Bone Age Assessment of Iranian Children in an Automatic Manner

Abstract

Background: Bone age assessment (BAA) is a radiological process with the aim of identifying growth disorders in children. The objective of this study is to assess the bone age of Iranian children in an automatic manner. Methods: In this context, three computer vision techniques including histogram of oriented gradients (HOG), local binary pattern (LBP), and scale-invariant feature transform (SIFT) are applied to extract appropriate features from the carpal and epiphyseal regions of interest. Two different datasets are applied here: the University of Southern California hand atlas for training this computer-aided diagnosis (CAD) system and Iranian radiographs for evaluating the performance of this system for BAA of Iranian children. In this study, the concatenation of HOG, LBP, and dense SIFT feature vectors and background subtraction are applied to improve the performance of this approach. Support vector machine (SVM) and K-nearest neighbor are used here for classification and the better results yielded by SVM. Results: The accuracy of female radiographs is 90% and of male is 71.42%. The mean absolute error is 0.16 and 0.42 years for female and male test radiographs, respectively. Cohen’s kappa coefficients are 0.86 and 0.6, \( P < 0.05 \), for female and male radiographs, respectively. The results indicate that this proposed approach is in substantial agreement with the bone age reported by the experienced radiologist. Conclusion: This approach is easy to implement and reliable, thus qualified for CAD and automatic BAA of Iranian children.

Keywords: Bone age assessment, computer vision operators, Iranian race, K-nearest neighbors, left-hand radiographic images, support vector machine

Introduction

Bone age assessment (BAA) is defined as a clinical procedure for indicating the stage of skeletal maturation in children. Therefore, BAA can be beneficial in diagnosing growth disorders as well as hormonal and genetic problems in children.[1] In general, bone age is assessed by measuring the maturity of bones through left-hand radiographs. Pediatricians typically assess the bone age of children through two basic methods, namely Greulich and Pyle (GP) method[2] and Tanner and Whitehouse (TW) method.[3] In the GP method, the pediatrician compares a patient’s radiographs with radiographs in the GP atlas. Then, the most corresponding radiograph to the patient’s radiograph is identified, and an estimated bone age is reported. In the TW method, which is a scoring method, a score is given to the maturation of a patient’s bones. Then, an estimated bone age is reported by translating the calculated score to the bone age by applying the standard table. The TW method is time-consuming, unbearable, and complex; therefore, 76% of pediatricians prefer to assess the bone age of children by GP atlas.[4] The average time needed to assess bone age by GP and TW methods is 1.4 and 7.9 min, respectively; hence, the manual BAA is considered as a time-consuming method.[5] Moreover, it depends on the pediatrician’s skill and is not accurate for neither inter-observation nor intra-observation.

BAA greatly depends on the race and sex of children; therefore, in recent years, researchers focusing on BAA have revealed an increasing interest in applying computer-aided diagnosis (CAD) systems for assessing the bone of children in different races. In 2014, El-Bakary et al. run a study on BAA of Egyptian children within 4–18 years’ range. In this study, the ratio between the total area of carpal bones and epiphyses of the ulna and radius (Bo) and
carpals (Ca) were used as age indicators. They reported a formula for assessing the bone age of Egyptian children as the following equation: \( \text{Age} = -0.998 + 18.708 \frac{\text{Bo}}{\text{Ca}} + 1.724 \ g(\text{Bo/Ca}) \), where \( g \) is a dummy variable equal to 1 for boys and 0 for girls. The median of the absolute values of residuals and standard error were -1.67 years and 1.85 years, respectively.

In 2016, De Luca et al. proposed a formula for assessing the bone age of Italian children. In their study, carpals and epiphyses of radius and ulna were used as age indicators. They reported the following equation for BAA of Italian children: \( \text{Age} = -1.7702 + 1.00889 \ g + 14.8166 \ (\text{Bo/Ca}) \). The median of the absolute values of residuals (observed age minus predicted age) was -0.38, and the standard error was 1.54 years.

