

Bone Age Classification Using the Discriminative Generalized Hough Transform

Markus Brunk¹, Heike Ruppertshofen^{1,2}, Sarah Schmidt^{2,3}, Peter Beyerlein³, Hauke Schramm¹

¹Institute of Applied Computer Science, University of Applied Sciences Kiel, Germany

²Institute of Electronics, Signal Processing and Communication Technology, Otto-von-Guericke-University Magdeburg, Germany

³Dept. of Engineering, Technical University of Applied Sciences Wildau, Germany

markus@brunk-net.de

Abstract. We present an approach for automatic bone age classification from hand x-ray images using the Discriminative Generalized Hough Transform (DGHT). To this end, a region, characteristic for the bone age (e.g. an epiphyseal plate), is localized and subsequently classified to determine the corresponding age. Both steps are realized using the DGHT, whereat the difference of the approaches lies within the employed models. The localization model is able to localize the target region over a broad age range and therefore focuses on the common features of all ages. The model for the classification, in contrast, focuses on the age discriminating features. The classification model consists of several submodels, one for each age class, where each submodel contains information about its age characteristics as well as discriminating features. In a first test the new method was applied to classify images into the two classes 11–12 and 14–15 years and achieved of 95% correct classifications.

1 Introduction

The determination of bone age in the clinic is necessary to diagnose early development disorders of children, as well as in legal cases to determine, if an adolescent has reached the age of majority.

The two most common approaches for bone age classification are the manual methods from Greulich & Pyle (GP) [1] and from Tanner & Whitehouse (TW) [2]. In the first method a physician compares all hand bones in an x-ray image to reference images from an atlas to determine the correct age. In the latter method a subset of hand bones are classified independently and these results are then weighted and combined to determine the age. Since both methods are performed manually, they are time-consuming, user-dependent and require a substantial amount of expert knowledge.

A semi-automatic method is given by computer-assisted skeletal age scores (CACAS) [2], which automates the TW approach. However, the classification still has to be carried out manually. A fully automatic method is given by the BoneXpert-System [3], which determines the age according to the method of

GP or TW using active appearance models and regression. Fischer et al. [4] developed another fully automatic method using content-based image retrieval. In this method the regions of interest are compared to an image database and the similarity is computed using cross-correlation. The age of the best matches are combined for the final result.

In this work, we apply the discriminative generalized Hough transform (DGHT) [5] to automatically locate and classify regions of interest for bone age determination. For this purpose, two models are learned in a fully automatic manner; one model is used for the localization of the age characteristic region, the other for its classification to the corresponding age group.

2 Materials and Methods

The DGHT [5] is a method which has been successfully applied to object localization in medical images. The method combines the Generalized Hough Transform (GHT) [6] with the Discriminative Model Combination (DMC) [7], which is used to assign individual weights to the model points of the GHT models. The employed point models, representing the target shape, are thereby generated from a number of images by extracting the edge points from a region around the target point. The localization results are very robust and the training and evaluation runs fully automatic [5]. Here, the DGHT is used to solve two tasks: (i) localization of an age characteristic region and (ii) classification to determine the patient's bone age.

