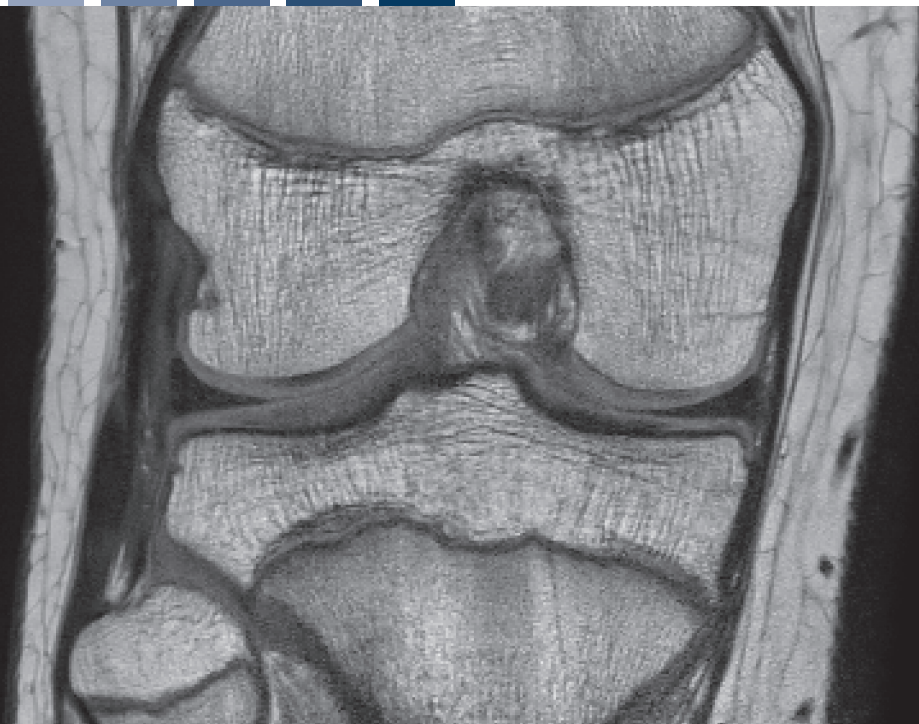


Towards Automated Age Estimation of Young Individuals

A New Computer-Based Approach Using 3D Knee MRI

Markus Auf der Mauer



**Towards Automated Age Estimation
of Young Individuals:
A New Computer-Based Approach
Using 3D Knee MRI**

Vom Promotionsausschuss der
Technischen Universität Hamburg
zur Erlangung des akademischen Grades
Doktor der Naturwissenschaften (Dr. rer. nat.)

genehmigte Dissertation

von

Markus Auf der Mauer

aus

Caracas, Venezuela

2020

1. Gutachter: Prof. Dr. habil. Michael M. Morlock
2. Gutachter: Prof. Dr. Dennis Säring

Tag der mündlichen Prüfung: 28. Februar 2020

Berichte aus der Medizinischen Informatik und Bioinformatik

Markus Auf der Mauer

**Towards Automated Age Estimation
of Young Individuals**

A New Computer-Based Approach Using 3D Knee MRI

Shaker Verlag
Düren 2020

Bibliographic information published by the Deutsche Nationalbibliothek

The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available in the Internet at <http://dnb.d-nb.de>.

Zugl.: Hamburg, Techn. Univ., Diss., 2020

Copyright Shaker Verlag 2020

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior permission of the publishers.

Printed in Germany.

ISBN 978-3-8440-7400-0

ISSN 1432-4385

Shaker Verlag GmbH • Am Langen Graben 15a • 52353 Düren

Phone: 0049/2421/99011-0 • Telefax: 0049/2421/99011-9

Internet: www.shaker.de • e-mail: info@shaker.de

Acknowledgements

There are no secrets to success. It is the result of preparation, hard work, and learning from failure.

Colin Powell

During the course of my PhD, I prepared each step of the project to be as efficient as possible, I worked hard to implement and evaluate my ideas, and I learned from failures to adapt and improve my strategies. I would like to thank the following people for making these steps possible through their guidance, support, and motivation.

First, my doctoral father Prof. Michael M. Morlock for his enthusiasm for the project and his academic guidance. I do not only appreciate his practical thinking but also the knowledge he transmitted during my studies.

My PhD supervisor Prof. Dennis Säring for his continuous, skilled, and dedicated support and guidance. His insight and knowledge in the fields of medical image processing and machine learning steered me through this research project. Our regular meetings and conversations were inspiring for me to think outside the box and pushed me in the right direction.

Eilin Jopp-van Well for her expertise in age estimation and together with my other research associates Jochen Herrmann, Michael Groth, Rainer Maas, Ben Stanczus, and Paul-Louis Pröve, for their collaborative effort and energy during data acquisition, journal publications and overall help throughout the project.

