<td>3 months &lt; v &lt; 6 months</td>
</tr>
<tr>
<td>3.</td>
<td>C</td>
<td>7 months &lt; v &lt; 1 year</td>
</tr>
<tr>
<td>4.</td>
<td>D</td>
<td>1 year &lt; v &lt; 2 year</td>
</tr>
<tr>
<td>5.</td>
<td>E</td>
<td>2 year &lt; v &lt; 4 year</td>
</tr>
<tr>
<td>6.</td>
<td>F</td>
<td>4 year &lt; v &lt; 5 year</td>
</tr>
<tr>
<td>7.</td>
<td>G</td>
<td>5 year &lt; v &lt; 5 years 5 months</td>
</tr>
<tr>
<td>8.</td>
<td>H</td>
<td>5 year 5 months &lt; v &lt; 6 year</td>
</tr>
<tr>
<td>9.</td>
<td>I</td>
<td>6 year &lt; v &lt; 7 year</td>
</tr>
<tr>
<td>10.</td>
<td>J</td>
<td>7 year &lt; v &lt; 8 year</td>
</tr>
</tbody>
</table>

### Table 3: Confusion Matrix

<table>
<thead>
<tr>
<th>Actual Class</th>
<th>Predicted Class</th>
<th>Positive</th>
<th>Negative</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Positive</td>
<td>True Positive (TP)</td>
<td>False Negative (FN)</td>
</tr>
<tr>
<td></td>
<td>Negative</td>
<td>False Positive (FP)</td>
<td>True Negative (TN)</td>
</tr>
</tbody>
</table>

### Table 4: Classification Metrics

<table>
<thead>
<tr>
<th>Class</th>
<th>Confusion Matrix</th>
<th>precision%</th>
<th>recall%</th>
<th>specificity%</th>
<th>accuracy%</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>10 0 0 80</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>B</td>
<td>10 0 0 80</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>C</td>
<td>9 1 80 0</td>
<td>90</td>
<td>100</td>
<td>99</td>
<td>99</td>
</tr>
<tr>
<td>D</td>
<td>1 82 1 9</td>
<td>89</td>
<td>89</td>
<td>99</td>
<td>98</td>
</tr>
<tr>
<td>E</td>
<td>0 81 1 8</td>
<td>100</td>
<td>90</td>
<td>100</td>
<td>99</td>
</tr>
<tr>
<td>F</td>
<td>8 2 82 2</td>
<td>80</td>
<td>80</td>
<td>98</td>
<td>96</td>
</tr>
<tr>
<td>G</td>
<td>8 2 82 2</td>
<td>80</td>
<td>80</td>
<td>98</td>
<td>96</td>
</tr>
<tr>
<td>H</td>
<td>7 83 1 2</td>
<td>78</td>
<td>88</td>
<td>98</td>
<td>97</td>
</tr>
<tr>
<td>I</td>
<td>8 1 82 1</td>
<td>89</td>
<td>89</td>
<td>99</td>
<td>98</td>
</tr>
<tr>
<td>J</td>
<td>9 1 81 1</td>
<td>90</td>
<td>90</td>
<td>99</td>
<td>98</td>
</tr>
</tbody>
</table>
value. In the FALSE case, (i.e.) if it is non-circular, we calculate the diameter of the epiphysis, R_diameter and check if it is above or below the threshold (which is calculated as the mean of the R_diameter values of both the extreme age classes) and assign a class to it accordingly.

3. Roughness

Roughness of the bone is measured by employing Fractal dimension. Fractal dimension is found to be a measure of roughness and is given by,

\[ \text{Roughness} = \frac{\log N}{\log 1/r} \]  

Where, \( N \) is the number of copies of a self similar set, which has been scaled down and \( r \) is the scaled ratio of the self similar set. If Roughness is above the threshold value (which is calculated as the mean of the Roughness values of both the extremes), it returns TRUE.

4. Capping

This is calculated as the difference between the horizontal diameter of the sample bone and the horizontal diameter of the corresponding epiphysis, as given in equation (3). If Capping value is negative, then it returns TRUE.

\[ \text{Capping} = \text{Diameter}_{\text{Bone}} - \text{Diameter}_{\text{epiphysis}} \]  

5. Fusion

Fusion is computed by simply measuring the Euclidean distance between the bone and the epiphysis. If it is negative, then it returns TRUE.

