*Alexander Oliver Mader, M.Sc.
alexander.o.mader@fh-kiel.de

GRAIL

Detection and Localization of Landmarks in the Lower Extremities Using an Automatically Learned Conditional Random Field

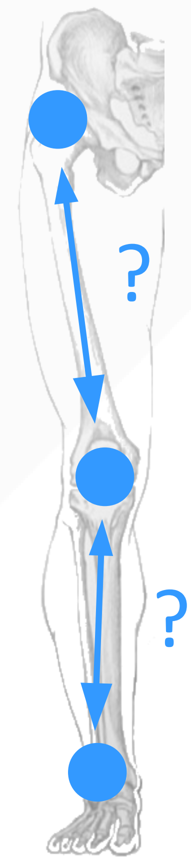
*Alexander O. Mader**, Cristian Lorenz, Martin Bergtholdt, Jens von Berg, Hauke Schramm, Jan Modersitzki, and Carsten Meyer

SPONSORED BY THE
Federal Ministry of Education and Research
PHILIPS
CAU Christian-Albrechts-Universität zu Kiel
FACHHOCHSCHULE KIEL University of Applied Sciences

Motivation

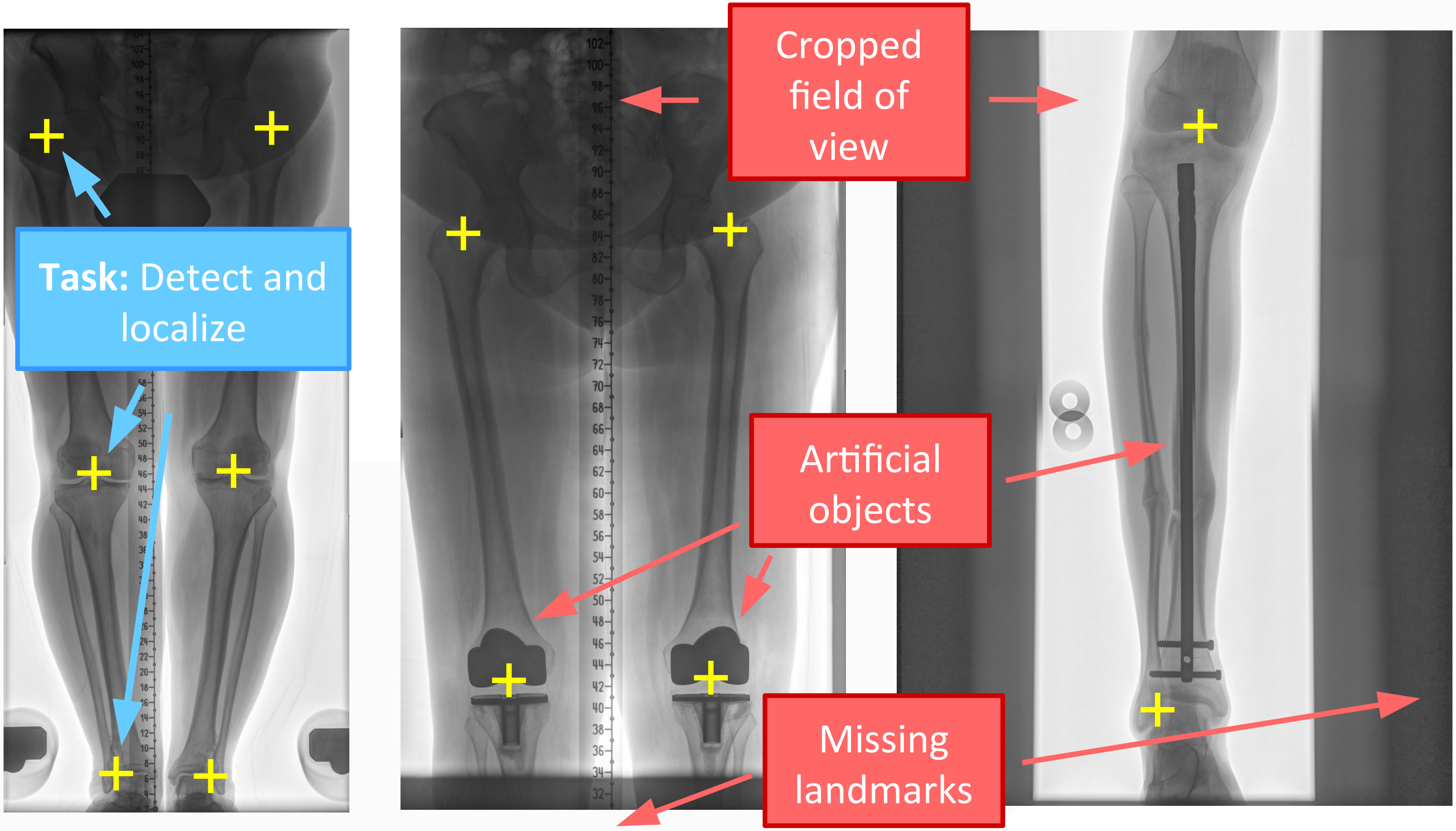
- Detection and accurate localization of anatomical landmarks is **clinically required by many tasks**
- High anatomical variability, outliers, restricted field of view, etc. render it a **rather hard task**