My former colleagues at the University of Applied Sciences of Wedel for their interest in my PhD work and the enriching experience working together for over three years.

Of course, my friends and family for their support, motivation, and understanding at all times. I am immensely grateful to my parents for laying the foundation to reach this milestone of my life. Nothing is more important than family.

Finally, I would like to express my deepest gratitude to my partner Julia Sabeike for always being there for me in both good and difficult moments. Your love, support, patience and understanding have made the success of this project possible.

Abstract

Background: Age estimation from medical images plays an important role in forensic medicine to determine the chronological age of individuals lacking legal documentation or to discriminate minors from adults. Current methods for imaging-based age estimation are labour-intensive, subjective, and involve radiation exposure. Recent studies indicate that magnetic resonance imaging (MRI) offers a viable alternative to established methods. The *goal of this work* is to develop a fully automated, computer-based, and non-invasive method to estimate the chronological age of male adolescents and young adults based on knee MRIs.

Materials and Methods: A total of 489 three-dimensional knee MRIs were acquired from 299 male Caucasian subjects aged 13 to 21 years. The dataset was expanded with numeric data of the subjects (anthropometric measurements and assessments of knee bone maturation). The proposed solution for automated age estimation is composed of three parts: (a) *pre-processing* to standardize the data, (b) *bone segmentation* via convolutional neural networks (CNNs) to extract age-relevant structures from the images, and (c) *age estimation*. Three different methods were investigated in part (c). *Method 1* (M1) is based on machine learning (ML) and uses the numeric data to solve the task. *Method 2* (M2) is composed of a CNN which takes in knee MRIs and outputs age predictions per image slice. Subsequently, an ML algorithm is trained on these predictions and on the numeric data to estimate a single and final age per subject. Finally, *Method 3* (M3) is a variant of M2 which incorporates the numeric data into the CNN trained on knee MRIs. Similar to M2, M3 predicts a final age per subject based on ML but using only the age predictions of the CNN.

Results: The best performing method is M2 and achieves a mean absolute error in age regression of 0.69 ± 0.47 years and an accuracy in majority classification of 90.93% using the 18-year-threshold.

Conclusions: The results demonstrate the potential of this approach for age estimation based on knee MRI and ML-techniques and is expected to improve further with the incorporation of additional datasets.

Keywords: Automated age estimation · MRI · Knee · Machine learning · Convolutional neural networks · Segmentation

Contents

Acknowledgements	i
Abstract	iii
List of Figures	vii
List of Tables	xi
List of Abbreviations	xiii
1 Introduction	1
1.1 Goal of the Work	4
1.2 Structure	5
2 State of the Art in Age Estimation	7
3 Materials	13
3.1 Study Population	13
3.2 Anthropometric Measurements	16
3.3 Knee MRIs	17
3.4 Growth Plate Ossification Stages	19
4 Image Pre-Processing	21
4.1 Image Import	22
4.2 Bias Field Correction	24
4.3 Automated Cropping	25
4.4 Normalization	30
4.5 Gold-Standard Segmentation	31
5 CNN-based Segmentation	35
5.1 Dataset Split	36
5.2 Augmentation	37
5.3 Resampling	39
5.4 CNN Architecture	40
5.5 Training	47

5.6	Post-Processing	49
5.7	Model Evaluation	50
6	Age Estimation	57
6.1	Method 1: ML-FEATS	58
6.1.1	Data Preparation	59
6.1.2	ML Setup	60
6.1.3	Training	63
6.2	Method 2: CNN-MRI	63
6.2.1	Data Preparation	65
6.2.2	CNN Architecture	69
6.2.3	Training	71
6.2.4	Age Regression	72
6.2.5	Majority Classification	72
6.3	Method 3: CNN-MIXED	73
6.4	Model Evaluation	74
7	Results	77
7.1	Preprocessing Results	77
7.2	Segmentation Results	84
7.3	Age Estimation Results	97
8	Discussion	111
9	Conclusions	123
A	Hardware and Software	125
B	Overview of MR Artefacts	127
C	Augmentation	131
D	Further Results on Segmentation	133
E	Further Results on Age Estimation	135
	Bibliography	149