The above features are extracted from the radius and ulna bones, leading to the following features: \( R_{\text{Presence}}, U_{\text{Presence}}, R_{\text{Diameter}}, U_{\text{Diameter}}, R_{\text{Roughness}}, U_{\text{Roughness}}, R_{\text{Capping}}, U_{\text{Capping}}, R_{\text{Fusion}}, \) and \( U_{\text{Fusion}} \). Thus at the end of the feature extraction phase, we end up with 11 morphological features for further processing.

IV. BONE AGE ESTIMATION

Feature extraction is followed by either the training phase or testing phase. In case of training, the classifier is trained with the range of values taken by each feature with respect to each age class. During testing, the extracted features are translated into corresponding skeletal bone age. The selected features occupy positions in the decision tree based on the order of ossification of the bones, which influence their contribution towards the bone age estimation procedure.

a) Decision Tree classifier:

The morphological features extracted from the wrist bones have a certain degree of overlapping between them, since the growth pattern is gradual. But the features can fall into only any one of the categories (i.e.) mutually exclusive. So the decision tree classifier [37] is more suitable for our system. The inputs to the decision tree (shown in Figure 4. (a) for males and Figure 4. (b) for females) are the 11 features modeled as Boolean values, and the output will be the skeletal class or category. The selected class is then mapped onto the skeletal bone age, based on the criteria shown in Table 1 for males and Table 2 for females.

V. RESULTS AND DISCUSSION

The performance of the proposed system in estimating the bone age was evaluated using a dataset consisting of 100 radiographs (50 for boys and 50 for girls). Much attention was dedicated to deploy robust techniques for preprocessing and reliable feature extraction. The ease and accuracy of feature extraction for the UROI were slightly imprecise when compared to the RROI, the reason being that the epiphysis of the ulna bone was smaller in size and less contributing. Also the accuracy of feature extraction tends to decline slightly as the age increases since at later stages the system relies much on the ulna ROI. On the other hand, the accuracy of feature extraction from RROI was outstanding, thus improving age estimation during early stages. The accuracy of the classification was measured in terms of four metrics, precision, recall, specificity and accuracy. The above four metrics are measured by using a confusion matrix representation for each class A – J, consisting of True Positive(TP), False Negative(FN), False Positive(FP), and True Negative(TN) cases, as shown in Table 3.

From the confusion matrix, the performance metrics are calculated using the following formulae: The performance of the system was validated by using the diagnoses results obtained for the data set from two skilled radiologists. The diagnoses results were obtained from two radiologists in order to overcome discrepancies. From the values of the metrics tabulated in Table 4, it is found that the minimum value of precision is 78% and obtained for the class H and that of recall is 80% obtained for classes F and G. The minimal value obtained for specificity is 98% obtained for classes F, G and H and that of accuracy is 96% obtained for the classes F and G. Classes F and G found to downgrade the overall performance of the classifier, the reason being the less contribution of UROI or poor ossification of the epiphysis of the ulna bone for those two age classes. Best performance was found in the earlier classes A and B, since the high contributing RROI are the only bones considered and good inter-class difference. The performance of the system is depicted in the graph shown in
Figure 5.

In order to further improve the performance of the system, the data set partition was changed from (Partition I) 120 train and 100 test images into 160 train and 60 test images (Partition II). Better results were obtained for this Partition II yielding 99% accuracy, 100% specificity, 92% precision and 93% recall. The partition set was further altered into Partition III with 180 train images and 40 test images. The system produced best results for this partition, achieving 100% in all the four performance metrics. Figure 6 depicts the comparison of the performance metrics of the system in all the three partitions.

![Figure 6. Comparison of Performance Metrics for three partitions.](image)

VI. CONCLUSIONS AND FUTURE WORK

An efficient computerized approach to estimate the skeletal bone age was proposed. The method is beneficial in that it requires only small amount of data to train the classifier and from the results it is evident that the performance of the system is very robust and appropriate for the age group of 0-10 years for males and 0-8 years for females. The system was validated with the results obtained from two radiologists. Future work will be focused on extending the system to work on the age group above 10 years, and broadening the system to include the further TW2 bones such as carpals, metacarpals, phalanges, etc. and also merging the system with Picture Archiving and Communication System.

VII. ACKNOWLEDGMENT

The authors would like to thank Dr. R. Saravanan, MD (Pediatrics), KKCT Hospital, Chennai for providing the database of hand radiographs, Dr. R. Shankar Anandh, MD (Radio Diagnosis) and Dr. S. Sanjitha, DMRD (Radio Diagnosis) of Barnard Institute of Radiology, Chennai for their valuable help in diagnosing the hand radiographs and validating the system.

VIII. REFERENCES

[25] A. Zhang, A. Gertych, B. Liu, “Automatic bone age assessment for young children from newborn to 7-year-old using car-


