X-ray-transform Invariant Anatomical Landmark Detection for Pelvic Trauma Surgery

Bastian Bier*[,1,2], Mathias Unberath*[,2], Jan-Nico Zaech[1,2], Javad Fotouhi[2], Mehran Armand[3], Greg Osgood[4], Nassir Navab*[,2], Andreas Maier*[,1]

[1] Pattern Recognition Lab, Friedrich-Alexander-Universität Erlangen-Nürnberg

[2] Computer Aided Medical Procedures, Johns Hopkins University

[3] Applied Physics Laboratory, Johns Hopkins University

[4] Department of Orthopaedic Surgery, Johns Hopkins Hospital

* Joint first authors listed alphabetically.

Percutaneous procedures can be difficult

- **Percutaneous surgery**: No direct view onto anatomy
- **Indirect feedback**: X-ray images

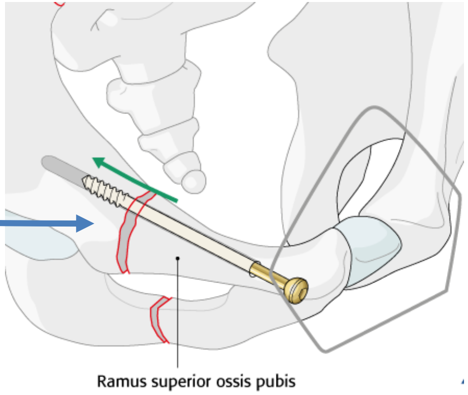
→ **Projective transformation**: Multiple views required

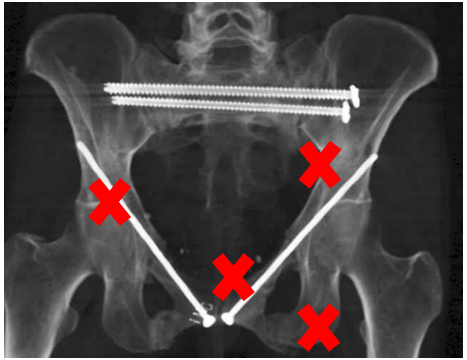
→ **Mental mapping**: Tool to patient from images

Percutaneous pelvis fixation

- **Difficult anatomy**: Complex fractures
- **Pre-op. 3D available**: Not used intra-op.
- **Guidance**: Mobile, non-robotic C-arms
- **Per screw**: > 1h and > 100 X-rays

→ **Very tedious procedures**





Anatomical landmarks: A powerful concept

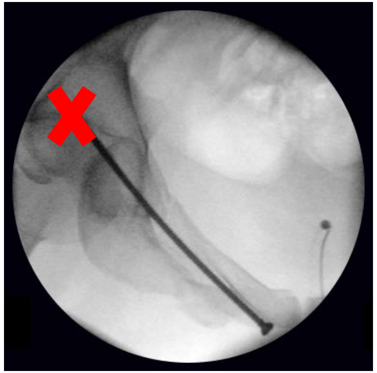
They …

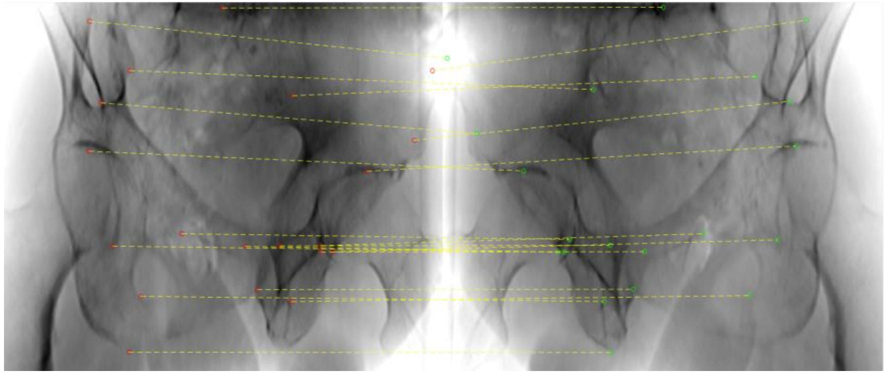
… **provide context** → Supports intra-operative decision making!