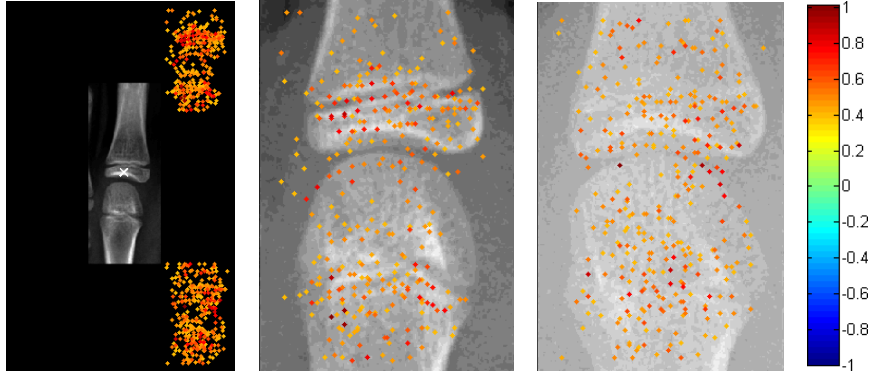
2.1 Localization

The characteristic region is defined as finger base joint of the middle finger. To obtain a general localization model, capable of robustly detecting the finger base joints in a wide range of ages, we used image data from various ages for initial model generation and training. Due to the strong similarity of the four finger joints over the different ages, we cannot generate a model, which localizes only the middle finger without any confusion with the other joints. However, by fitting a mean distance model of the joints to the peaks visible in the Hough space the position of the middle finger can be robustly determined.

2.2 Classification

In the second step, a region of interest around the localized point is extracted and used for classification. For this purpose, a new kind of GHT models is trained, containing a combination of age-specific submodels (Fig. 1). An initial model is generated in the same way as described in Sec. 2.4 using training data from all considered age classes. Thereafter this model is duplicated for each age class and combined into a single classification model using class specific offsets. Each class of the training data is assigned the same offset as the corresponding submodel. In the training step, each submodel is adapted to the corresponding class using

Fig. 1. Illustration of the trained classification model overlaid on an example image extract and the two submodels displayed upon example images of the same class. We show the model point weights color-coded with the scale visible on the right.



the DMC weighting procedure; thereby class specific model points get a higher weight than common points; points belonging to rivaling classes might also be assigned negative weights.

Using the trained model in the GHT, it extracts the age specific characteristics of an image by voting for these characteristics in the resulting Hough space. The most distinct region of the Hough space, also containing the highest vote, is the region with the highest number of matching characteristics and corresponds to the submodel of the correct class.

2.3 Material and Experimental Setup

For localization we used 158 left hand male and female x-ray images of the ages 11–15. The skeletal age was manually classified by a medical expert using the GP-method and is used as ground-truth. Each image has an isotropic resolution of 0.1 mm and a size of approx. 2000×2000 pixel. For classification 69 images of only female patients are used to account for gender-specific differences. In a first study we try to separate two classes: (1) 11–12 and (2) 14–15 years. Example images of the two classes are given in Fig. 2.

2.4 Discriminative Generalized Hough Transform

To train a localization model, we used 75 images, from which 15 are taken for model generation. For validation of the localization functionality 83 images in the age range from 11 to 15 are used. Each image is down-sampled twice and the Hough space is quantized by a factor of two. Choosing these settings leads to a better performance without degrading classification rate.

For the classification we cut out the region of the localized middle finger and automatically assigned the image extract to one of the two considered age classes. We used 8 images per class for model generation, 14 images per class

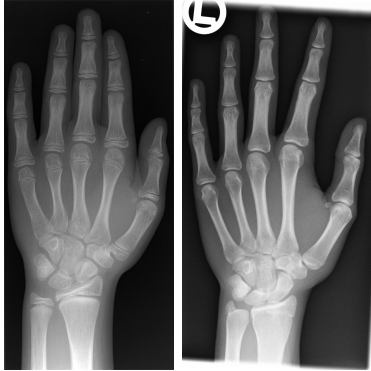


Fig. 2. Illustration of the test dataset containing x-ray images of the left hand. Left: 11 years, female; right: 15 years, female.

for training and performed the test on the remaining 41 images. These images are also downsampled twice. Since the classification task does not require a high precision, the Hough space was quantized by a factor of 10.

3 Results

In the localization experiments, we found the joint in 95% of the 83 test images. A localization is considered to be successful, if it is sufficiently exact to cut out the whole joint for subsequent classification.