In our previous work, the histogram of oriented gradients (HOG), local binary pattern (LBP), and scale-invariant feature transform (SIFT) are applied for feature extraction from carpal region of interest (ROI) and eROI (which belongs to the epiphyseal center of proximal phalanx) [Figure 1]. As a dataset, 442 left-hand radiographs were applied from the University of Southern California (USC) hand atlas. The HOG, LBP, and dense SIFT were concatenated to improve BAA. Support vector machine (SVM) and 5-fold cross-validation were applied for classification. The accuracy of female radiographs was 73.88% and of male was 68.63%. The mean absolute error (MAE) was 0.5 years for both genders’ radiographs. The accuracy within 1-year range was 95.32% for female and 96.51% for male radiographs. The accuracy within 2-year range was 100% and 99.41% for female and male radiographs, respectively. According to the results, in spite of using two ROIs, the higher accuracy and lower error were obtained in comparison with the other published methods in this field. Therefore, we decided to run a new study to assess the bone age of Iranian race in an automatic manner using the aforementioned methods.

In this study, radiographs from the USC hand atlas are used for training the CAD system. The HOG, LBP, and dense SIFT are applied in carpal ROI and eROI for feature extraction. Iranian left-hand radiographs are entered to the CAD system, and an estimated bone age is reported as a result. SVM and K-nearest neighbor (KNN) are used for classification. According to the GP atlas, bone evolution rate is significantly higher in females than in males; consequently, this study is run on both genders’ radiographs in a separate manner.

**Methods**

The CAD bone age chain here consists of preprocessing, ROI extraction, background subtraction, feature extraction, and classification [Figure 2]. To implement the above stages, carpal ROIs and eROIs are extracted from Iranian radiographs, followed by applying background subtraction on them. The HOG, LBP, and dense SIFT features are extracted from ROIs and then concatenated to yield better features as to BAA. Eventually, SVM and KNN are used for classification. In this study, the USC radiographs are used for training the CAD system where the Iranian radiographs are used as a test set for BAA.

All images of the USC dataset consist of two reports from two radiologists. In this study, 442 radiographs (220 radiographs for females and 222 radiographs for males) within 0–18 years’ range are applied as a training set for CAD system. For assessing the bone age of Iranian children, 17 radiographs (10 radiographs for females and 7 radiographs for males) are applied as a test dataset. All images consist of a report from an experienced radiologist in Isfahan University of Medical Sciences.

The implemented stages are as follows:

**Preprocessing**

In this stage, anisotropic diffusion filtering is applied for the purpose of noise reduction in radiographic images. Moreover, homomorphic filtering is applied in radiographs to normalize the brightness of them and increase their contrast as well.

**Region of interest extraction**

The ROI extraction method is not in an automatic manner. In this study, two ROIs including carpal ROI and eROI are extracted from left-hand radiographs. The carpal area which consists of carpal bones, distal radius, and ulna contains discriminative features for BAA of young children. However, at older ages, this area does not have a desirable performance for BAA. Therefore, an epiphyseal center which belongs to the proximal phalanx is extracted to improve the performance of the proposed approach.
Background subtraction

In this stage, top-hat transform operation is applied on ROIs to remove objects from the image, correct the nonuniform brightness of the images, and increase the efficiency of this approach.\cite{6} The output of this stage is applied for the purpose of feature extraction.

Feature extraction

From prior work, three computer vision methods including HOG, LBP, and dense SIFT are used for feature extraction. HOG and LBP are known as object detection methods which deal with detecting the instances of semantic objects of a certain class in digital images. In general, there exist no bones in the carpal area at birth and the number of which is completed upon growth. Hence, HOG-LBP feature extraction method is applied as an object detection method in young children. Dense SIFT is known as an object recognition which deals with identifying objects in an image through their identities such as size or scale. Since the number of carpal bones will be completed in children older than 5–7 years, the bone maturation index in this age range is the size of the bones in the carpal area.\cite{6} Therefore, SIFT feature description method is added to extract more accurate features for BAA of Iranian children.