→ However, **frequent co-occurrence and spatial correlation** of anatomical landmarks can be exploited!



Alexander Oliver Mader
alexander.o.mader@fh-kiel.de

Introduction



Alexander Oliver Mader
alexander.o.mader@fh-kiel.de

FACHHOCHSCHULE KIEL
University of Applied Sciences

Basic Idea

1. Locally generate landmark hypotheses:

   For each landmark, independent of its existence, generate **multiple localization hypotheses** considering only local information

2. **Select hypotheses with Conditional Random Field (CRF)**:

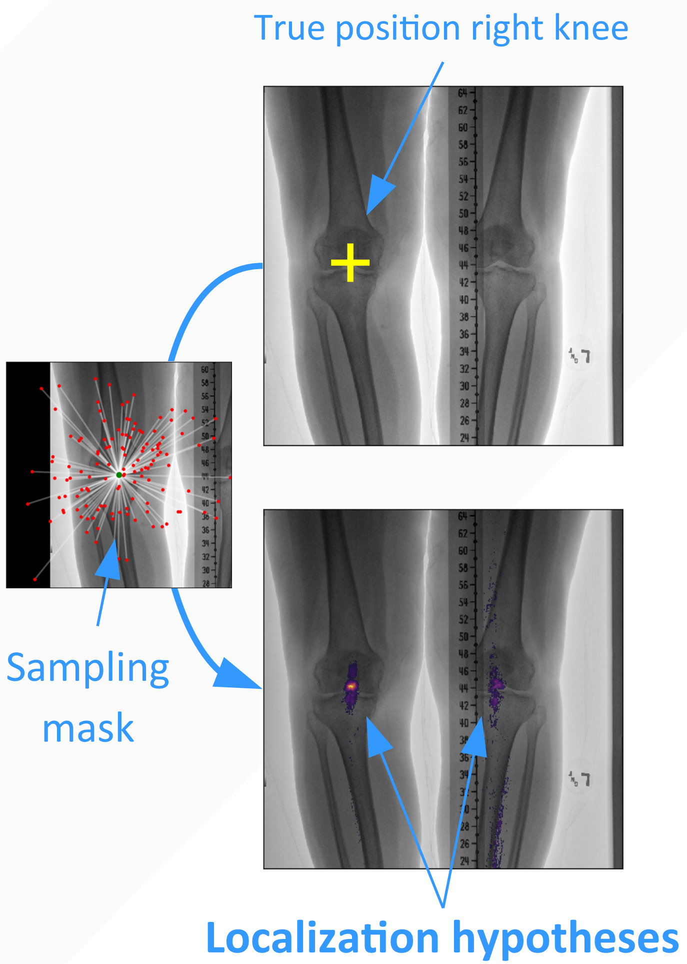
   Use **spatial knowledge between landmarks** to make an **informed selection** of localization hypotheses or assigning “missing” for each landmark

→ Going from *local* to *global* context!

