LOCALIZATION AND LABELING OF POSTERIOR RIBS IN CHEST RADIOGRAPHS USING A CRF-REGULARIZED FCN WITH LOCAL REFINEMENT

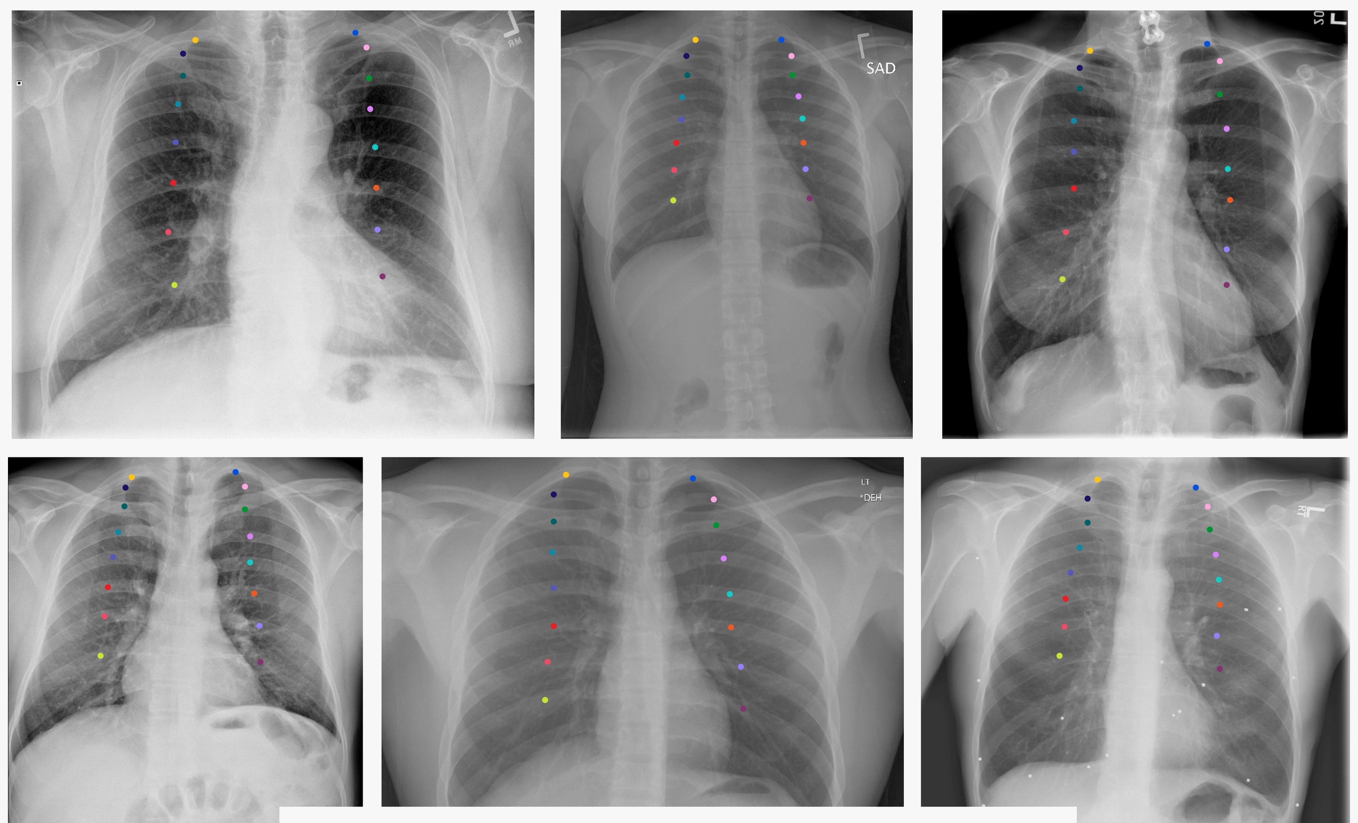
Mader A.O.,[1,2,3] von Berg J.,[3] Fabritz A.,[2] Lorenz C.,[3] and Meyer C.[1,2,3]

