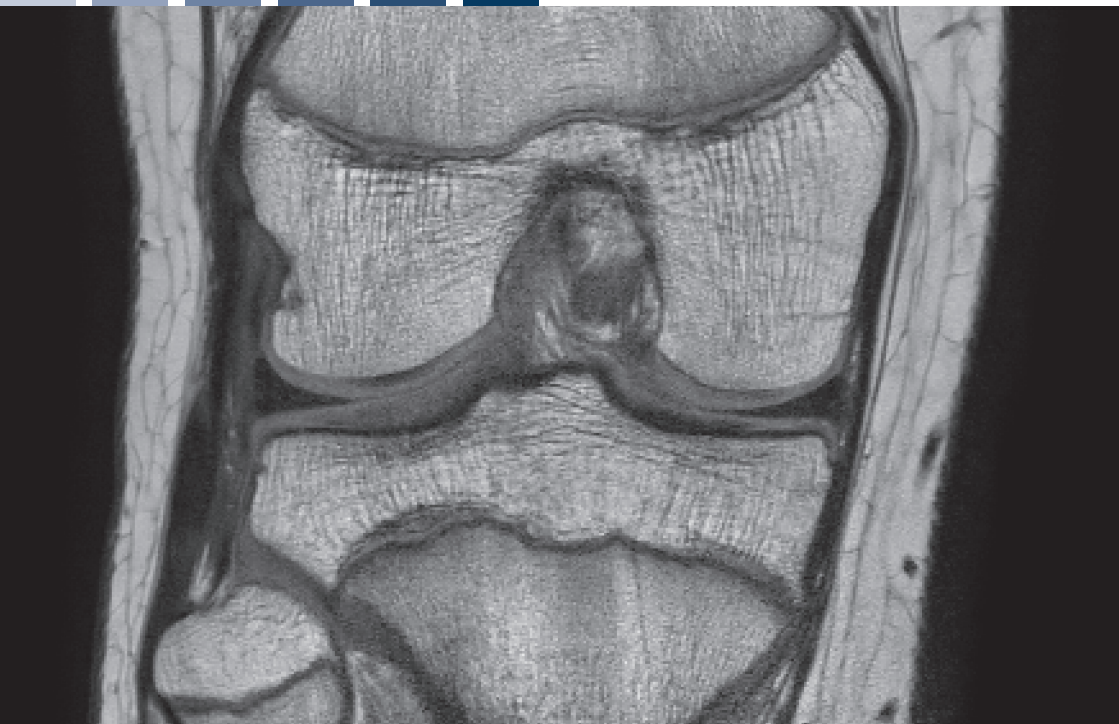
Towards Automated Age Estimation of Young Individuals

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

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

List of Figures

List of Tables

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

Abbrv.		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