… **supply semantic information** → Defines correspondences (camera pose)!

… **foster machine understanding** → Groundbreaking steps towards autonomy!

…

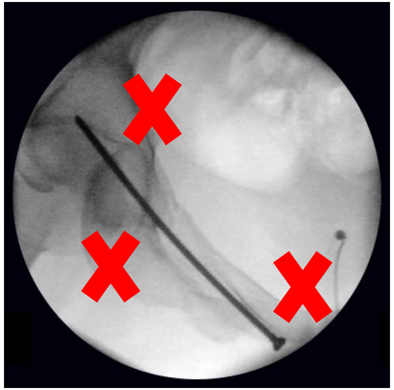




Detect anatomical landmarks in X-ray images *independent from* the viewing direction

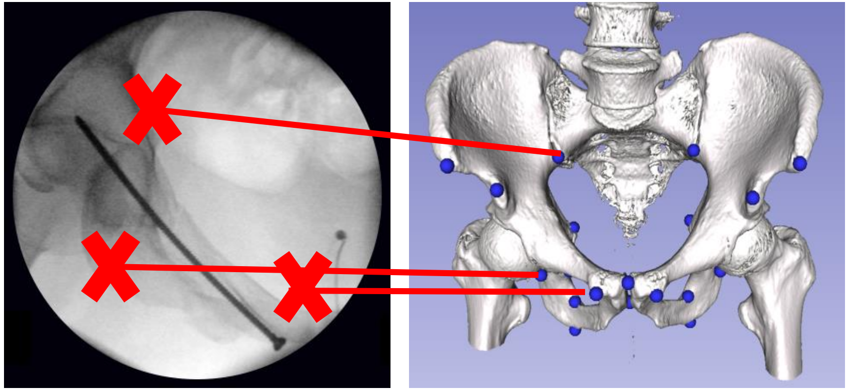
Step 1

Detection



Step 2

Semantic knowledge enables initialization of 2D/3D registration

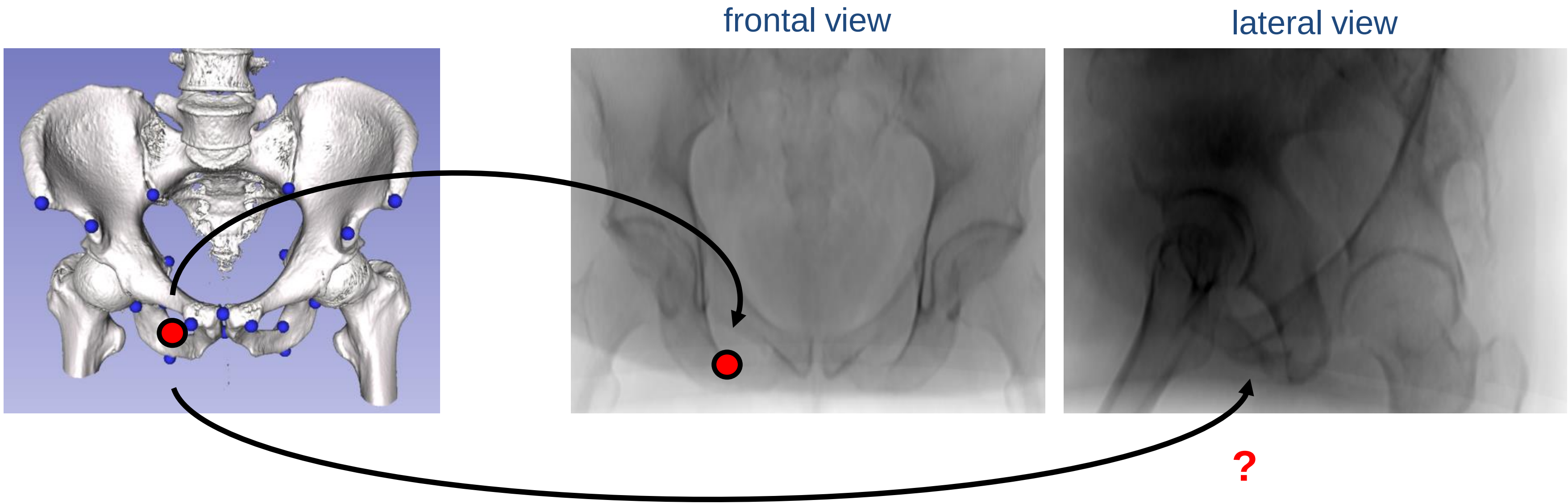


Landmark appearance can change substantially