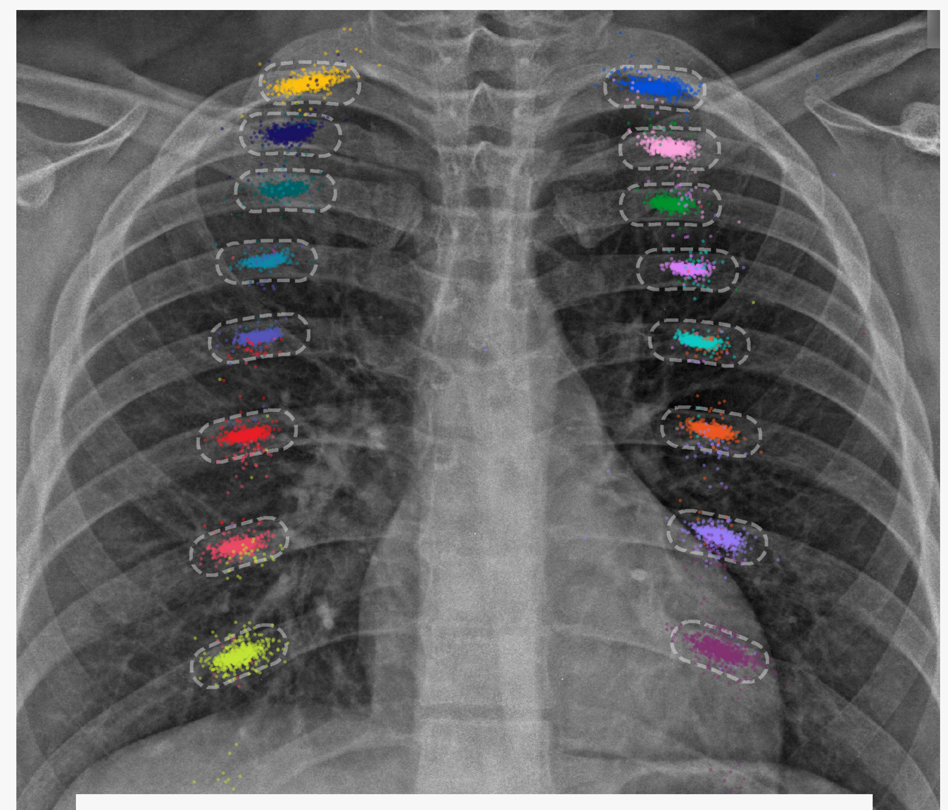
[1] Institute of Computer Science, Kiel University of Applied Sciences, Kiel, Germany
[2] Department of Computer Science, Faculty of Engineering, Kiel University, Kiel, Germany
[3] Department of Digital Imaging, Philips Research Hamburg, Hamburg, Germany

