Deep Learning for Fluoroscopic Feature Detection

Mid Point Presentation

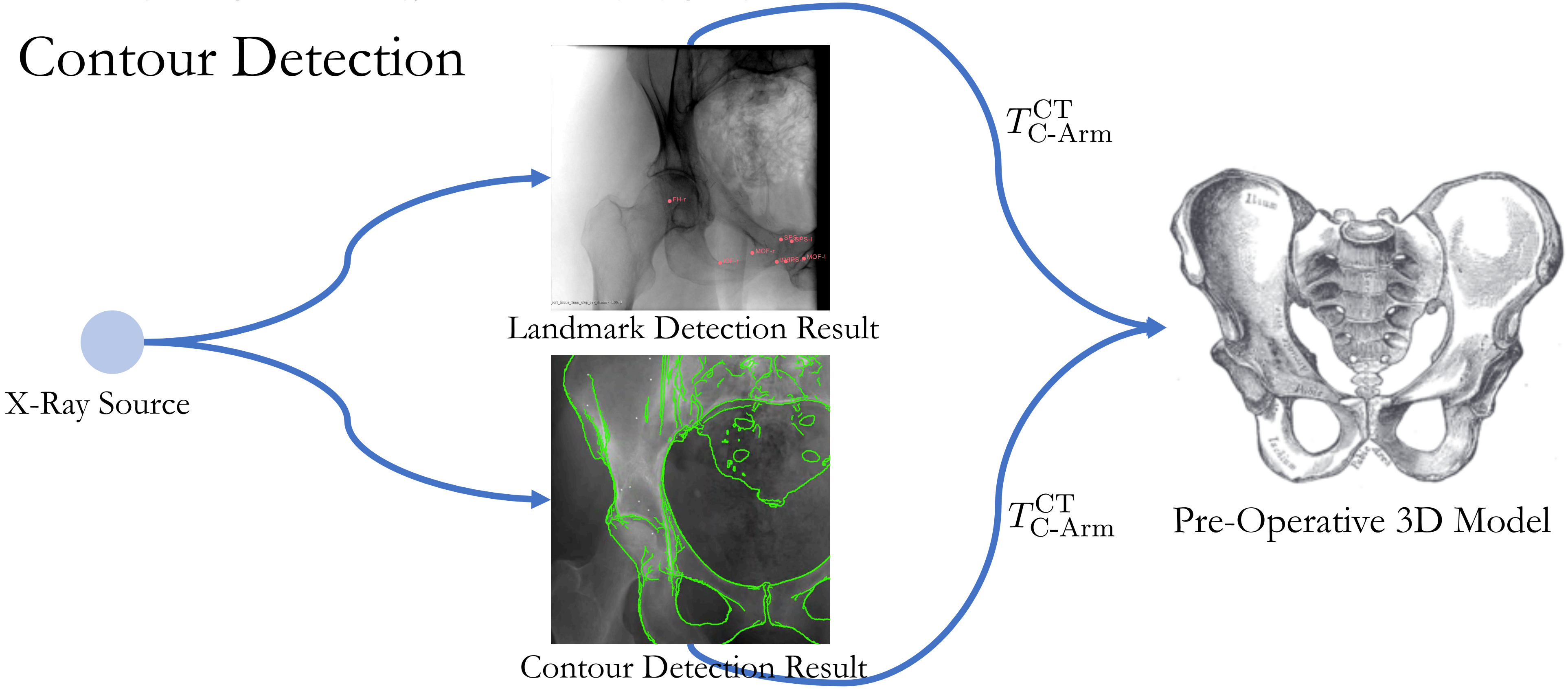
Liujiang Yan

Mentor: Robert Grupp, Professor Russell Taylor

JOHNS HOPKINS WHITING SCHOOL of ENGINEERING

Motivation

- Feature Detection for 2D-3D Registration
  - Anatomical Landmark Detection
  - Contour Detection



[1] Wikipedia contributors. (2018, March 17). Pelvis. In Wikipedia, The Free Encyclopedia. Retrieved 17:09, April 9, 2018.

JOHNS HOPKINS
WHITING SCHOOL of ENGINEERING

Deliverable Status

- Minimum Deliverable Achieved
- Medium / Maximum Deliverable Ongoing

Task	Status	02/22	03/01	03/15	03/29	04/05	04/12	04/19	04/26	05/03	05/10
Project Proposal	Done									**Done**	
Environment Set Up	Done									**Ongoing**	
Initialize Neural Network	Done									**Not Start**	
Seminar Report	Ongoing										
Training on Larger data	Done										
Evaluating Accuracy	Done										
Final Report / Poster	Not Yet										
Tools in field of view	Ongoing										
Simulation Software	Not Yet										
Contour Detection	Not Yet										