Not comparable with reflection imaging

→ **Structures overlap** and **edges vanish** in transmission imaging

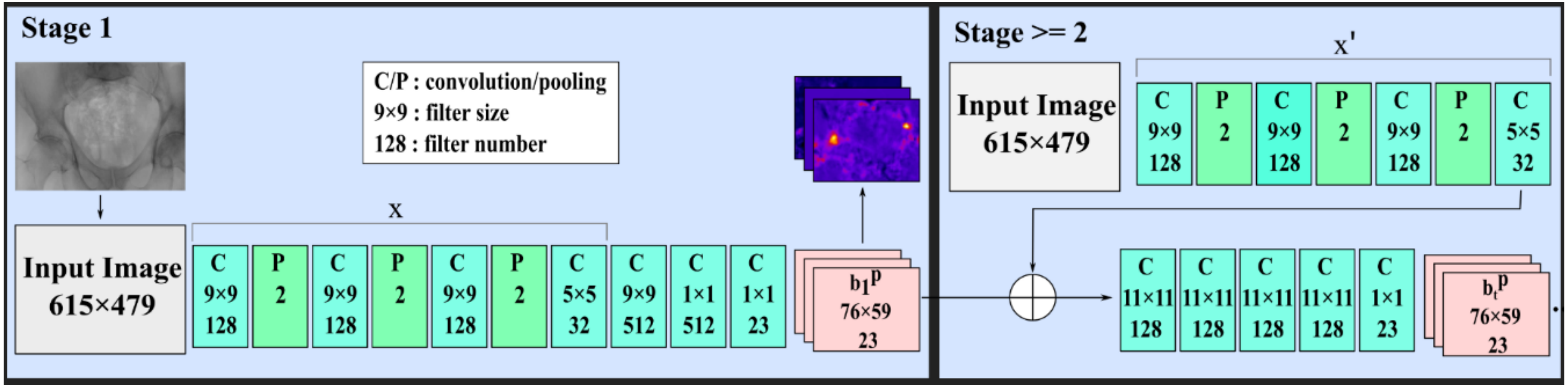
→ Common techniques (e.g. local descriptors) not generally appropriate



Architecture: Sequential predication framework

Input: X-ray image

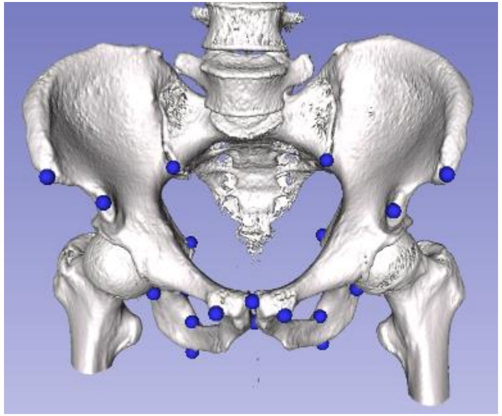
Output: One heatmap per landmark (multi-task learning)



- **Large receptive field** → Reveal global configurations
- **Stage wise application** → Resolve ambiguities

How to train the network?