Alexander Oliver Mader
alexander.o.mader@fh-kiel.de

1. Landmark Localization using Regression Tree Ensembles

- Use Ensemble of Decision Tree Regressors to *regress the “probability”* of position **x** being the true position of the landmark, given a feature vector $f(\mathbf{x})$
- Use *sampling mask* to compute local *intensity differences* as feature vector $f(\mathbf{x})$
- Repeat for each pixel **x** and get a pseudo probability map with potentially *many local maxima* (localization hypotheses)

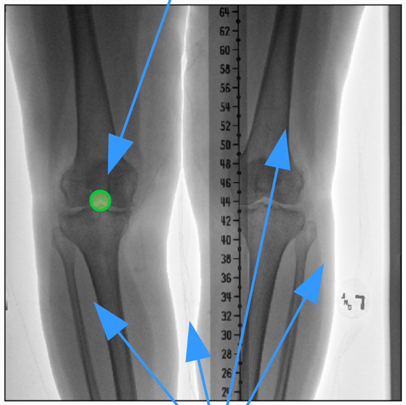


Alexander Oliver Mader
alexander.o.mader@fh-kiel.de

1. Discriminative Training of a Regression Tree

- Use *positive samples* to train a first intermediate tree
- Use intermediate tree to find *incorrect high responses* and use them as *negative samples*
- Iterating over all images, a more *refined tree is trained* after each sample generation
- Final tree is added to the ensemble

Positive samples: Features and regression values of all pixel with an Euclidean distance less than *R* to the landmark

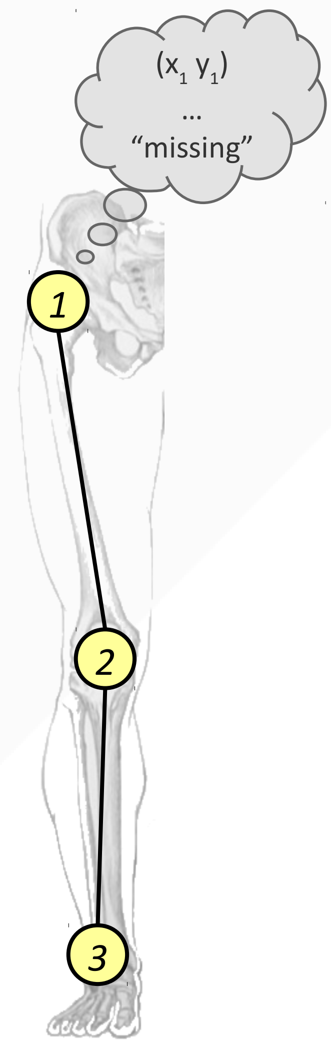


Negative samples: Apply intermediate tree and select incorrect high responses

Alexander Oliver Mader
alexander.o.mader@fh-kiel.de

2. Informed Landmark Selection using a CRF

- Use a CRF to model the *landmark positions as random variables*
- State space constrained to top *n* localization hypotheses per landmark
- *Additional "missing" state* for each landmark representing its absence
- Different energy potentials describe the quality of a joint selection over all landmarks
- Inference is used to find the best configuration, solving localization and detection in one step



Alexander Oliver Mader
alexander.o.mader@fh-kiel.de

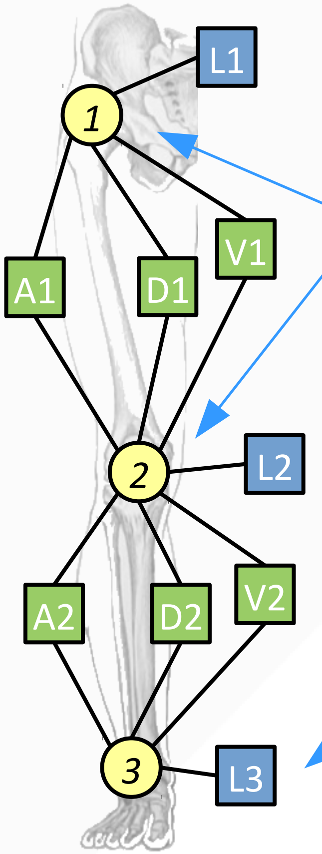
2. Informed Landmark Selection using a CRF