Scale-invariant feature transform

The procedure of feature extraction in SIFT contains two major steps which are feature detection and feature description. In feature detection, keypoints, which represent the most informative parts of the image, are determined through an algorithm. In feature description, a local descriptor is computed for each keypoint which can be selected in sparse or dense manner.\cite{11}

Histogram of oriented gradients

HOG is a feature description technique and is reliable for the purpose of object detection. This method applied the distribution of local gradients or edges to describe an image, even if there is no accurate information about the position of gradients or the direction of edges.\cite{12}

Local binary pattern

LBP is a feature description technique which determines microstructures such as edges, spots, flat areas, and lines.\cite{13,14} This method is one of the best descriptors for texture description.\cite{15}

Classification

In this study, SVM executed with linear kernel function and KNN is applied with City block and Euclidean distances for female and male radiographs, respectively.\cite{16,17}

Validation experiments

The size of feature vectors of three computer vision techniques is related to the image size. Therefore, all the ROIs extracted from radiographs are rescaled to 48 × 48 pixels. This size is recognized as the best as to feature extraction time and BAA accuracy.\cite{6}

In this study, radiographs from the USC dataset are used for training the CAD system, and Iranian radiographs are used as a test dataset. Bone growth is of different rates in females and males; therefore, this approach is run on both their radiographs, separately.

In this study, the accuracy of BAA in Iranian race and MAE is calculated to evaluate the performance of the proposed approach for both the genders’ radiographs. The accuracy is calculated through Eq. 1:\cite{18}

\[
\text{Accuracy} = \frac{tp + tn}{n}
\]  

(1)

Where \( tp \) is the true positive, \( tn \) is the true negative, and \( n \) is the number of data.

The MAE is calculated through Eq. 2:

\[
\mu_{\text{MAE}} = \left( \frac{1}{H} \sum_{h=1}^{H} |a^{\text{read}}[h] - a^{\text{predict}}[h]| \right)
\]  

(2)

Where \( \mu \) is the MAE. Moreover, \( a^{\text{read}} \) is the predicted age reported by the experienced radiologist in Isfahan University of Medical Sciences for each hand \( h \). Definition \( a^{\text{predict}} \) is the predicted age of this proposed method and \( H \) is the total number of hand radiographs. The predicted age is calculated according to the equation presented by Kashif et al. as follows:

\[
\text{Age} = \frac{1}{2} (UB[c] + LB[c])
\]  

(3)
Where \(c\) is the predicted age class and UB (\(c\)) and LB (\(c\)) are the upper and lower bands of this class.\(^{[19]}\)

### Implementation

The whole procedures in this study are run completely in matrix laboratory (MATLAB). The computer vision toolbox of MATLAB is used here which provides built-in support for HOG and LBP. The procedures of keypoint detection and description of SIFT are applied through the VLFEAT library version 0.9.21 (Andrea Vedaldi and Brian Fulkerson in 2007, Oxford University of England, United Kingdom).

### Results

The feature vectors’ size of the HOG, LBP, and dense SIFT are 900, 360, and 1024, respectively. The HOG, LBP, and dense SIFT features are concatenated to improve the performance of this proposed method for BAA of Iranian children. Moreover, the performance of this proposed approach is considered when combining Iranian radiographs and the USC dataset.

The bone age reported by the experienced radiologist and the predicted age by this proposed method for female radiographs using SVM are tabulated in Table 1.

The bone age reported by the experienced radiologist and the predicted age by this proposed method for male radiographs using SVM are tabulated in Table 2.

The accuracy and MAE of this proposed method for female radiographs for SVM and KNN classifiers are tabulated in Table 3.

The bone age reported by the experienced radiologist and the predicted age by this proposed method for male radiographs using SVM are tabulated in Table 4.

The bone age reported by the experienced radiologist and the predicted age by this proposed method for male radiographs using KNN are tabulated in Table 5.

The accuracy and MAE of this proposed method for male radiographs for SVM and KNN are tabulated in Table 6.

According to the obtained results, the SVM classifier outperforms the KNN; therefore, in this study, Cohen’s kappa statistical test is calculated for the bone age assessed using SVM to provide a more accurate BAA for Iranian children. The results yielded from Cohen’s kappa for female radiographs is 0.86 (confidence interval [CI] 95%, 0.60–1.11), \(P\)-value < 0.05, which indicate the perfect agreement between the results here and the gold standard (the BAA reported by an experienced radiologist). The results obtained from Cohen’s kappa for male radiographs is 0.6 (CI 95%, 0.14–1.05), \(P\)-value < 0.05, which indicate the moderate agreement between the results of this study and the gold standard.

To consider the results of cross-validation for BAA when combining both USC and Iranian datasets, we used SVM and 5-fold cross-validation. For this classification, SVM is executed with a radial basis function (RBF) kernel and one-against-all approach. Twenty percent of the total data
of Iranian radiographs is much lower than that of the USC dataset, so it does not have a significant effect on the results. However, because of combining two different datasets, the accuracy decreased in this experiment.