alexander.o.mader@fh-kiel.de

ABSTRACT

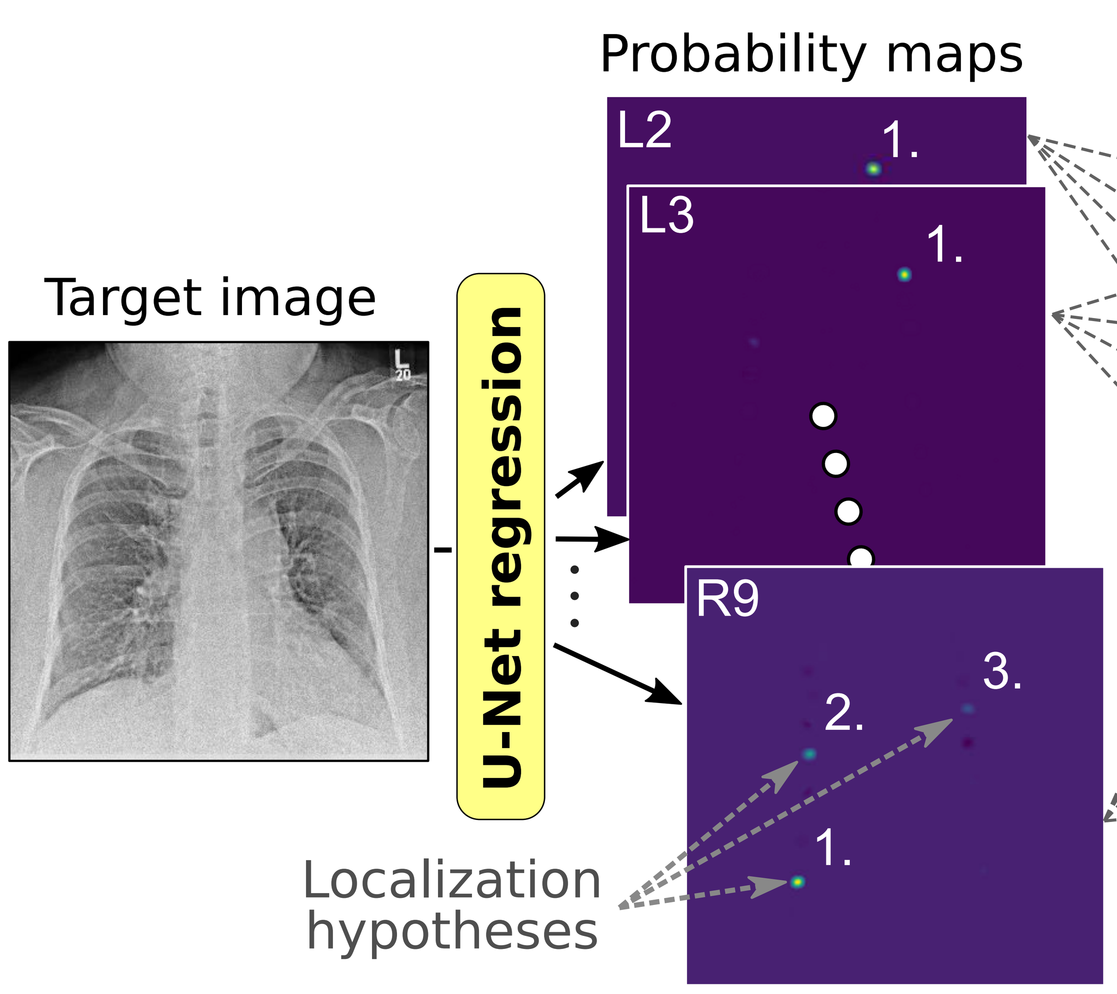
We propose a **general approach for localizing spatially correlated landmarks** using a combination of a **fully convolutional network** (FCN) as hypotheses generator and a **conditional random field** (CRF) as spatial regularizer. To overcome potentially incorrect hypotheses, we introduce a **third refinement step** based on a **novel "refine" label** introduced in the CRF. We use our method to **solve the detection of posterior ribs**, achieving a **success rate of 94.6%** on the public Indiana chest X-ray collection.