JOHNS HOPKINS
WHITING SCHOOL of ENGINEERING

Pipeline Summary

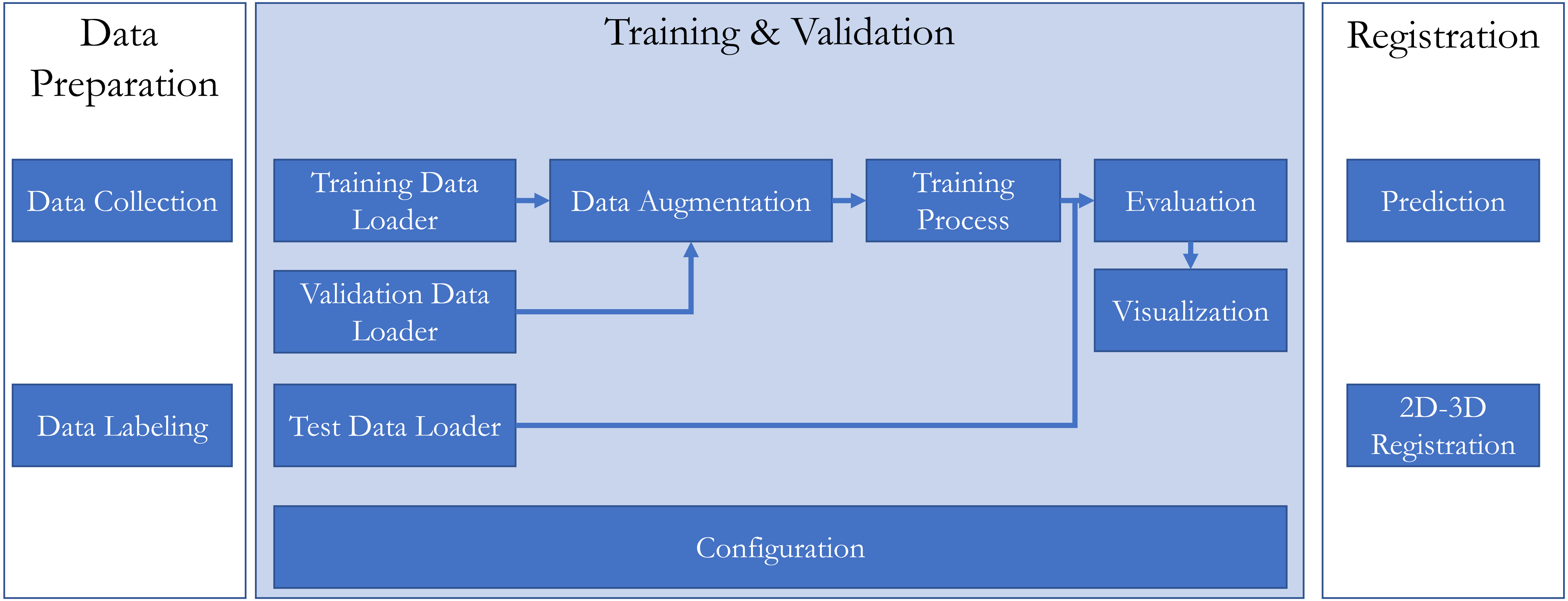
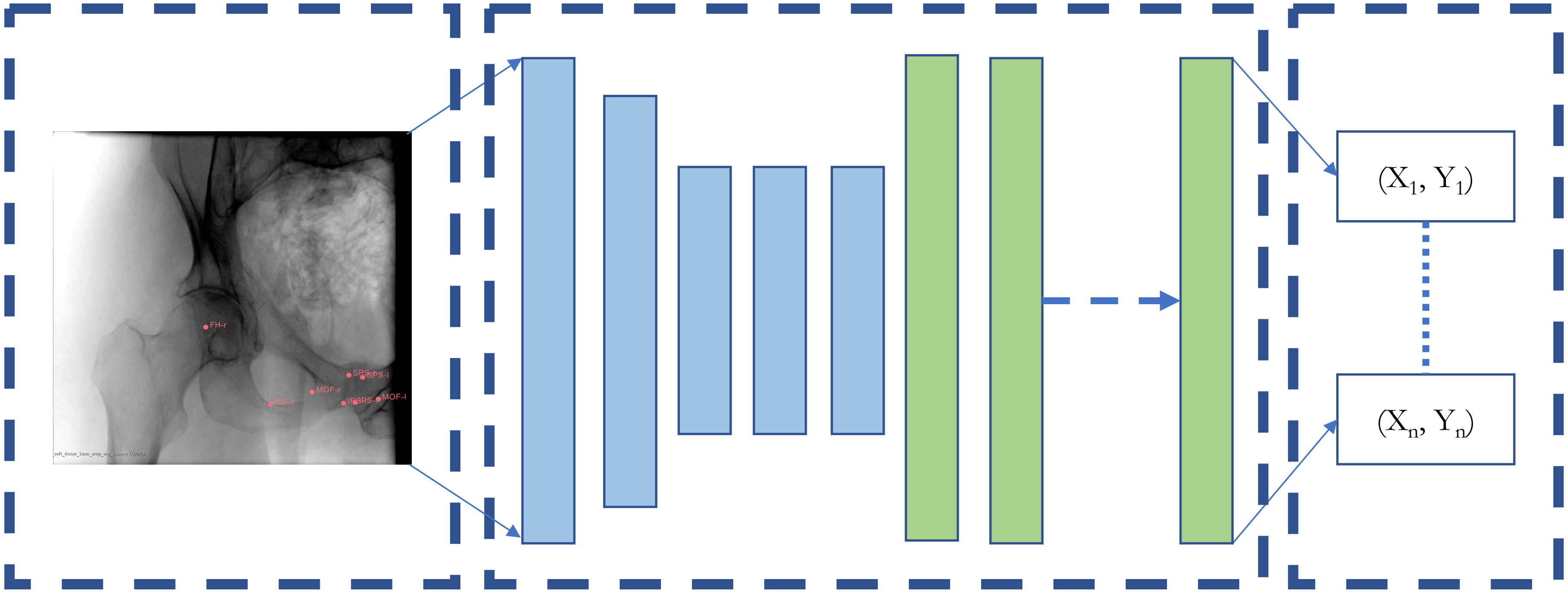


Fig. Pipeline Layout

Approach - Position Regression

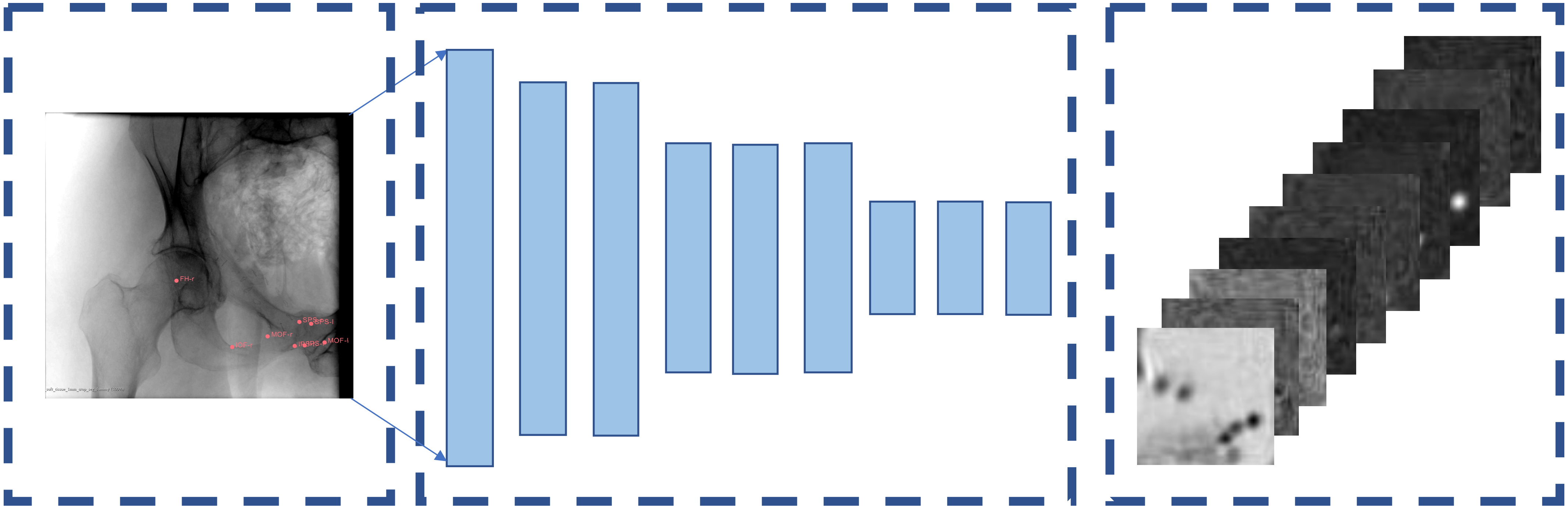
- Problem identification: fixed number key point detection



JOHNS HOPKINS
WHITING SCHOOL *of* ENGINEERING

Approach - Heatmap Regression

- Problem identification: fixed number key point detection
- Heatmap: Intensity given by <u>Gaussian distribution with position as mean</u>



