Fully Automatic Bone Age Estimation from Left Hand MR Images

Darko Stern[1,⋆], Thomas Ebner[1,2], Horst Bischof[1], Sabine Grassegger[2,3], Thomas Ehammer[2,3], and Martin Urschler[1,2,3]

[1] Institute for Computer Graphics and Vision, BioTechMed, Graz University of Technology, Austria
[2] Ludwig Boltzmann Institute for Clinical Forensic Imaging, Graz, Austria
[3] Department of Legal Medicine, Medical University of Graz, Austria

Abstract. There has recently been an increased demand in bone age estimation (BAE) of living individuals and human remains in legal medicine applications. A severe drawback of established BAE techniques based on X-ray images is radiation exposure, since many countries prohibit scanning involving ionizing radiation without diagnostic reasons. We propose a completely automated method for BAE based on volumetric hand MRI images. On our database of 56 male caucasian subjects between 13 and 19 years, we are able to estimate the subjects age with a mean difference of 0.85 ± 0.58 years compared to the chronological age, which is in line with radiologist results using established radiographic methods. We see this work as a promising first step towards a novel MRI based bone age estimation system, with the key benefits of lacking exposure to ionizing radiation and higher accuracy due to exploitation of volumetric data.

1 Introduction

Bone age estimation (BAE) of living individuals or human remains has recently received conspicuous attention due to increasing demands in clinical and legal medicine, like: growth predictions for prognostic and therapeutic purposes, diagnosis of endocrinological diseases [1], victim identification after disasters [2], assessing asylum seekers entering a country without proper identification documents [3], or preventing age manipulation in junior-level sports competitions [4]. Being based on bone ossification, the clinically established BAE methods employ conventional X-ray examinations of the hand bones to provide means for an objective and reliable age estimation up to 19 years. In cases involving subjects near to the legal majority age, this examination is accompanied by a CT of the clavicle bone, which is one of the last bones to finish ossification, and a panoramic X-ray of the third molar teeth, to enable BAE up to 24 years. By examining the ossification and epiphyseal plate fusion in X-ray images of the hand, radiologists perform BAE based on the hand bones according to the Greulich-Pyle [5] (GP)

⋆ This work was partly supported by the city of Graz (A16-21628/2013) and a European Community FP7 Marie Curie Intra European Fellowship (331239).

P. Golland et al. (Eds.): MICCAI 2014, Part II, LNCS 8674, pp. 220–227, 2014.

or the Tanner-Whitehouse [6] (TW) systems. Building upon the visual comparison of the whole hand with a reference atlas, the GP method is fast and easy to use compared to TW, but shows lower accuracy and larger inter-observer variability [3]. Exploiting the fact that the aging progress is not the same for all bones of the hand, the BAE is improved in TW method by visually comparing each individual hand bone to the X-ray image based atlas and combining these scores to a final age estimate. Automated image analysis methods have recently started to appear, most prominently the X-ray image based BoneXpert method [7], which successfully mimic the atlas matching procedure.

A severe drawback of radiographic BAE techniques is radiation exposure, which can not be justified for screening healthy children and adolescents, especially in applications of legal or sports medicine. Thus, non-invasive magnetic resonance imaging (MRI) has gained in importance for BAE [1, 4], since many countries prohibit scanning involving ionizing radiation without diagnostic reasons. The higher cost of MRI compared to a single X-ray examination might seem as a drawback, but the possibility of acquiring hand bones, teeth and clavicle in a single MR scan session amortizes this cost. Another benefit of MRI compared to projective 2D X-ray examinations is its volumetric nature, which may provide a foundation for more accurate and reliable BAE. Up to our knowledge, all current methods proposed for BAE in MR images are restricted to best-view cross sections to imitate the estimation methods developed for X-ray images [1, 8].

We present a completely automated method for BAE based on volumetric MR images of the hand. We see this work as a first step towards establishing a novel MR imaging based method for age estimation. Key benefits are the lack of exposure to ionizing radiation and higher accuracy due to exploiting 3D volumetric data. Due to our intended application, i.e. age estimation of asylum seekers without proper documents, the method is evaluated on adolescents, however, the algorithm design supports future extension to a more comprehensive age group involving children as well. The bone age can be estimated from the fusion stages of the epiphyseal gap located between epiphysis and metaphysis of the hand bones. We propose to locate the region of the epiphyseal gap in individual bones and extract nearby features that discriminate the fusion stages of the epiphyseal gap in our data set based on the known chronological age. Detection of the gap region and mapping the extracted features to the chronological age can be seen as a regression task, which we model using the powerful and efficient random forest (RF) framework [9]. Similarly to the TW method [6], but in an automatic fashion, we obtain an estimated age by fusing the BAE of individual hand bones.