- Energy-based formulation to describe one selection w.r.t. different knowledge sources
- Weighted sum over T potentials $\varphi_j(\mathbf{S})$ or missing energies β_j for selection **S**:

$$E(\mathbf{S}) = \sum_{j=1}^{T} \lambda_j \cdot \begin{cases} \beta_j & \text{if } s_i = 0 \text{ for any } i \in \mathrm{Scope}(\phi_j) \\ \phi_j(\mathbf{S}) & \text{else} \end{cases}$$

- Use inference (exact A* search) to find the selection with the lowest energy:

$$\hat{\mathbf{S}} = \arg\min_{\mathbf{S} \in \mathcal{S}} E(\mathbf{S})$$



Landmarks as random variables; state space made of localization hypotheses plus "missing"

Different types of unary and binary potentials (knowledge sources)

Alexander Oliver Mader
alexander.o.mader@fh-kiel.de

2. Informed Landmark Selection using a CRF

- Energy-based formulation to describe one selection w.r.t. different knowledge sources
- Weighted sum over T potentials $\varphi_j(\mathbf{S})$ or missing energies β_j for selection $\mathbf{S}$:

$$E(\mathbf{S}) = \sum_{j=1}^{T} \lambda_j \cdot \begin{cases} \beta_j & \text{if } s_i = 0 \text{ for any } i \in \text{Scope}(\phi_j) \\ \phi_j(\mathbf{S}) & \text{else} \end{cases}$$

- Use inference (exact A* search) to find the ... ith the lo...

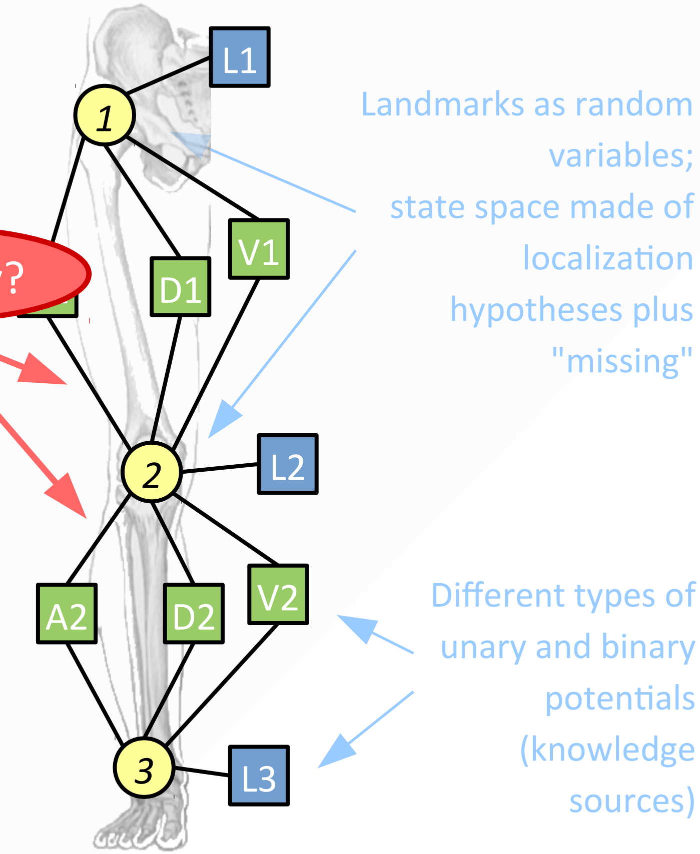
$$\hat{\mathbf{S}} = \arg\min_{\mathbf{S}\in\mathcal{S}} E(\mathbf{S})$$

"Missing" energy?

Topology?

Weights?

Important potentials?