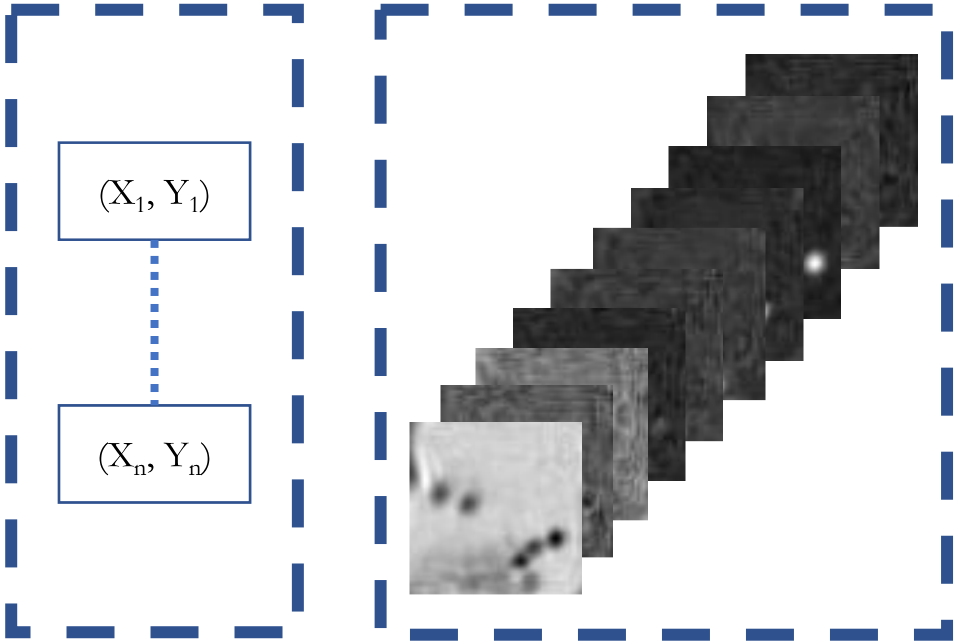
Input

Architecture
CNN

Heatmap Regression

Approach

Why regress heatmap instead of position? What is the tradeoff?



Position Regression

Heatmap Regression

Network Architecture

- Convolutional Pose Machine
- Multi Stage Model
- Intermediate Supervision

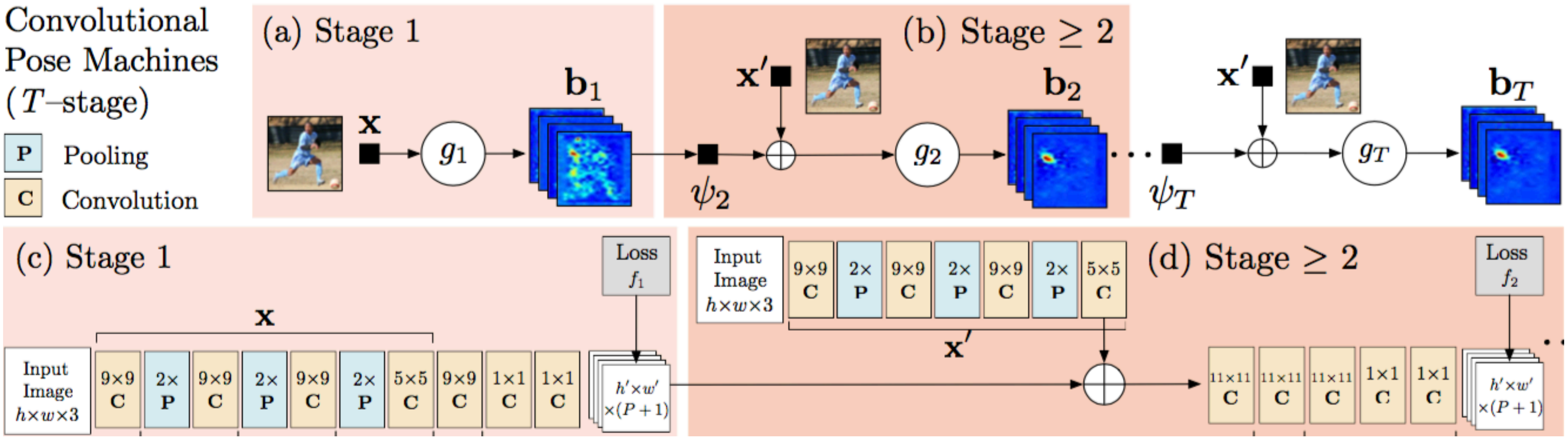


Fig. Convolutional Pose Machine[1]

[1] Wei, Shih-En, et al. "Convolutional pose machines." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016.

Network Architecture

- Convolutional Pose Machine
- Multi Stage Model
- Intermediate Supervision

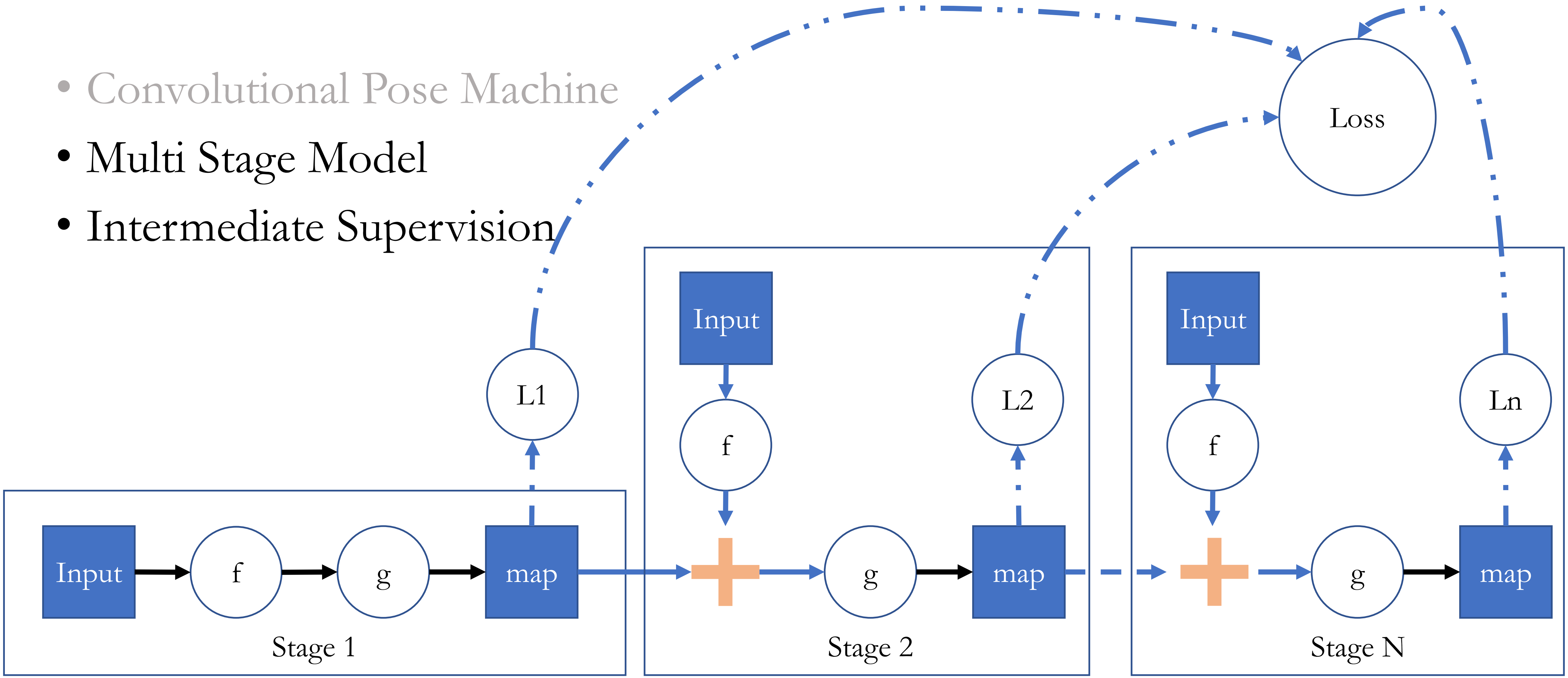


Fig. Architecture Details

Hyperparameters - Stage Number

- More stages:
  - Larger parameter space; larger receptive field;
  - Overfitting; Harder to train; Takes more time and memory;