List of Figures

1.1	Proposed solution for automated age estimation	5
2.1	Three-stage system for the ossification degree of knee growth plates	8
3.1	Average growth rates of boys and girls around puberty	14
3.2	Stacked age distribution of <i>Dataset A</i> , <i>Dataset B</i> , and <i>Dataset C</i>	15
3.3	Sitting height and lower leg length	16
3.4	MR image slices of all datasets	17
3.5	Three-stage system for the ossification degree of knee growth plates	20
4.1	Image pre-processing for 3D knee MRIs	21
4.2	Contents of a MetaImage header file	22
4.3	Medical image filename template	24
4.4	Bias field correction example	25
4.5	Extracting a standardized VOI from MRIs	26
4.6	Characteristic region for patch matching of coronal MRIs	27
4.7	Characteristic region for patch matching of sagittal MRIs	27
4.8	Automated cropping of knee MRIs using a patch matching algorithm . . .	28
4.9	Image segmentation tool developed to generate gold-standard segmenta- tions of 3D knee MRIs	32
4.10	Gold-standard segmentation and label map example for a knee MRI slice	33
5.1	CNN-based segmentation used for the bone detection in knee MRIs	35
5.2	Augmentation of knee MRIs	38
5.3	A multilayer perceptron	40
5.4	U-Net, a popular CNN architecture for segmentation	41
5.5	Convolution	42
5.6	Common activation functions of neural networks	42
5.7	Max pooling	43
5.8	The final architecture for CNN-based segmentation of knee MRIs	44
5.9	The building blocks of the CNN for segmentation	45
5.10	Training process of a neural network	47
5.11	Post-processing to enhance the segmentation results of the CNN	49
5.12	The building blocks of the 3D CNN for segmentation	54
6.1	Three methods for age estimation of male adolescents and young adults .	57

6.2	Correlation between anthropometric measurements and chronological age	59
6.3	Boxplots of chronological age vs. ossification stages	59
6.4	Analysis of parameters of machine learning algorithms	62
6.5	Train vs. validation losses for the age regression task using 2D MRIs without the bone segmentation step	64
6.6	Image preparation for the age estimation task via masking	65
6.7	Removal of sparse bone information	66
6.8	Augmentation of the training set of the CNN for age estimation	69
6.9	CNN architecture for age regression based on masked 2D knee MRIs	69
6.10	A “multi-input and mixed data” CNN architecture for age estimation	73
7.1	Bias field correction results	78
7.2	Bias field correction of images affected by MR artefacts	79
7.3	Automated cropping results of coronal MRI slices	82
7.4	Automated cropping results of sagittal MRI slices	83
7.5	Segmentation results for coronal MRI slices	85
7.6	Discrepancies between predicted and ground truth segmentations	86
7.7	Intermediate sum of feature maps	88
7.8	Visualization of low level features of the segmentation network and activation maximization	89
7.9	Visualization of high-level features of the segmentation network and activation maximization	89
7.10	Segmentation quality of a model trained on noisy data	91
7.11	Segmentation of uncropped coronal MRI	92
7.12	Segmentation of sagittal MRI using a merged and fine-tuned model	93
7.13	Training and validation loss for the merged model	94
7.14	Training vs. validation loss for age estimation models based on unmasked and masked MRIs	98
7.15	Absolute error between the true and predicted age per image slice	99
7.16	Absolute error between predicted and actual age per age group	100
7.17	Correct classification of a under-age subject	101
7.18	Predicted vs. true chronological age of test subjects from all five folds	108
7.19	ROC curve for the best model on majority classification	109
B.1	Motion artefacts in knee MRIs	128
B.2	Wrap-around artefacts observed in knee MRIs	129
B.3	Ringling artefacts	129
B.4	Intensity distortions	130
C.1	Augmentation before vs. after cropping	131

D.1	DSC score distribution from a model for segmentation	133
E.1	Age vs. ossification degree of the growth plates of the knee	136
E.2	Change in SKJ accumulated over a 2-year period	136
E.3	Distribution of the age prediction errors of a model using CNNs only vs. using CNNs and ML algorithms	137
E.4	Occlusion method to visualize important regions in the knee MRIs used for age estimation	139

List of Tables

3.1	Anthropometric measurements gathered for male subjects	16
3.2	Overview of study population, datasets, and MRI sequences	18
5.1	Data split into three sets for the segmentation task	37
6.1	Split per dataset and age group into three sets for the age estimation task based on coronal MRIs ($N = 185$)	68
7.1	Execution times of N4ITK algorithm	80
7.2	Execution times of the automated cropping step	81
7.3	Performance of various models on the segmentation task	95
7.4	Age regression performance of several model variants from Method 1	102
7.5	Age regression performance of several model variants from Method 2 using coronal knee MRIs	103
7.6	Age regression performance of several model variants from Method 2 using sagittal knee MRIs	104
7.7	Age regression performance of several model variants from Method 3	104
7.8	Performance on majority classification of several model variants from Method 1	105
7.9	Performance on majority classification of several model variants from Method 2 using coronal MRIs	106
7.10	Performance on majority classification of several model variants from Method 2 using sagittal MRIs	106
7.11	Performance on majority classification of several model variants from Method 3	107
8.1	Comparison of the performance of various segmentation models of the current work to other studies	115
8.2	Comparison of age regression performance between the current work and other studies	120
8.3	Comparison of majority classification performance between the current work and other studies.	121
A.1	Essential hardware available for this work	125
A.2	Most important Python and C++ libraries and frameworks	125
E.1	Performance of multiple models from Method 1 on age regression	140