**Discussion**

According to the previous research, BAA depends greatly on the race and sex of children. Since there exists no study on automatic BAA of Iranian children, we decided to find a manner in assessing the bone age of Iranian race automatically. In our previous work, the obtained results indicated that HOGLBPdense SIFT feature extraction techniques and background subtraction were promising measures approach in BAA. Hence, in this study, attempt is made to assess the bone age of Iranian children using our previous successful approach.

The results of accuracy and MAE for the SVM classifier reveal that this proposed approach is significantly reliable for BAA of Iranian children rather than the KNN. According to the obtained results, carpal area is more reliable than epiphyseal phalanx for BAA. The performance of the CAD system is considerably improved after applying two ROIs simultaneously. The results of the estimated bone age reveal that better results are obtained in case of female radiographs which assuring the fact that female bones develop faster than that of males. Although the proposed approach yielded good results especially for female radiographs, more data is needed to affirm the efficiency of this approach on BAA of Iranian children.

**Conclusion**

In this article, for the first time, attempt is made to assess the bone age of Iranian children in an automatic manner. Here, two different datasets are used: the USC dataset for training and Iranian dataset for evaluating the performance of this proposed approach for BAA of Iranian race. SVM and KNN are applied here for classification. According to the obtained results, the SVM classifier is more reliable than the KNN for assessing the bone age of Iranian children. The accuracy of female and male radiographs using this classifier is 90% and 71.42%, respectively, when using carpal ROI and eROI simultaneously. To evaluate the performance of this CAD system more accurately, MAE is calculated. The MAE is 0.16 years for female test radiographs and 0.42 years for male test radiographs.

The Cohen’s kappa statistical test is 0.86 for female and 0.6 for male radiographs, \( P < 0.05 \). The results here indicate that there is a significant agreement between the bone age assessed by this newly proposed approach and the bone age assessed by an experienced radiologist. The results of accuracy and MAE for the SVM classifier reveal that this proposed approach is significantly reliable for BAA of Iranian children rather than the KNN. According to the obtained results, the carpal area is more reliable than epiphyseal phalanx for BAA. The performance of

<table>
<thead>
<tr>
<th>Table 4: The bone age reported by the experienced radiologist and the result of the proposed computer-aided diagnosis system for male radiographs using support vector machine</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>The report of bone age by radiologist (year)</strong></td>
</tr>
<tr>
<td>Carpal ROI</td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>3.5</td>
</tr>
<tr>
<td>14.5</td>
</tr>
<tr>
<td>10.5</td>
</tr>
<tr>
<td>3.5</td>
</tr>
<tr>
<td>1.5</td>
</tr>
<tr>
<td>11.5</td>
</tr>
<tr>
<td>12.5</td>
</tr>
</tbody>
</table>

CAD – Computer-aided diagnosis; ROI – Region of interest; eROI – Epiphyseal region of interest

<table>
<thead>
<tr>
<th>Table 5: The bone age reported by the experienced radiologist and the result of the proposed computer-aided diagnosis system for male radiographs using K-nearest neighbor</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>The report of bone age by radiologist (year)</strong></td>
</tr>
<tr>
<td>Carpal ROI</td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>3.5</td>
</tr>
<tr>
<td>14.5</td>
</tr>
<tr>
<td>10.5</td>
</tr>
<tr>
<td>3.5</td>
</tr>
<tr>
<td>1.5</td>
</tr>
<tr>
<td>11.5</td>
</tr>
<tr>
<td>12.5</td>
</tr>
</tbody>
</table>

CAD – Computer-aided diagnosis; ROI – Region of interest; eROI – Epiphyseal region of interest

<table>
<thead>
<tr>
<th>Table 6: Validation results for male radiographs</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Outcome</strong></td>
</tr>
<tr>
<td>SVM Accuracy</td>
</tr>
<tr>
<td>MAE</td>
</tr>
<tr>
<td>KNN Accuracy</td>
</tr>
<tr>
<td>MAE</td>
</tr>
</tbody>
</table>