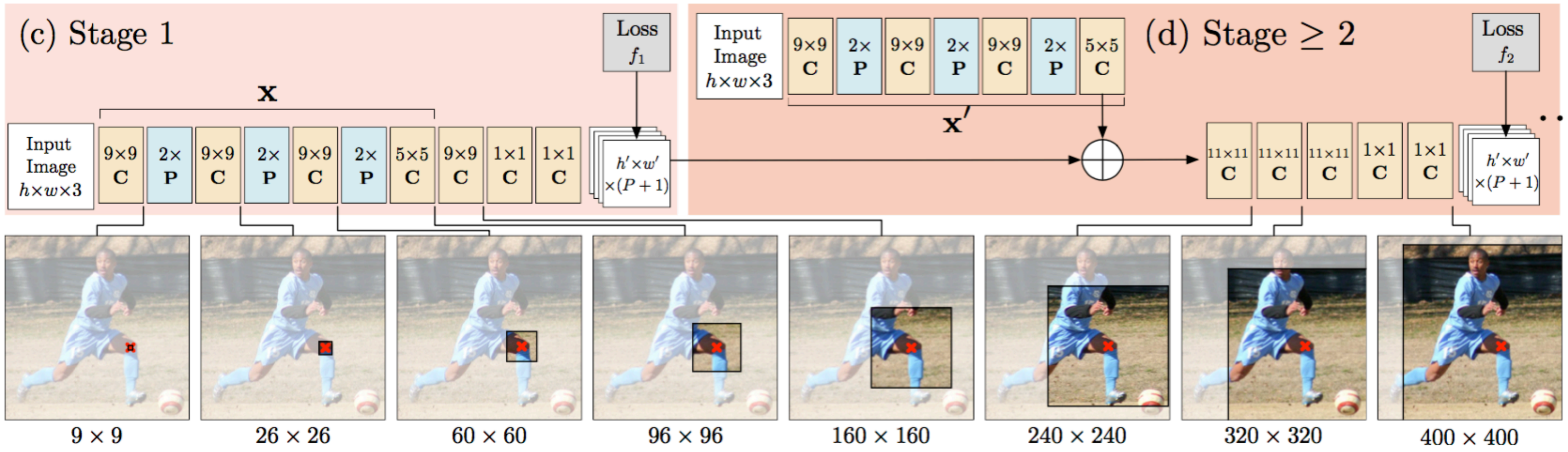


Fig. Convolutional Pose Machine[1]

[1] Wei, Shih-En, et al. "Convolutional pose machines." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016.

Johns Hopkins Whiting School of Engineering

Hyperparameters - Heatmap

- Output Size:
  - Larger size means more accurate localization; More resources to train;
- Gaussian Variance:
  - Larger variance means more information; less accurate localization;