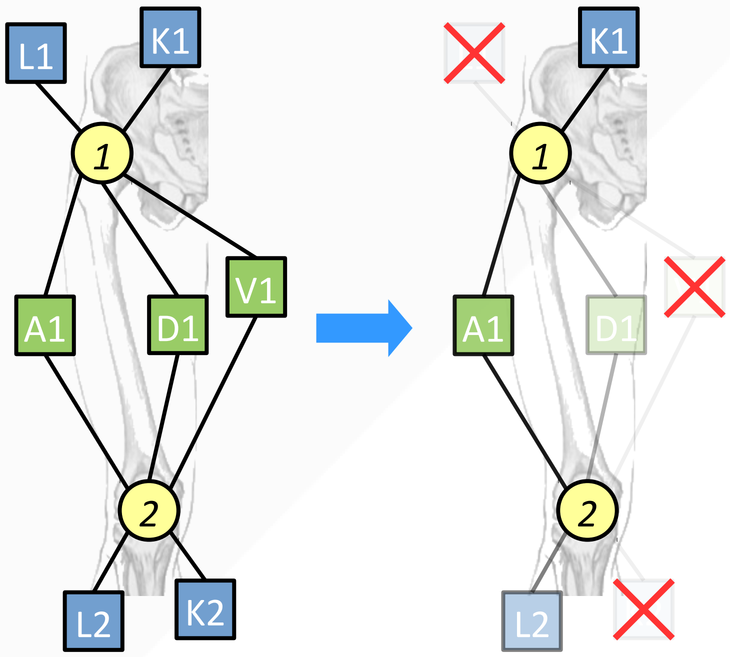
Alexander Oliver Mader
alexander.o.mader@fh-kiel.de

2. Learning Weights, Missing Energies and Topology

- **Central idea:** Starting from a *fully connected graph* containing instances of *all potentials from a "pool of potentials" (φ)*, learn weights (λ), energies (β) and deduce topology (λ = 0) in one novel step

- Pool may contain *arbitrary potentials* of arbitrary arity, but we limit it to unary localizer and binary probabilistic potentials using geometric features, i.e., angle, distance and vector

→ Ultimate goal is to **reduce complexity** while **increasing detection and localization accuracy** by automatically **learning important model components**!



Alexander Oliver Mader
alexander.o.mader@fh-kiel.de

2. Learning Weights, Missing Energies and Topology

- *Max-margin approach* to increase energy gap between correct and "best" incorrect selection, exact A* search by Bergtholdt et al. [1] to find the latter
- Optimization using *stochastic gradient descent* in form of the Adam algorithm [2] to efficiently perform the optimization

Margin to improve generalization

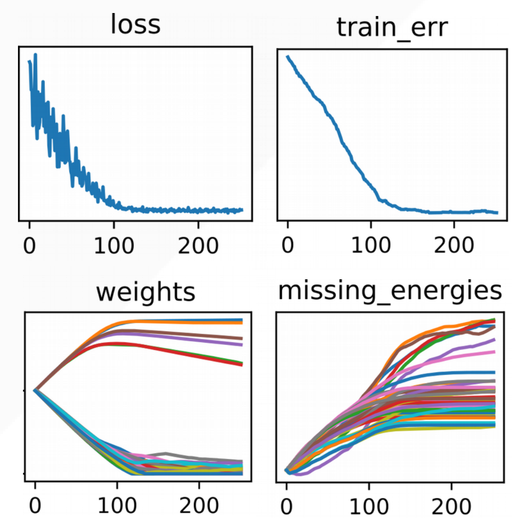
Correct selection

"Best" incorrect selection

$$L(\mathbf{\Lambda}, \boldsymbol{\beta}) = \frac{1}{K}\sum_{k=1}^{K} \max\left(0, m + E(\mathbf{S}_k^*) - E(\mathbf{S}_k^-)\right) + \theta \cdot \sum_{j=1}^{T} |\lambda_j|$$