2 Method

Our proposed approach, as illustrated in Fig. 1, first locates individual hand bones from MR images, followed by a localization of the expected epiphyseal gap position in the bone. A bounding box capturing the region of the epiphyseal gap is extracted at this location, defined by the diameter of the bone and a statistical height estimate derived from training data. The bounding box allows to crop the

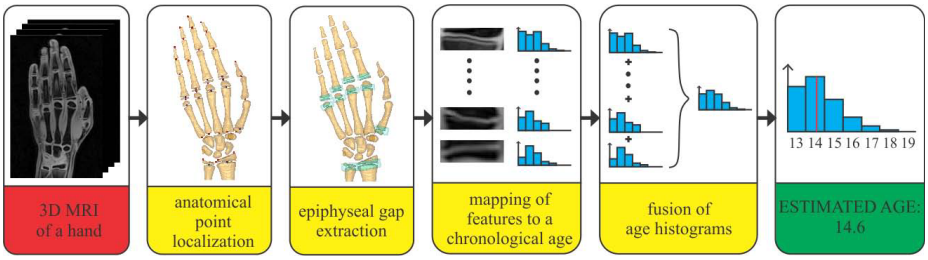


Fig. 1. Flowchart diagram for the proposed method

3D region, where relevant features for bone age estimation can be expected. After feature extraction a regression step maps features to a chronological age, and the fusion of per-bone age estimates determines the final BAE.

2.1 Extraction of Epiphyseal Gap Images

Localization of hand bones and epiphyseal gaps was inspired by the work on anatomical structure localization in [9] and [10]. Based on regression RF, we developed a fully automated method for localization of anatomical landmarks in 3D MR images [11]. Its main idea is to gradually decrease the area where landmark positions are expected. This is achieved with a two-step multiscale approach, that first makes a prediction of the coarse landmark positions analyzing the whole shape of the hand and using global feature information from the whole image. After finding the area, where the precise landmark locations are expected, the second step uses more localized information for prediction. Thus, we obtain more accurate landmark localizations compared to [9] and also avoid to explicitly model the spatial relations of anatomical landmarks with a Markov Random Field as in [10]. With this approach we localize 28 landmarks on the hand bones, which we use as anchor points to crop each individual bone based on its maximally expected extent as estimated from our data set and to rotate each obtained image to align the bone axes to a standard orientation.

For each bone image now the same method for anatomical landmark localization is used to identify the position of the epiphyseal gap c_G. The bounding box of the epiphyseal gap is defined with the center in c_G, the estimated diameter of the bone in the axial slice where c_G is located, and a statistically estimated average height of the gap. Since bones have a cylindrical structure at the location of the gap, the diameter of the bone can be estimated as an average distance of bone edges to the center of the bone, i.e. c_G. Bone edges are localized by radially emitting rays originating from c_G, and finding the maximum of the dot product between the image gradient vector and the unit ray vector along the rays. The gap image is generated by cropping the bone image with the obtained bounding box and then normalizing the size of the image. Thus, all gap images are aligned to allow features discriminating age to be generated in the fixed coordinate system of the epiphyseal gap.

2.2 Age Estimation for Individual Bone

To cope with the large number of features that can be generated from the epiphyseal gap region and due to their non-linear relation with bone age, we again use a regression RF to perform feature selection in the gap image. Thus, we let the RF select those features that best map the fusion stage of the epiphyseal gap to the chronological age. We treat each gap image and its selected features together with the chronological age as one data input to train the regression RF.

At each node of the RF, features are generated in one of the three following ways: the intensity value at a randomly generated point, the intensity difference between two randomly generated points, or the average intensity value along the line that connects two randomly generated points. A random selection between these three types is performed. All these features can be efficiently computed and provide the required discriminative capabilities for differentiating epiphyseal gaps over the age range. According to the maximization of information gain (IG), a feature $f(\boldsymbol{\theta_j})$ and threshold τ_i that best discriminates over the ages in the set of the gap images S reaching the node, are stored in the forest:

$$IG = |Var(S)| - \sum_{i \in \{L,R\}} \frac{|S_i|}{|S|} |Var(S_i)|. \tag{1}$$

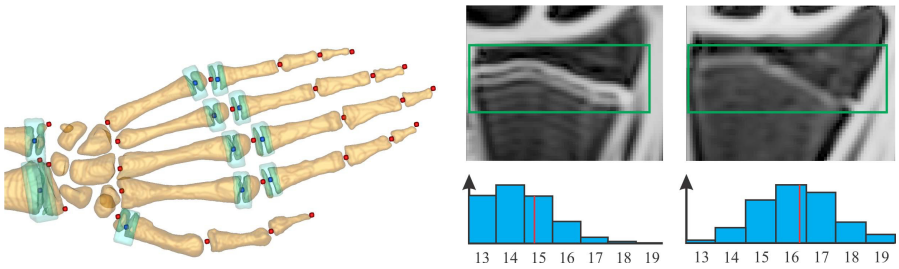
Here $\boldsymbol{\theta_j}$, $j \in \{1..N_F\}$, are the parameters of the features, $Var(\cdot)$ is the variation of a set of gap images, S_L, S_R are left and right split subsets of the gap images, respectively, and node splits are defined according to the binary test $f(\boldsymbol{\theta_j}) > \tau_i$, $i \in \{1..N_T\}$. Node splitting is done recursively and stops, when the maximum tree depth D is reached or there is no improvement in IG. The histogram of the age distribution for the gap images that reach the leaf node are generated and stored in the tree, to be available when testing an image.

During testing, each gap image is pushed through all trees in the RF. Until the gap image reaches the leaf node, a feature response of the gap image is generated based on the feature parameters $\boldsymbol{\theta_j}$ and the image is passed to the left or right child node depending on the result of the comparison with the threshold value τ_i. The estimated age of a bone $EAB(b)$ is obtained as a sum of the histogram $h(b;t)$ in the reached leaf nodes of the trees $t \in \{T\}$.

2.3 Age Estimation

The aging progress of epiphyseal gaps is not the same for all bones of the hand [6]. Thus, the aging of the epiphyseal gap in the phalanges and carpal bones is faster than aging of the ulna and radius bone, except for the thumb metacarpal bone which lies in between. The final estimated age of a subject EA is therefore obtained as a sum of the estimated age of the individual bones:

$$EA = \sum_b w(b) \cdot EAB(b) = \sum_b w(b) \cdot \sum_{t=1}^{T} h(b;t), \tag{2}$$



(a) (b)

Fig. 2. (a) Ground truth of the 28 anatomical hand bone landmarks (red) and 11 landmarks (blue) and bounding boxes (green) of the epiphyseal gaps. (b) Epiphyseal gap of a radius bone (green box) presented in the best view cross-section of a 14.6 (left) and 17.4 (right) years old subject. The estimated age of the radius bone (left: 15.4 years; right: 16.7 years, see red lines) is obtained as mean value of the resulting histograms of the regression RF. The final age is estimated from the 11 hand bones.

where the weight coefficient of the bone $w(b) \in \{0, 1\}$ includes or excludes a bone b from the estimation depending on the prior knowledge about the limitations in maximum age estimation of each bone.

3 Materials and Experimental Setup

Materials. Our dataset of left hand T1-weighted 3D gradient echo MR images consists of scans from 56 caucasian male volunteers of known chronological age equally distributed in seven age groups between 13 and 20 years (mean±std 16.5 ± 2.0, min 13.0, max age 19.9 years). The average dimension of the volumes is $294 \times 512 \times 72$ voxels and the average voxel size is $0.45 \times 0.45 \times 0.9 mm^3$. A scientist well experienced in the analysis of 3D MR images manually annotated 28 anatomical landmarks of the hand bones and additionally the center position of the epiphyseal gap for 11 bones (radius bone, ulna bone, five metacarpal bones and four proximal phalanges bones) that are used in BAE (see Fig. 2(a)).

Experimental Setup. For localization of 3D anatomical landmarks we use forests with maximal tree depths of $D = 15$ and a small number of trees $T = 8$. At each split node $N_F = 20$ random candidate features and $N_T = 10$ candidate thresholds are generated. We use this design, since all voxels of the respective training images vote for distance estimates to the landmarks, and we found larger forests to not improve the detection performance. The results of our localization forests are evaluated in a cross-validation setup with five rounds. In each round, we randomly split the 56 input volumes into 44 training and 12 testing images.