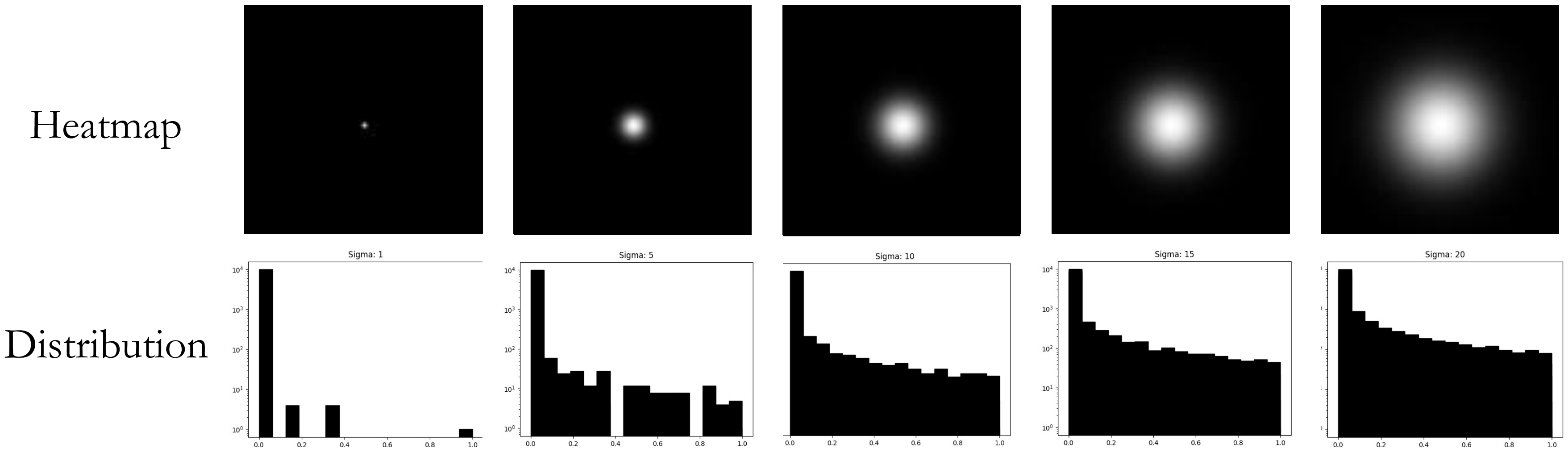


Fig. Heatmap with different variance and corresponding intensity histogram

JOHNS HOPKINS
WHITING SCHOOL of ENGINEERING

Evaluation and Visualization

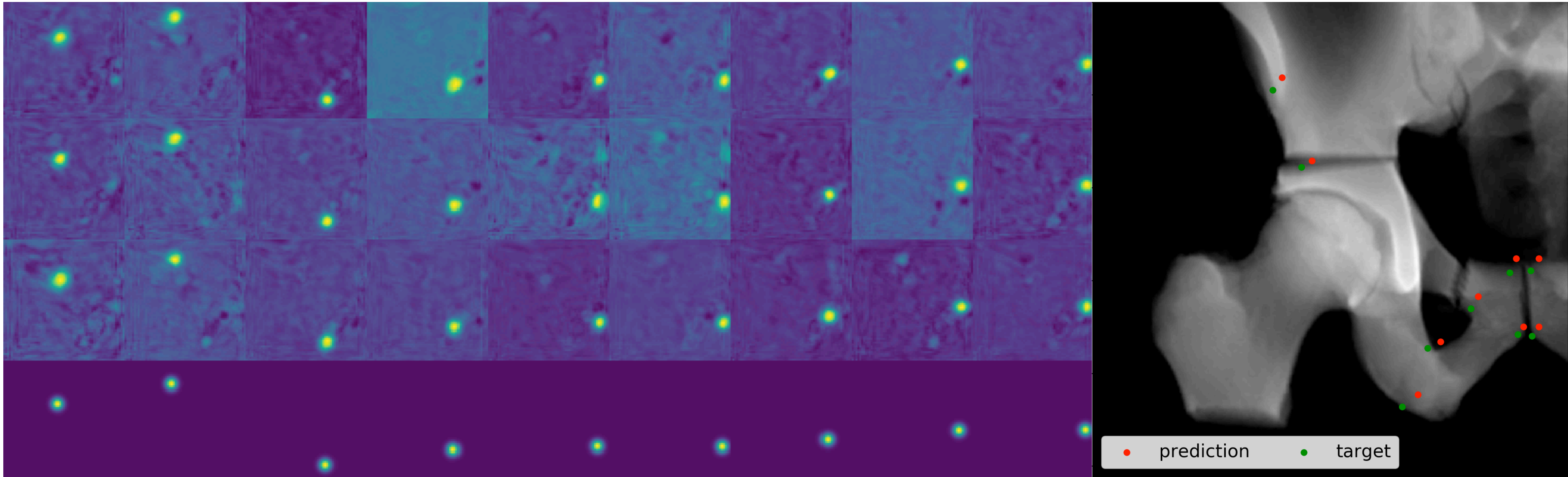
- Multistage Visualization - Intuition
- 2D Error Plot - Quantitative Analysis
- Good Prediction / Bad Prediction - Case Study

Stage1

Stage2

Stage3

Target



Evaluation and Visualization

- Multistage Visualization - Intuition
- 2D Pixel Error Plot - Quantitative Analysis
  - Discretization: 10x downsample
    - Measure in original image size (768 by 768);
    - Heatmap size (63 by 63);
- Good Prediction / Bad Prediction - Case Study