Optimize loss function

L1 regularization to accelerate potential sparsification

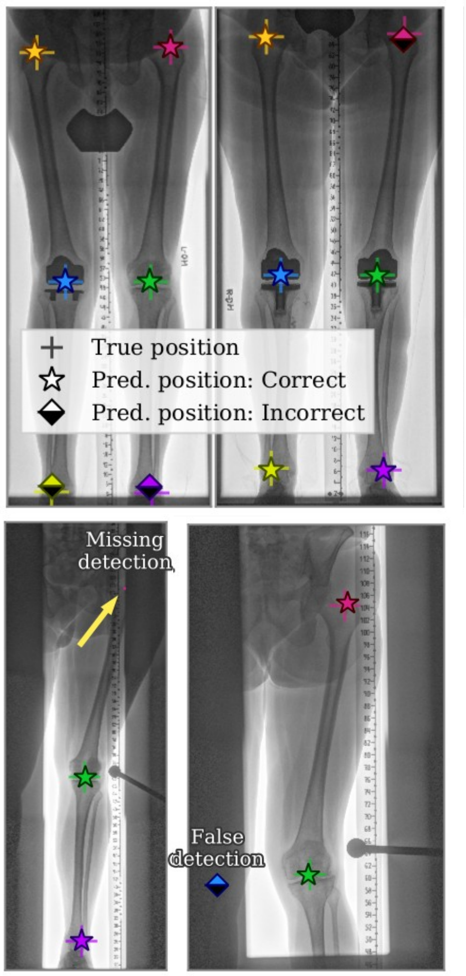


Alexander Oliver Mader
alexander.o.mader@fh-kiel.de

FACHHOCHSCHULE KIEL
University of Applied Sciences

Results

- 5-fold cross validation on an in-house dataset of 660 X-ray images of the lower extremities
- Due to restricted field of view, 2 to 6 different landmarks per image with 11.3% being altered by prostheses or pathologies
- Highly **effective detection with 98.1%** over 660 images and up to 6 different landmarks
- Correctly **detected and localized (< 10mm) on average 92.8%,** with landmark-specific **rates from 90.0% to 97.4%**
- Significantly **outperformed previous results** [3] by Ruppertshofen et al.
- **Removed** on average **23 of 51 potentials,** effectively **reducing the inference** time by **20.1%**

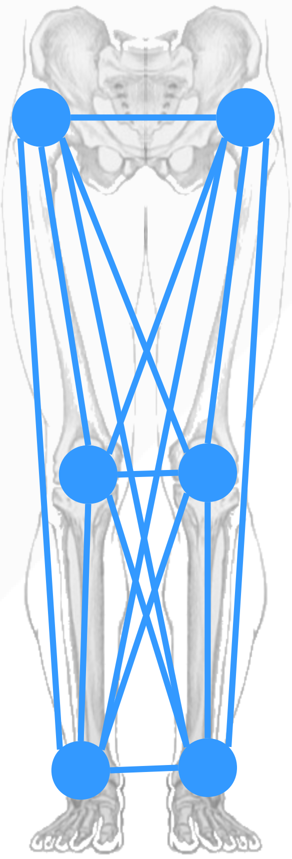


Alexander Oliver Mader
alexander.o.mader@fh-kiel.de

FACHHOCHSCHULE KIEL
University of Applied Sciences

Conclusions

- **Automated framework** for detecting and localizing spatially correlated landmarks
- Starting from a pool of potentials: Efficient one-shot **optimization of underlying CRF**, i.e., potential selection and weighting, energies for “missing”-state and topology
- **Allows application** to different datasets with little manual effort



Future Work:

- Usage of different localizers (e.g., CNN-based) to further improve accuracy
- Inclusion of higher-order potentials (using, e.g., higher-order clique reduction) and potentials using image information rather than only spatial features

Alexander Oliver Mader
alexander.o.mader@fh-kiel.de
FACHHOCHSCHULE KIEL
University of Applied Sciences

Questions / Discussion

References

1. M. Bergtholdt, J. Kappes, C. Schnörr. *“Learning of graphical models and efficient inference for object class recognition.”* Pattern Recognition. 2006.

2. D. Kingma, J. Ba. *"Adam: A method for stochastic optimization."* ICLR. 2015.

3. H. Ruppertshofen, C. Lorenz, S. Schmidt, P. Beyerlein, Z. Salah, G. Rose, H. Schramm. *“Discriminative Generalized Hough transform for localization of joints in the lower extremities.”* Computer Science - Research and Development. 2011.