For the BAE step, we construct forests for radius, ulna, thumb metacarpal and a single forest for the other metacarpals, as well as a single forest for the phalanx bones. We group these bones since they show a similar appearance and aging progress of the epiphyseal gap. Due to the limited number of images in

our data set, our forests for age estimation contain trees with a small maximum depth ($D = 4$) but with a large number of trees ($T = 1000$). Each leaf node of the forest stores a histogram over the age range with a one year resolution between 13 and 19 years. Extremal age values are determined by the youngest subject in our data set and the age when hand bone ossification is considered to be finished. Age estimation results are computed in a leave-one-out cross-validation.

We implemented our methods in C++ on an 8-core Intel(R) Core(TM) i7 CPU, the runtime for the localization of anatomical landmarks in 3D MR images is about 400s and for the age estimation less than 5s. Training of the RF for the localization of anatomical landmarks took around 24 hours and the training of the RF for the age estimation around 10s. Our algorithms currently are not optimized and could be sped up further by GPU based implementation.

4 Results

The performance of the method for localization of anatomical landmarks in 3D MR images was measured quantitatively by computing the Euclidean distance (±standard deviation) between the obtained and the ground truth landmark position for the 28 anatomical landmarks of the bone, as well as for the epiphyseal gap position in the 11 hand bones. The overall localization error for the 28 anatomical hand bone landmarks was 1.4±1.5mm [11] and for the epiphyseal gap position in the 11 hand bones it was 1.1±0.8mm. The mean difference between the diameter of the bounding box and the size of the bone at the location of the gap was 0.6±0.5mm. The mean difference between ground truth chronological and estimated age in our data set was 0.85±0.58 years. Figure 3 shows box-whisker plots that incorporate the results for all images and for each year separately. The error in age estimation for each bone separately is also shown in Fig. 3. Detailed results of radius bone age estimation for a selected 14.6 and 17.4 years old subject can be seen in Fig. 2(b).

5 Discussion

Our results show that the use of the regression RF framework enables not only the localization of the epiphyseal gap regions in 3D MR images of the hand, but also allows finding discriminative features in the gap regions, which are relevant for age estimation. For detecting epiphyseal gaps of individual bones, we rely on an anatomical landmark localization procedure determining first landmarks between the bones (see Fig. 2(a)), and from there accurately locate the epiphyseal gap regions [11]. This is different compared to related work in [9], where bounding boxes are directly extracted using global feature information from all over the image, while we focus on more local, and thus more precise feature information. Our regression RF shows an accuracy of 1.4±1.5mm in landmark localization compared to 1.5±1.1mm with the approach in [10], but reported on hand CTs. We find that epiphyseal gap localization with its accuracy of 1.1±0.8mm is sufficient for the following age estimation related feature extraction steps, since

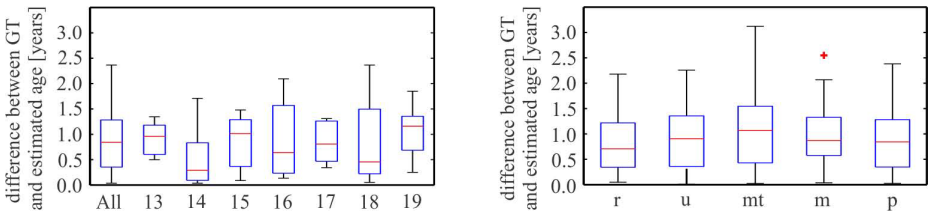


Fig. 3. (Left) Difference between the estimated and ground truth age for all subjects, separately for each age group. (Right) Differences of estimated and chronological age separately for each bone.

the goal is to roughly, but robustly locate the gap region, where the regression forest extracts the features to learn the age from.