Evaluation and Visualization

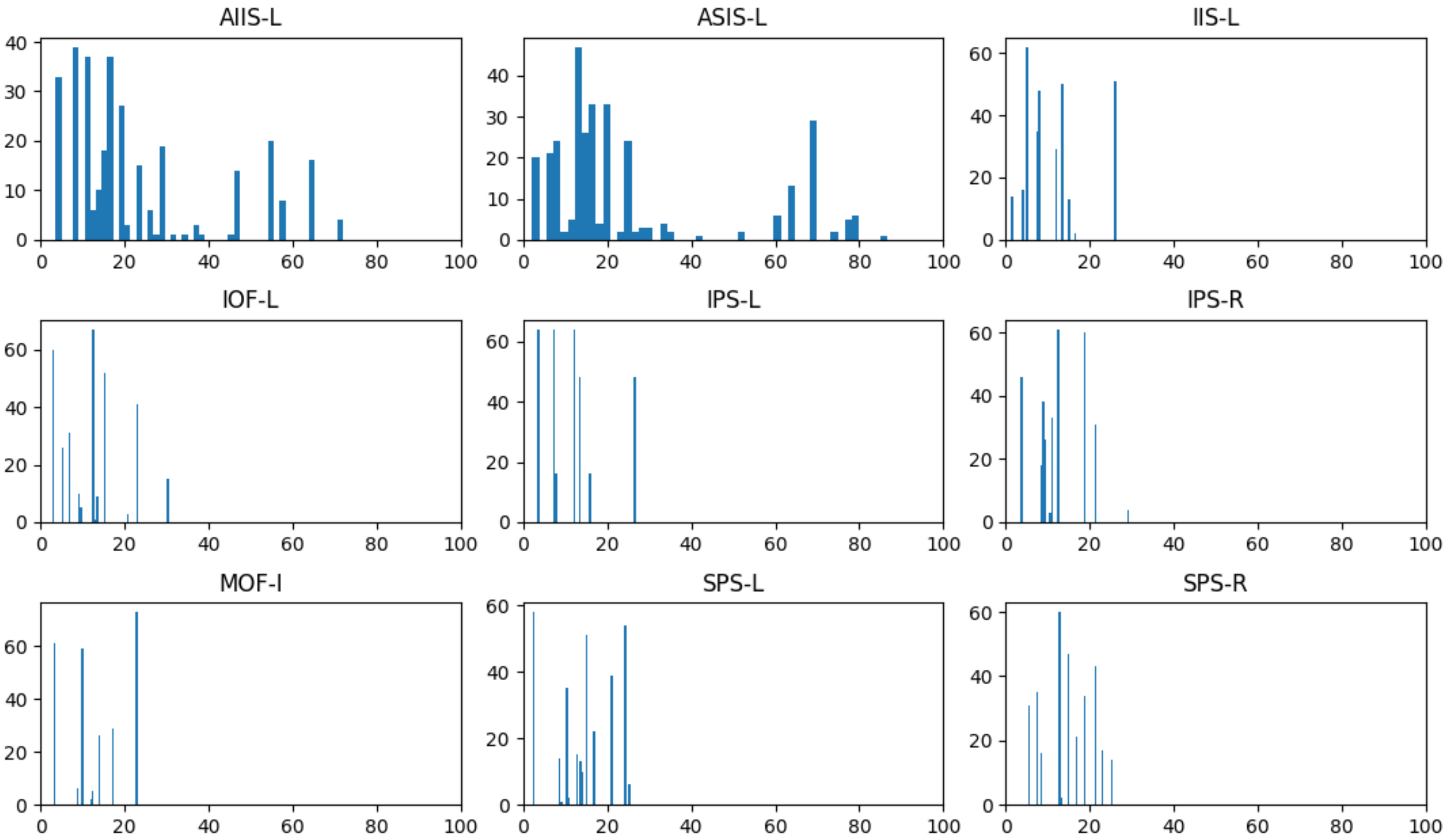


Fig. 2D Pixel Error

JOHNS HOPKINS
WHITING SCHOOL of ENGINEERING

Evaluation and Visualization

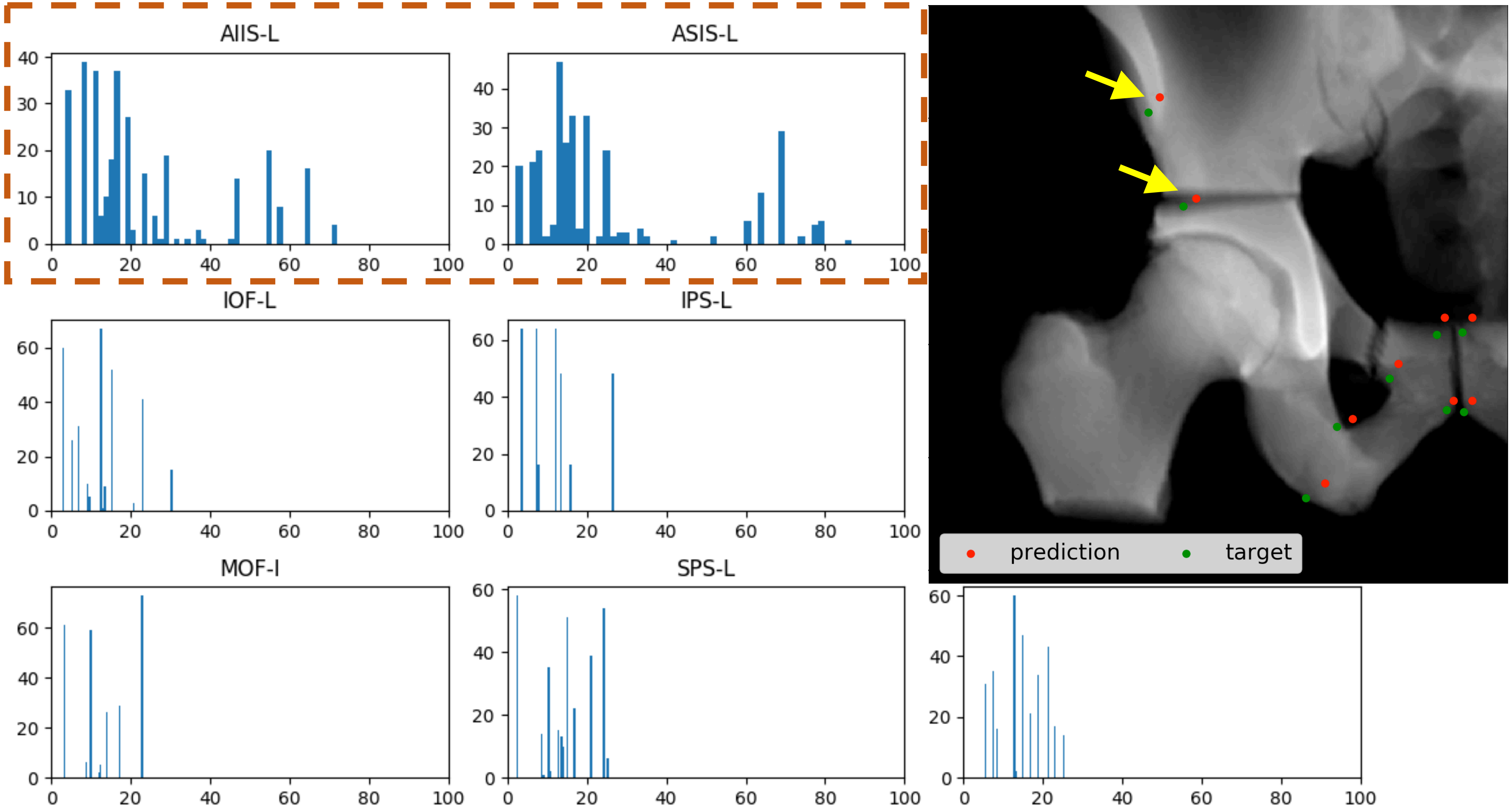
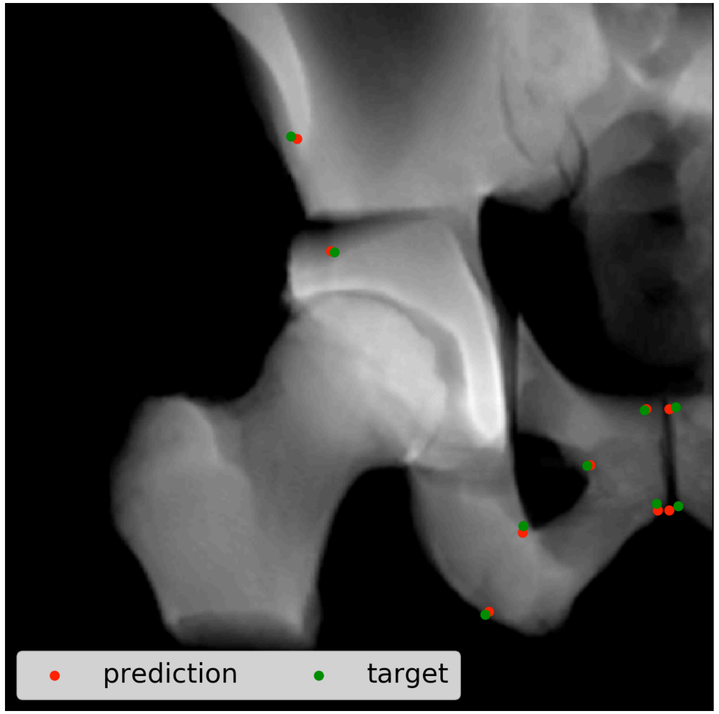
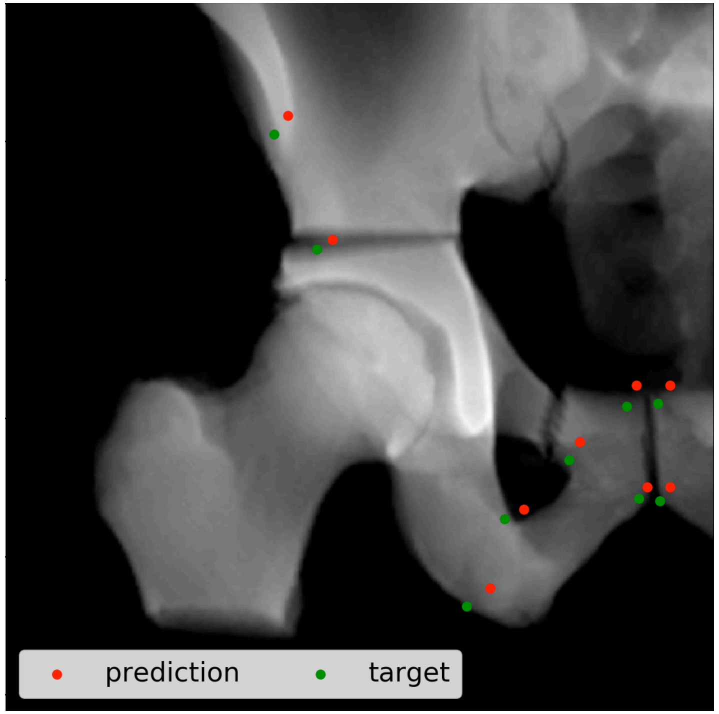


Fig. 2D Pixel Error

Evaluation and Visualization

- Multistage Visualization - Intuition
- Spatial Error Plot - Quantitative Analysis
- Good Prediction / Bad Prediction - Case Study





Limitations & Data Augmentation

- Intra-operation scenario: Tool in the view
- Hard to synthesize samples with tool in the view