In Fig. 1a the trained classification model with given offsets is illustrated. In the center an example x-ray image of a finger base joint is displayed with highlighted reference point. The two submodels for the age classes 11–12 and 14–15 are shown in the right top and bottom, respectively. Figures 1b and 1c show the two submodels in enlarged view with underlaid image extracts for better illustration. From these images it becomes obvious that the submodel for 11–12 years focuses on the epiphyseal plate, which has very strong positive weights in this region. The epiphyseal plate is less adhered at this age than at the age of 14–15 as can be seen in the underlying example images. In this test 95% of the test images were classified correctly; only two images were misclassified.

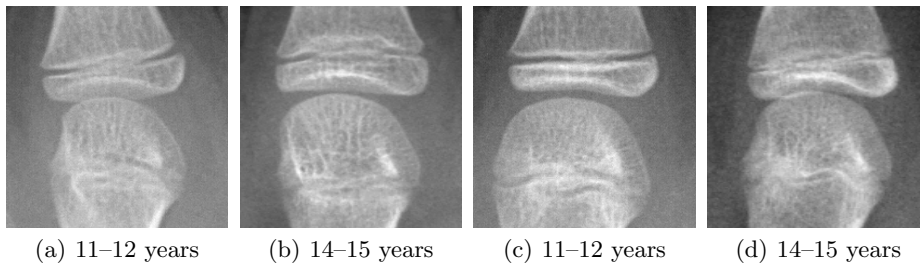


Fig. 3. Illustration of the two misclassified joints (left) in comparison with two training images of the two classes (right).

4 Discussion

The localization and classification results achieved with the described technique are promising. Both tests yielded a result of 95% of right localizations and classifications on unknown images.

In the localization experiment the 5% failures are due to confusions with the thumb joint and rotations of the hand. These mislocalizations will most likely be eliminated in future tests by improving the mean distance model used to identify the middle finger as described in Sec. 2.4.

In the classification test we observed two misclassifications in the class 11–12. These two images are shown in Fig. 3(a) and 3(b). Figure 3(c) and 3(d) show example images of correct classifications. Comparing these images it becomes obvious that in the range of the epiphyseal plate the bone is much more adhered in the misclassified images than would be expected for this age. From this, one can conclude that either the ground-truth is wrong or one joint is not sufficient for age classification since the ground-truth was obtained from the whole image with the GP method using all bones for classification. Potentially the untested joints or the carpal bone might appear more like the ground-truth class.

In the future, we will combine the classification results of all joints and the carpal bone to obtain a better and more detailed estimate for groups with a difference in age of less than one year. The results of each classification should then be combined in a suitable weighted manner, e.g. following the TW method. Furthermore, additional age specific submodels will be included into the classification model, allowing for an age classification with a larger number of classes. Further experiments will be conducted with larger number of images, also using public databases for better comparison of classification performance.

Acknowledgement. We thank Dr. Kayser, University Medical Center Schleswig-Holstein, Germany for providing the x-ray images used in this study.

References

1. Greulich WW, Pyle SI. Radiographic atlas of skeletal development of the hand and wrist. 2nd ed. Stanford University Press; 1976.
2. Tanner JM, Healy MJR, Goldstein H, et al. Assessment of skeletal maturity and prediction of adult height (TW3 method). 3rd ed. WB Saunders; 2001.
3. Thodberg HH, Kreiborg S, Juul A, et al. The boneXpert method for automated determination of skeletal maturity. *IEEE Trans Med Imaging*. 2009;28(1):52–66.
4. Fischer B, Brosig A, Welter P, et al. Content-based image retrieval applied to bone age assessment. In: *Proc SPIE*. vol. 7624; 2010. p. 35.
5. Ruppertshofen H, Lorenz C, Beyerlein P, et al. Fully automatic model creation for object localization utilizing the generalized Hough transform. *Proc BVM*. 2010; p. 281–5.
6. Ballard DH. Generalizing the Hough transform to detect arbitrary shapes. *Patt Recogn Lett*. 1981;13(2):111–22.
7. Beyerlein P. Discriminative model combination. In: *Proc ICASSP*; 1998. p. 481–4.