E.2	Performance of multiple models from Method 2 on age regression using coronal MRIs	141
E.3	Performance of multiple models from Method 2 on age regression using sagittal MRIs	142
E.4	Performance of multiple models from Method 3 on age regression	143
E.5	Performance of other age regression models using coronal MRIs	144
E.6	Performance of other age regression models using sagittal MRIs	144
E.7	Performance of multiple models from Method 1 on majority classification	145
E.8	Performance of multiple models from Method 2 on majority classification using coronal MRIs	146
E.9	Performance of multiple models from Method 2 on majority classification using sagittal MRIs	147
E.10	Performance of multiple models from Method 3 on majority classification	148

List of Abbreviations

Abbrv.	Meaning
3D	= Three-Dimensional
AE	= Absolute Error
AGFAD	= international and interdisciplinary study Group on Forensic Age Diagnostics of the German Society of Legal Medicine
AI	= Artificial Intelligence
AM	= Anthropometric Measurements
ANN	= Artificial Neural Network
AUC	= Area Under the Curve
BAMF	= Bundesamt für Migration und Flüchtlinge
BFC	= Bias Field Correction
BL	= Baseline
BN	= Batch Normalization
CCW	= Counter-Clockwise
CNN	= Convolutional Neural Network
COG	= Center of Gravity
CPU	= Central Processing Unit
CT	= Computed Tomography
CV	= Cross-Validation
CW	= Clockwise
DCNN	= Deep Convolutional Neural Network
DF	= Distal Femur
DICOM	= Digital Imaging and Communications in Medicine
DO	= Dropout
DSC	= Dice Similarity Coefficient
DTC	= Decision Tree Classifier
EASO	= European Asylum Support Office
ELU	= Exponential Linear Unit
ETC	= Extremely Randomized Trees Classifier
ETR	= Extremely Randomized Trees Regressor
EU	= European Union

Abbrv.	Meaning
FC	= Fully-Connected
FN	= False Negative
FNR	= False Negative Rate
FOV	= Field of View
FP	= False Positive
FPR	= False Positive Rate
FU	= Follow-UP
GAP	= Global Average Pooling
GBC	= Gradient Tree Boosting Classifier
GBR	= Gradient Tree Boosting Regressor
GIGO	= Garbage In, Garbage Out
GMP	= Global Max Pooling
GNB	= Gaussian Naive Bayes
GP	= Greulich and Pyle
GUI	= Graphical User Interface
ID	= Identification
IQR	= Interquartile Range
IoU	= Intersection-over-Union
ITK	= Insight Segmentation and Registration Toolkit
KJC	= Knee Joint Cavity
KNC	= K-Nearest-Neighbors Classifier
KNN	= K-Nearest-Neighbors
LLL	= Lower Leg Length
LOOCV	= Leave-One-Out Cross-Validation
LR	= Linear Regression
LReLU	= Leaky Rectified Linear Unit
MAE	= Mean Absolute Error
MFS	= Magnetic Field Strength
ML	= Machine Learning
MLP	= Multilayer Perceptron
MRI	= Magnetic Resonance Imaging
MSE	= Mean Squared Error
N3	= Nonparametric nonuniform intensity normalization algorithm
N4ITK	= Improved N3 algorithm for ITK

Abbrev.	Meaning
NCC	= Normalized Cross-Correlation
OC	= Ossification Classes
PCA	= Principal Component Analysis
PCL	= Posterior Cruciate Ligament
PF	= Proximal Fibula
PReLU	= Parametric Rectified Linear Unit
PT	= Proximal Tibia
ReLU	= Rectified Linear Unit
RF	= Random Forests
RFC	= Random Forests Classifier
RFR	= Random Forests Regression
RMSE	= Root Mean Squared Error
ROC	= Receiver Operating Characteristic
ROI	= Region of Interest
SENSE	= SENSitivity Encoding
SGD	= Stochastic Gradient Descent
SKJ	= Score of the Knee Joint
SVC	= Support-Vector Classification
SVM	= Support-Vector Machine
SVR	= Support-Vector Regression
TE	= Echo Time
TN	= True Negative
TNR	= True Negative Rate
TP	= True Positive
TPR	= True Positive Rate
TR	= Repetition Time
TSE	= Turbo Spin Echo
TW2	= Tanner-Whitehouse method 2
UKE	= University Medical Center Hamburg-Eppendorf
VOI	= Volume of Interest
VTK	= Visualization Toolkit