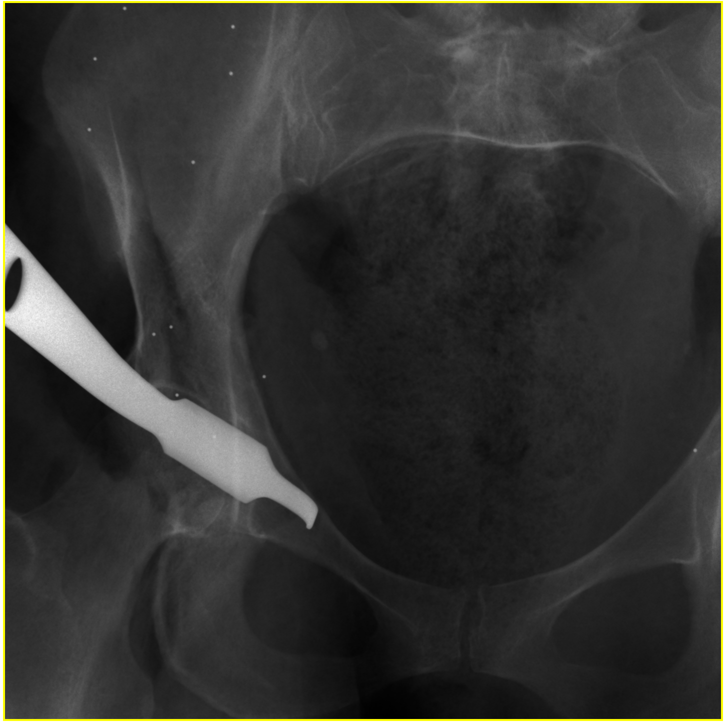


Fig. Tool in the view

Limitations & Data Augmentation

- Intra-operation scenario: Tool in the view
- Random Mask (shape; size; location; rotation; brightness)
  - Input: mask some connected region with random constant value;
  - Target: keep all landmarks visible;

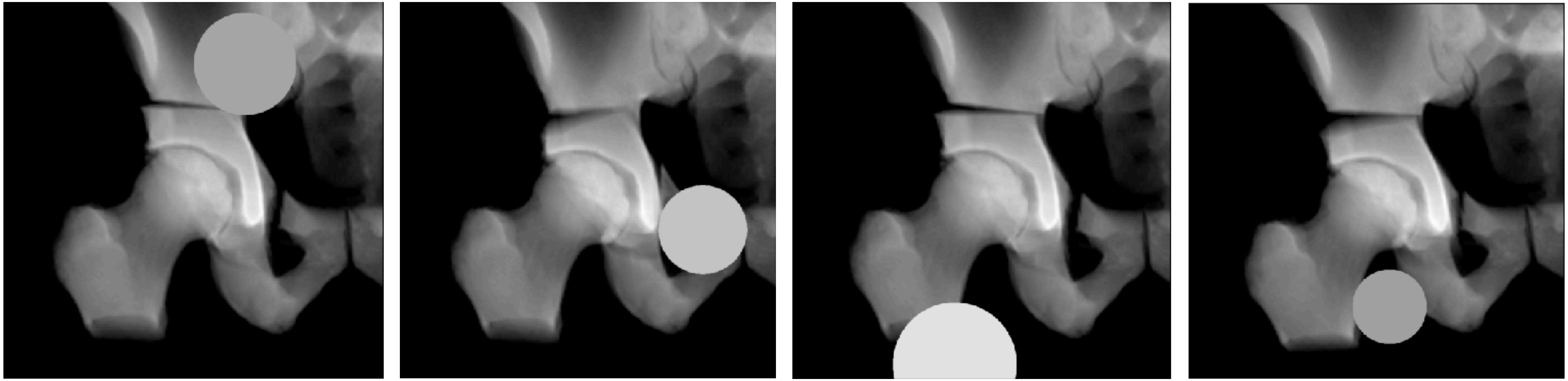
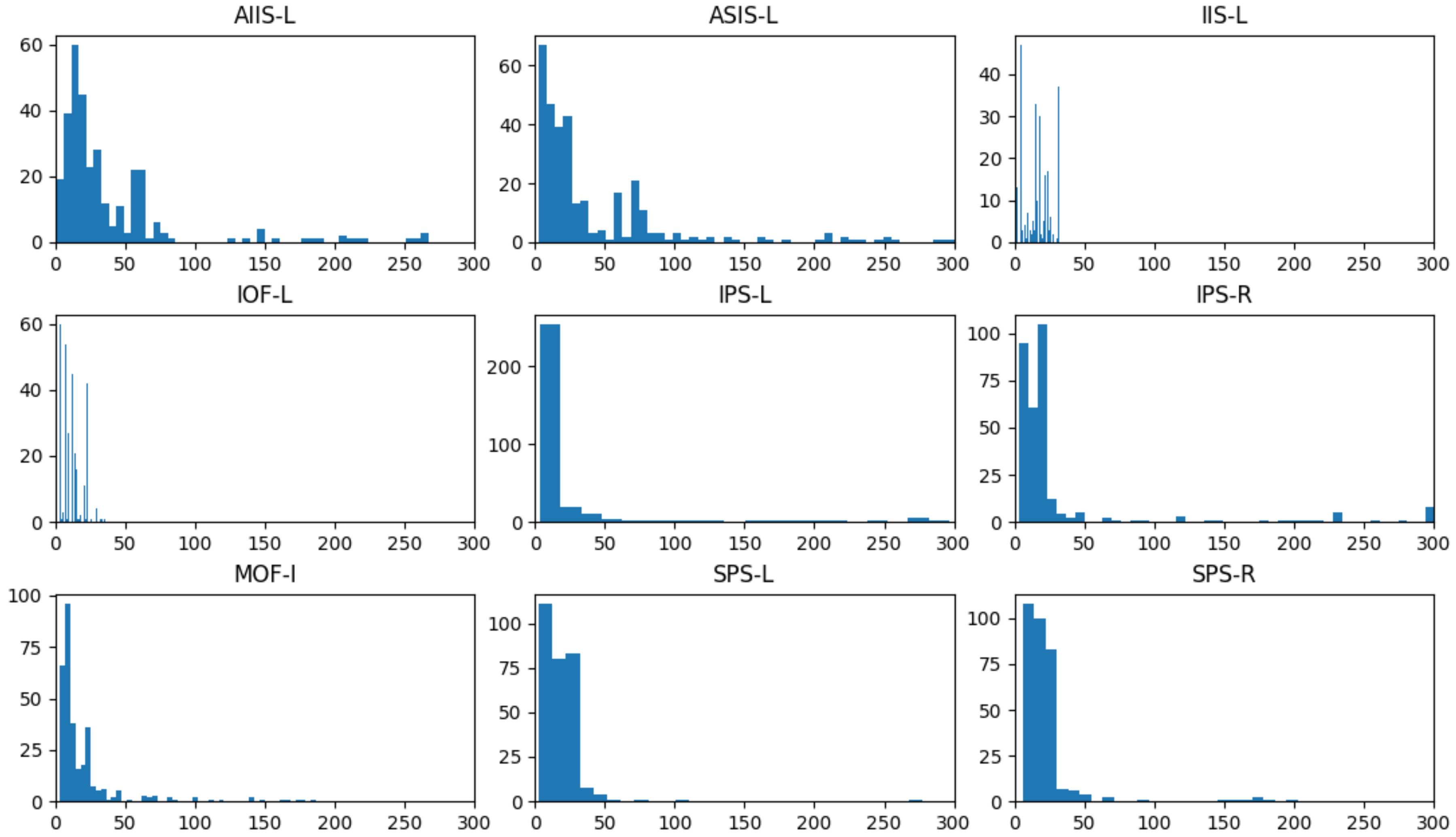


Fig. Samples with Random Mask (not necessarily circular)