MAE – Mean absolute error; ROI – Region of interest; eROI – Epiphyseal region of interest; SVM – Support vector machine

is selected on a random basis as a validation set to tune the hyperparameters of RBF kernel and select appropriate features as well. The results obtained in this step are applied on the remaining 80% of the total data for classification. The accuracy of female radiographs was 69.36% and of male was 59.22%. The accuracy within 1-year range was 85.55% for female and 92.89% for male radiographs. The accuracy within 2-year range was 90.75% and 98.82% for female and male radiographs, respectively.

is selected on a random basis as a validation set to tune the hyperparameters of RBF kernel and select appropriate features as well. The results obtained in this step are applied on the remaining 80% of the total data for classification. The accuracy of female radiographs was 69.36% and of male was 59.22%. The accuracy within 1-year range was 85.55% for female and 92.89% for male radiographs. The accuracy within 2-year range was 90.75% and 98.82% for female and male radiographs, respectively.

is selected on a random basis as a validation set to tune the hyperparameters of RBF kernel and select appropriate features as well. The results obtained in this step are applied on the remaining 80% of the total data for classification. The accuracy of female radiographs was 69.36% and of male was 59.22%. The accuracy within 1-year range was 85.55% for female and 92.89% for male radiographs. The accuracy within 2-year range was 90.75% and 98.82% for female and male radiographs, respectively.
the CAD system is considerably improved after applying two ROIs simultaneously. The results of the estimated bone age reveal that better results are obtained in case of female radiographs, which assures the fact that female bones develop faster than that of males.\textsuperscript{[19]} Although the proposed approach yielded good results, especially for female radiographs, it is low time-consuming, and easy to implement, more data is needed to affirm the efficiency of this approach on BAA of Iranian children.

Financial support and sponsorship

The authors would like to thank the University of Isfahan and the Legal Medicine Research Center, Legal Medicine Organization, Tehran, Iran, for their support. The grants of this research have been registered under the 95-12442 and 95-143566 numbers in the University of Isfahan and Legal Medicine Organization of Iran, respectively.

Conflicts of interest

There are no conflicts of interest.

References

BIOGRAPHIES

**Farzaneh Dehghani** received the B.Sc. Degree in Biomedical Engineering from the Islamic Azad University of Khomeinishahr Branch, Isfahan, Iran, in 2013. She also received her M.Sc. in Biomedical Engineering from the University of Isfahan, Isfahan, Iran, in 2018. Her main research area is image processing, machine learning, and computer vision.

**Email:** farzaneh70d@gmail.com

**Alireza Karimian** is an associate professor in the Department of Biomedical Engineering, University of Isfahan, Iran. He has passed successfully a 1-year fellowship research in the field of medical physics in La Sapienza University, Rome, Italy, under grant of ICTP. He was successful to carry out his 9-month sabbatical and post Ph.D. fellowship at Johns Hopkins University, USA, in the field of mathematical modeling in medicine (2016–2017). His research interests are image and signal processing, and serious entertainment development, focusing on mental/physical therapy, and education.

**Email:** rasti@eng.ui.ac.ir

**Mehri Sirous** received her degree in Doctor of Medicine and Radiology specialization from Isfahan University of Medical Sciences, Isfahan, Iran. She is an associate professor in the Department of Radiology, Isfahan University of Medical Sciences. She successfully passed a 6-month CT course and a 2-month Doppler course in Shahid Beheshti University of Medical Sciences in 2001. Furthermore, she completed a 6-month MRI course in Isfahan University of Medical Sciences in 2014.

**Email:** sirous@med.mui.ac.ir

**Javad Rasti** received his Ph.D. in Computer Engineering from University of Isfahan in 2013, with thesis entitled “An Image Segmentation Method Based on Intelligent Algorithms for Object Recognition in Outdoor Scenes.” He also received a M.Sc. in Computer Architecture and his B.Sc. in Computer Hardware Engineering from University of Isfahan in 2003 and 2006, respectively.

He is currently an assistant professor in Biomedical Engineering Department in University of Isfahan and also the funder/head of University of Isfahan Center of Video Games. His main research interests are image and signal processing.

**Email:** rasti@eng.ui.ac.ir

**Ali Soleymanpour** received his degree in Doctor of Medicine and Forensic Medicine and Toxicology specialization from Tehran University of Medical Sciences, Tehran, Iran. His main research interests are forensic sciences, medical ethics, physical examination, and crime scene investigation. He has been the lecturer of some courses such as medical ethics and forensic sciences too.

**Email:** asoleymanpour@yahoo.com