83.0% fully correct cases


94.6% correct landmarks

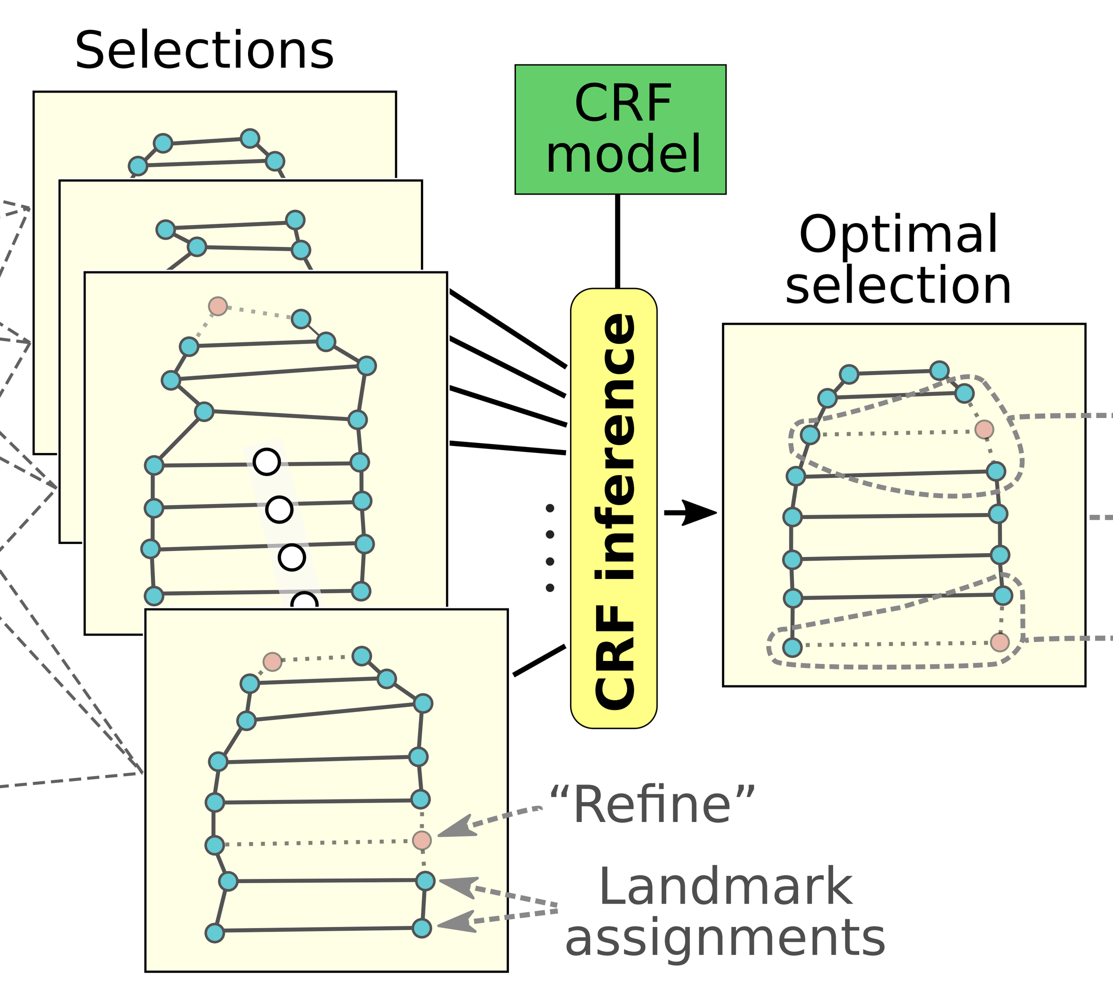
METHOD

1. Use U-Net to generate **localization hypotheses**



- Drop soft-max layer of U-Net to regress heatmaps (instead of classification)
- Sum-squared-error loss with Gaussian at true positions as target
- Use non-maximum suppression to extract 15 localization hypotheses from each regressed heatmap
- **One network to handle all landmarks at the same time (36ms)**
- But poor accuracy on its own with just 30.7% fully correct cases