We need the 2D locations of all 23 landmarks in each projection
→ **Infeasible to manually label in 2D** (time, accuracy, consistency)



CT volume with 23 labels

Projection of the volume

Physics-based simulation via **DeepDRR**
(Poster ID **W-6**; Available on *Github*)

Projection of the 3D labels

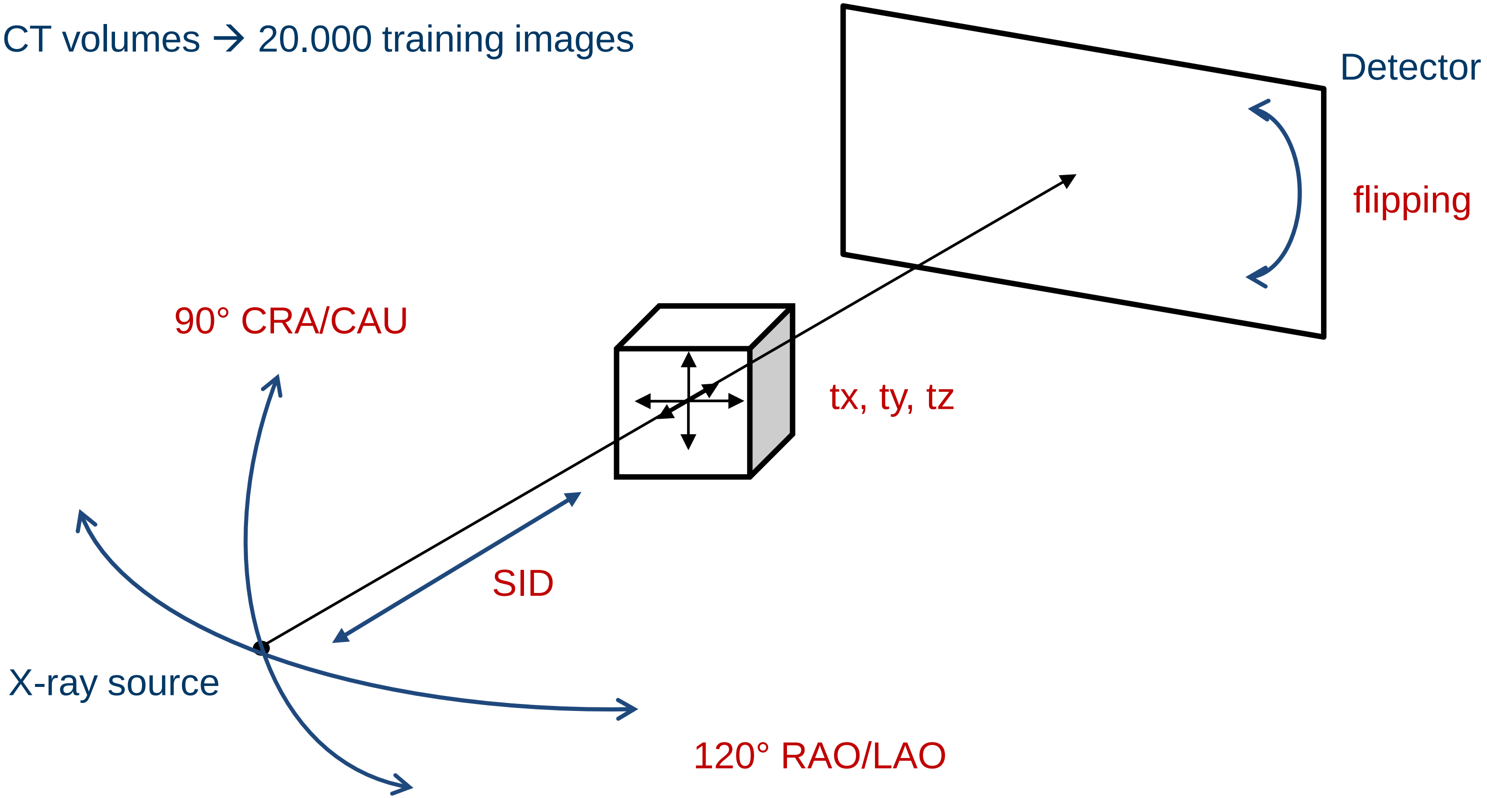
Training data

X-ray

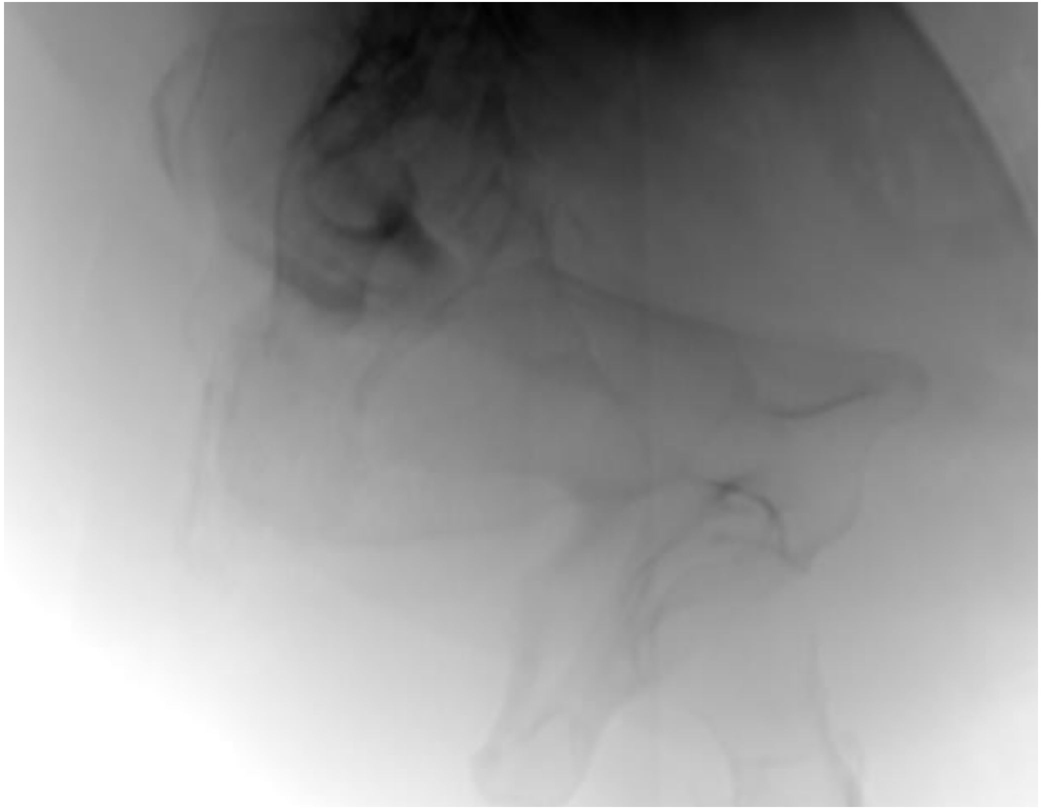
Heatmaps

View invariance needs augmentation

20 CT volumes → 20.000 training images



Results



Average error:

9.1 ± 7.4 pixels

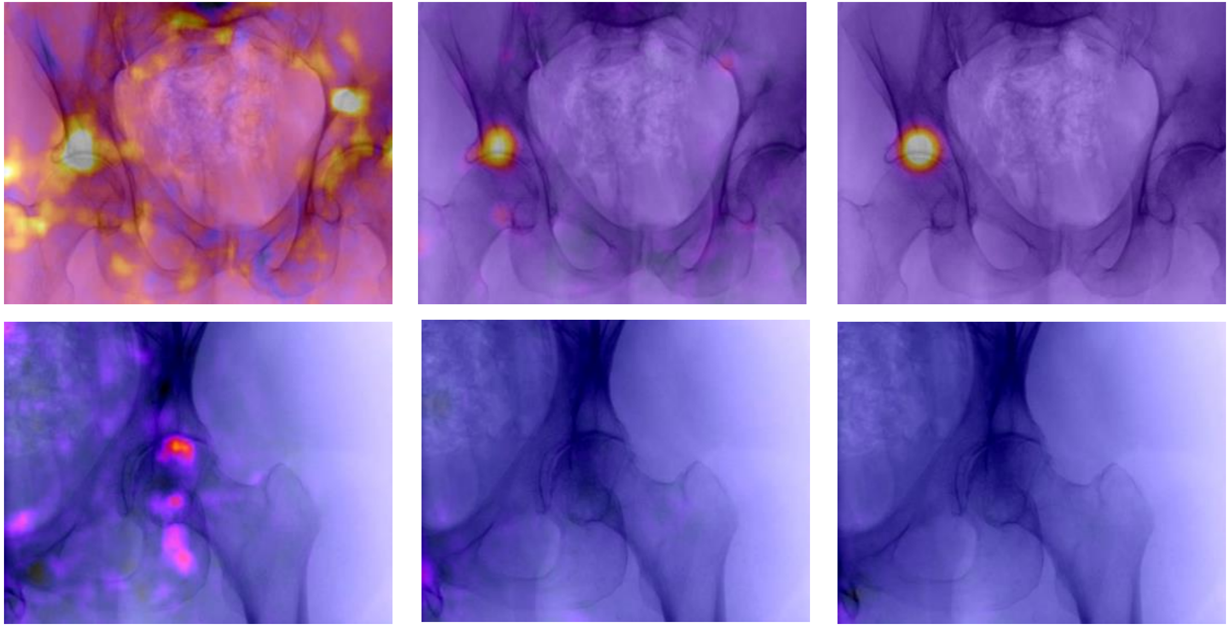
5.6 ± 4.5 mm

Resolving ambiguities

Task: Detect the tip of the right femur

Example 1
Landmark in FOV

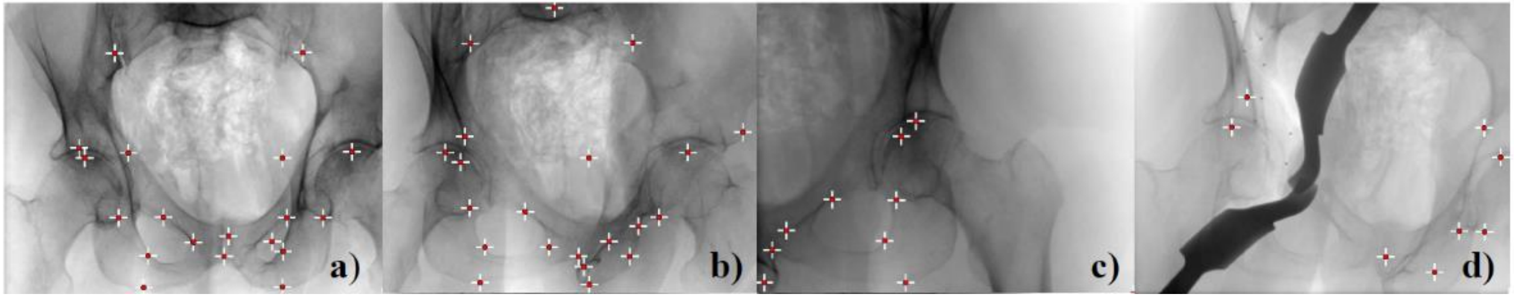
Example 2
Landmark **not** in FOV



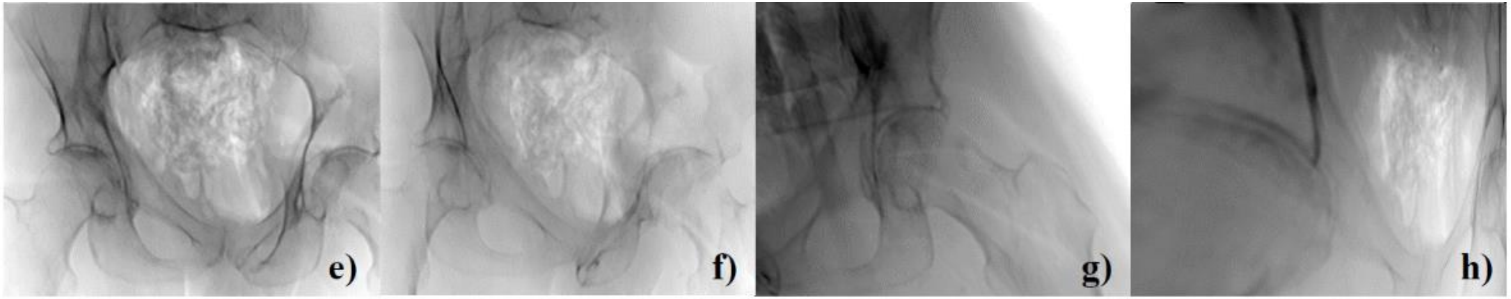
Stage 1 Stage 3 Stage 6

Network transfers to real data

Landmark predictions

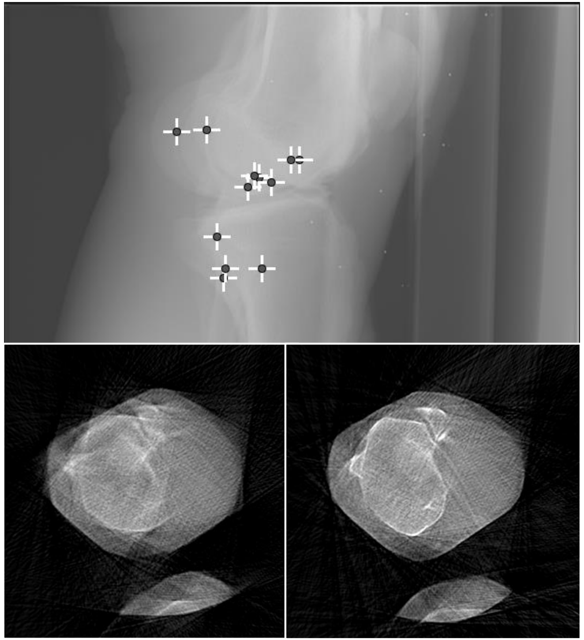


Estimated 2D/3D pose: Analytic solution via direct linear transform → Initialize registration (*global!*)