Our regression forests for age estimation are easy to implement, since they do not require an intermediate bone segmentation step as in the BoneXpert system [7]. Their complex statistical shape/appearance model to segment bones from X-rays is very demanding to be extended to 3D, especially due to the tedious annotation required for training the generative model. In addition generative models are known to suffer in the presence of outliers, while our discriminative model promises robustness to outliers. The overall results for age estimation of adolescents (0.85 ± 0.58 years) are comparable with clinically established methods like manual Greulich-Pyle [5] or Tanner-Whitehouse [6] X-ray atlas comparison, with reported precisions from 0.5 up to 2.0 years depending on the age, sex and origin of the examined population [3]. BoneXpert [7], the most prominent automatic method for BAE from X-ray images, reports a deviation of 0.72 years, however, this result is obtained by comparing with the manual Greulich-Pyle [5] atlas matching method performed by radiologists, not with the chronological ground truth age as in our case. Therefore, no valid comparison can be made, since Greulich-Pyle solely gives an age estimate as well. Nevertheless, the data base on which the BoneXpert system is trained, i.e. more than 1700 X-ray images of boys and girls in the age range from 2 to 17 years, significantly exceeds our data set size. From the box-whisker plots of each separate age (Fig. 3), it can be seen that a larger difference between the estimated and ground truth age is present for the extreme ages in the histogram, i.e. 13 and 19 years. This may be caused by the notched representation of their age distribution in the histogram (see Fig. 2(b), left). Further, when the epiphyseal gap is in the final stage of the fusion, the gap features are scattered in the gap region and therefore harder to detect. This may be the reason for the larger difference between the estimated and ground truth age for 18 years. Additional uncertainty in extraction of the features relevant for age estimation may also come from motion artifacts that are manifested inside the bone with an image intensity similar to the intensity value of the gap at the age of 16 (Fig. 2(b), left). From Fig. 3, it can be seen that radius and ulna bone for the whole age range, give similar estimation error as the metacarpals and proximal phalanges bone by the age of 17. The highest

uncertainty is obtained for the thumb metacarpal bone, since it was trained separately from other metacarpal bones, i.e. on a smaller data base.

6 Conclusion and Outlook

Up to our knowledge, we have presented the first fully automatic skeletal bone age estimation method from 3D hand MR data. On our database of 56 male caucasian subjects between 13 and 19 years, we are able to estimate the subjects age with a mean difference of 0.85 ± 0.58 years compared to the chronological age, which is in line with results using established methods based on X-ray projections, thus involving radiation exposure. In future work we will evaluate our approach on a larger database and we will investigate how to extend our method to a larger age range involving children as well.

References

1. Terada, Y., Kono, S., Tamada, D., Uchiumi, T., Kose, K., Miyagi, R., Yamabe, E., Yoshioka, H.: Skeletal age assessment in children using an open compact MRI system. Magnet. Reson. Med. 69(6), 1697–1702 (2013)
2. Bassed, R.B.: Advances in forensic age estimation. Forensic. Sci. Med. Pathol. 8(2), 194–196 (2012)
3. Ritz-Timme, S., Cattaneo, C., Collins, M., Waite, E., Schuetz, H., Kaatsch, H., Borrman, H.: Age estimation: The state of the art in relation to the specific demands of forensic practise. Int. J. Legal Med. 113, 129–136 (2000)
4. Dvorak, J., George, J., Junge, A., Hodler, J.: Age determination by magnetic resonance imaging of the wrist in adolescent male football players. Brit. J. Sport Med. 41(1), 45–52 (2007)
5. Greulich, W.W., Pyle, S.I.: Radiographic atlas of skeletal development of the hand and wrist, 2nd edn. Stanford University Press, Stanford (1959)
6. Tanner, J.M., Whitehouse, R.H., Cameron, N., Marshall, W.A., Healy, M.J.R., Goldstein, H.: Assessment of skeletal maturity and predicion of adult height (TW2 method), 2nd edn. Academic Press (1983)
7. Thodberg, H.H., Kreiborg, S., Juul, A., Pedersen, K.D.: The BoneXpert method for automated determination of skeletal maturity. IEEE Trans. Med. Imag. 28(1), 52–66 (2009)
8. Tomei, E., Sartori, A., Nissman, D., Al Ansari, N., Battisti, S., Rubini, A., Stagnitti, A., Martino, M., Marini, M., Barbato, E., Semelka, R.C.: Value of MRI of the hand and the wrist in evaluation of bone age: Preliminary results. J. Magn. Reson. Im. 39, 1198–1205 (2013)
9. Criminisi, A., Robertson, D., Konukoglu, E., Shotton, J., Pathak, S., White, S., Siddiqui, K.: Regression forests for efficient anatomy detection and localization in computed tomography scans. Med. Image Anal. 17(8), 1293–1303 (2013)
10. Donner, R., Menze, B.H., Bischof, H., Langs, G.: Global localization of 3D anatomical structures by pre-filtered hough forests and discrete optimization. Med. Image Anal. 17(8), 1304–1314 (2013)
11. Ebner, T., Stern, D., Donner, R., Bischof, H., Urschler, M.: Towards automatic bone age estimation from MRI: Localization of 3D anatomical landmarks. In: Golland, P., Hata, N., Barillot, C., Hornegger, J., Howe, R. (eds.) MICCAI 2014. LNCS, vol. 8674, pp. 418–425. Springer, Heidelberg (2014)