2. Use CRF to **assign position or "refine"** to each landmark



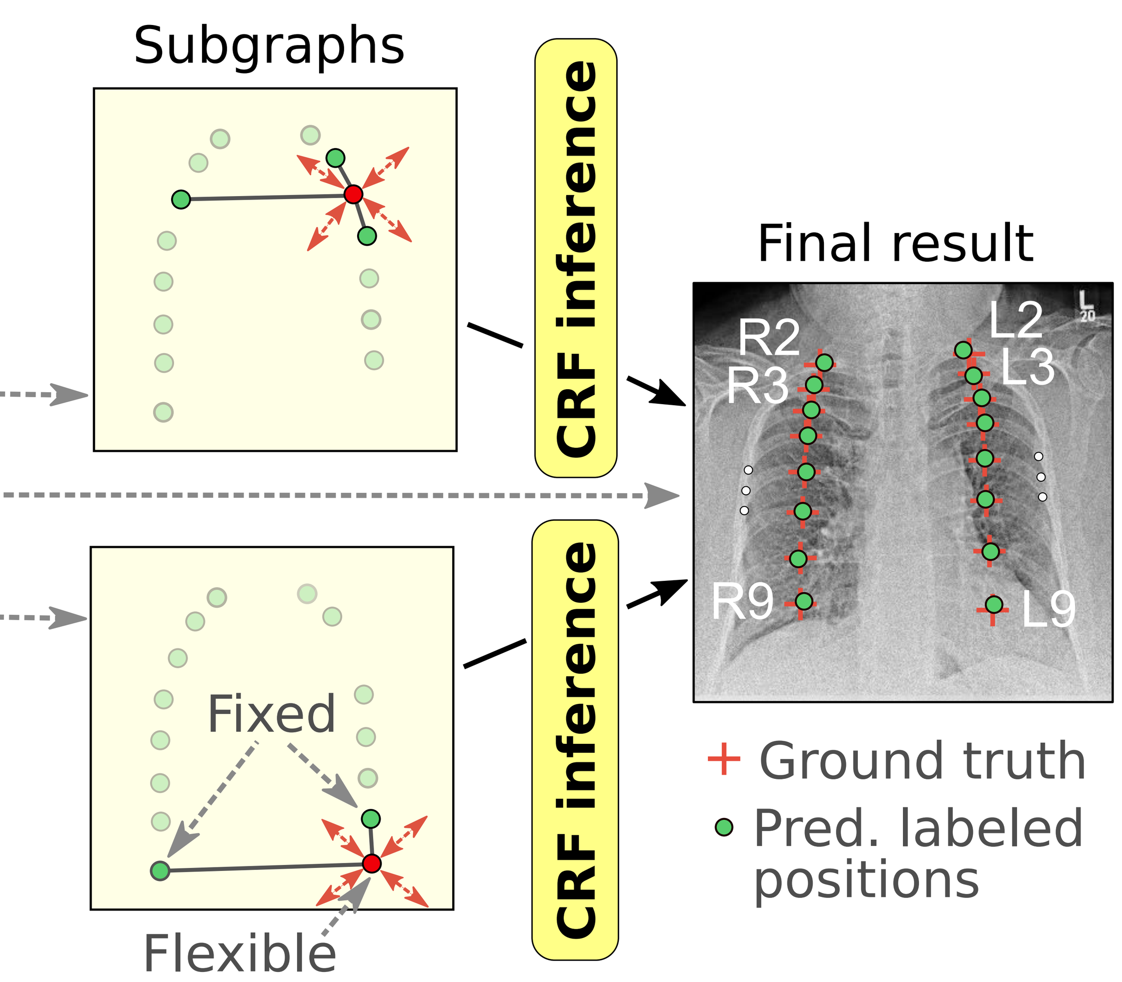
- Landmarks as CRF nodes and localization hypotheses plus novel "refine" label as states
- Unary potentials set to heatmaps
- Binary potentials assess spatial relations: angle, distance, vector
- **Potential weights and "refine" energies are learned from data:**

$$E(\mathbf{S}) = \sum_j \lambda_j \cdot \begin{cases} \beta_j & \exists i \in \mathrm{Scope}(\phi_j)\ s_i = \mathrm{refine} \\ \phi_j(\mathbf{S}) & \end{cases}$$

$$L(\boldsymbol{\lambda}, \boldsymbol{\beta}) = \frac{1}{K} \sum_k \max\left(0, m + E(\mathbf{S}_k^+) - E(\mathbf{S}_k^-)\right)$$

- CRF inference (61ms) increases fully correct cases to 57.3% (+26.6%)

3. Refine selected landmarks over all image pixels



- Extract small subgraphs around "refine"-labeled nodes
- Fix locations for all subgraph nodes except the "refine" node
- Perform CRF inference over all positions instead of 15 hypotheses
- **Sub-problem is feasible $O(n)$ in contrast to full problem $O(n^N)$**
- Greedy heuristic determines order
- Refinement (73ms) increases **fully correct cases to 83.0%** (+25.7%)

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

This work has been financially supported by the Federal Ministry of Education and Research under the grant 03FH013IX5. The liability for the content of this work lies with the authors.