Discussion and Conclusion

“The first tool for anatomical landmark detection in X-ray images independent of the view.”



Improving accuracy

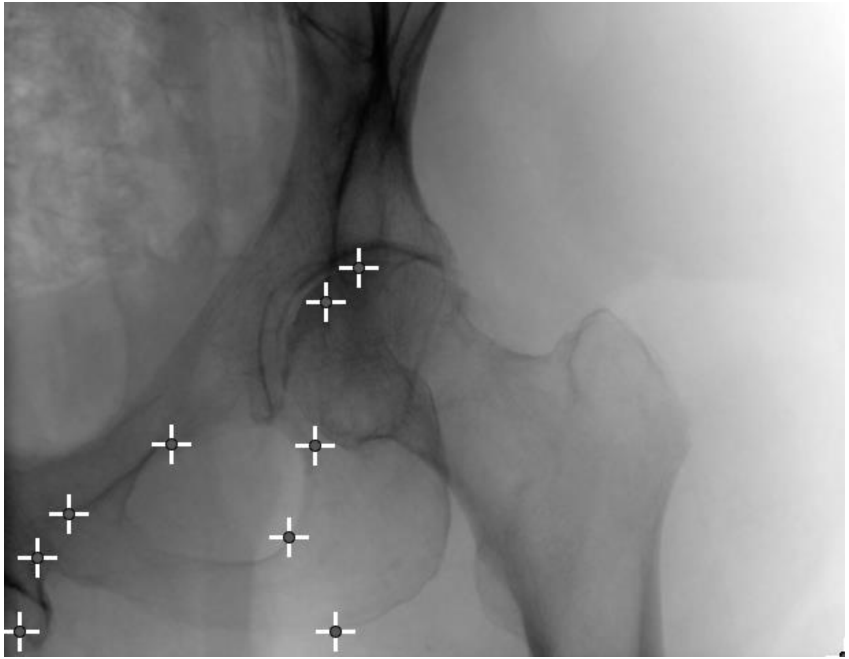
Towards robust performance in presence of tools

Push deployment for 2D/3D registration

Strong prospect for other anatomies and applications

Thank you!

Find us at poster W-1



Clinical collaborator
Prof. Greg Osgood M.D