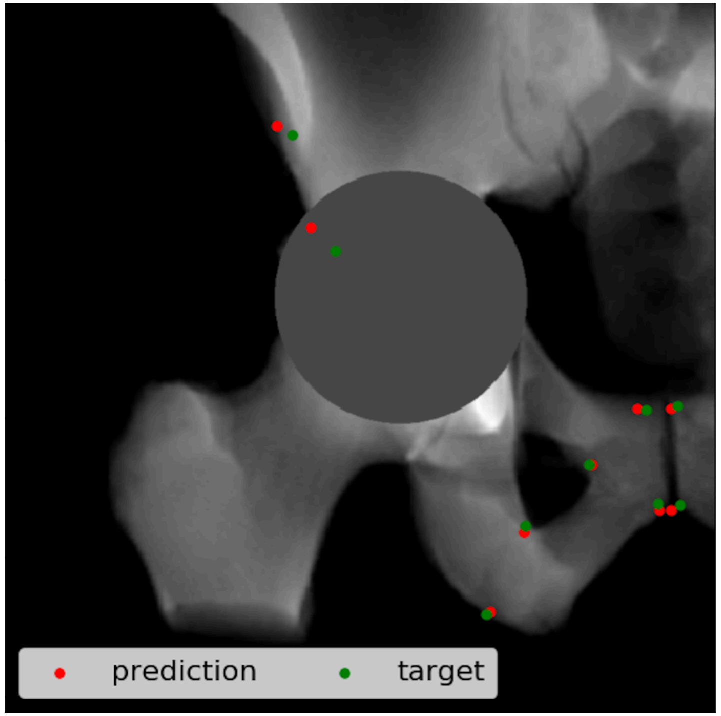
Spatial Error Plot

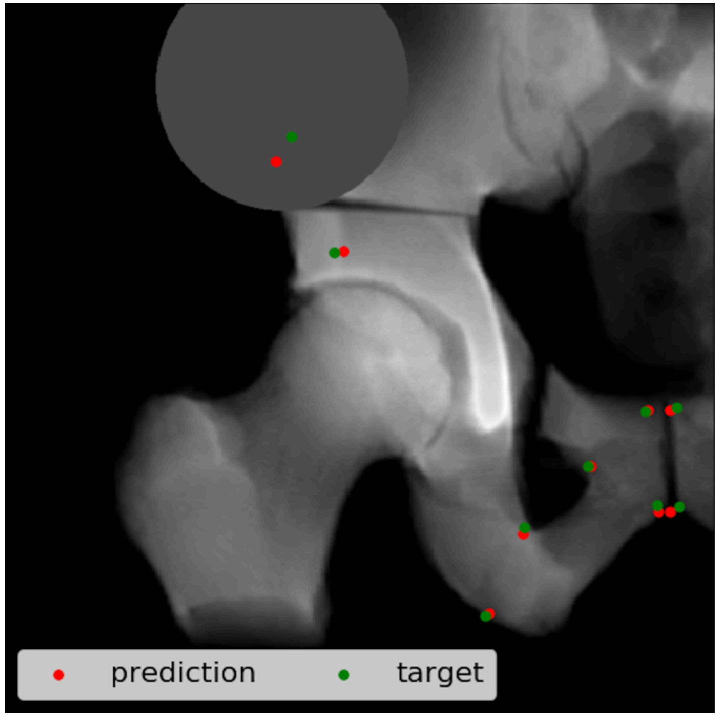


Landmark Spatial Error (**With Mask**)

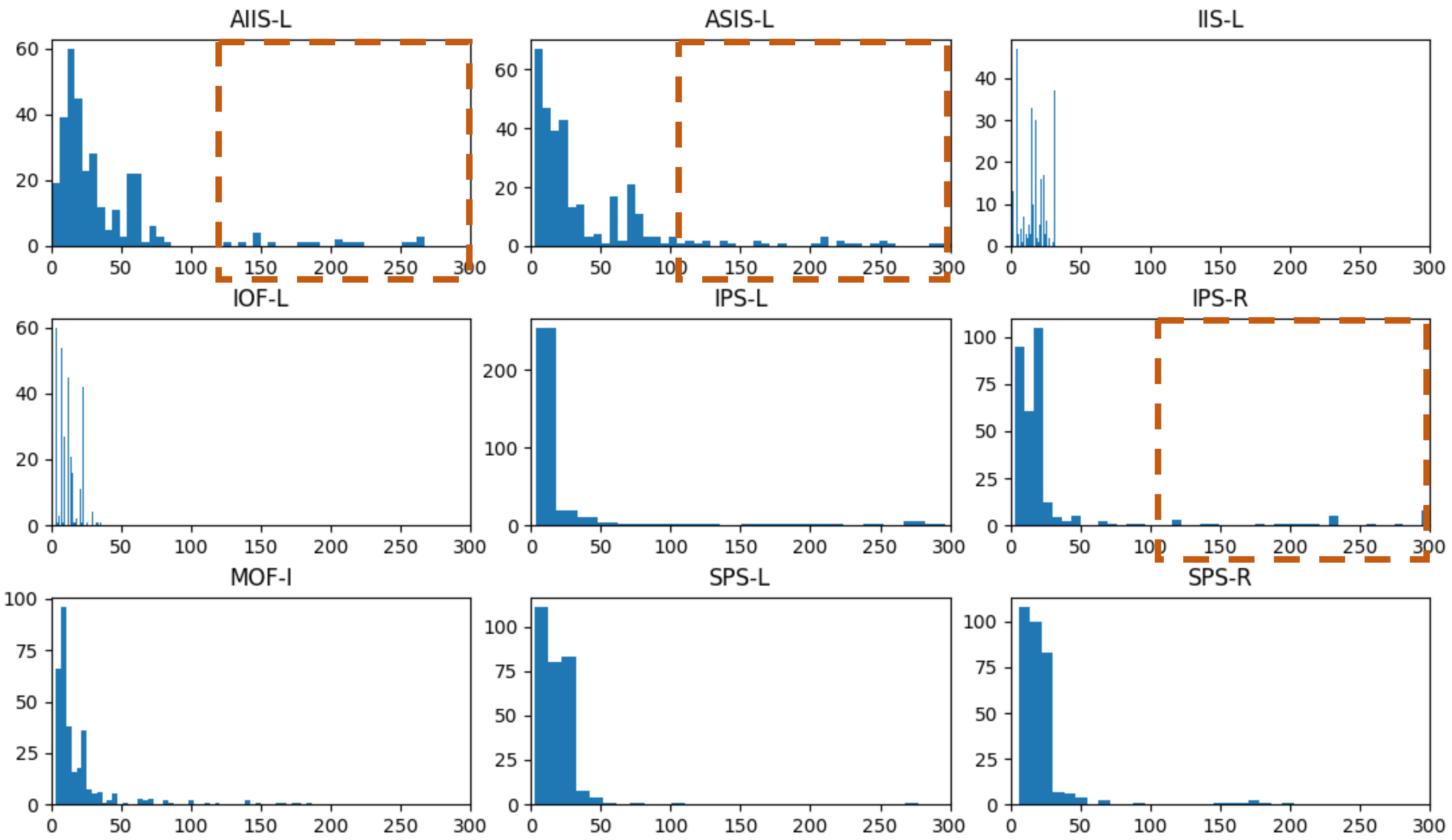
Evaluation and Visualization

- Multistage Visualization - Intuition
- Spatial Error Plot - Quantitative Analysis
- Good Prediction / Bad Prediction - Case Study





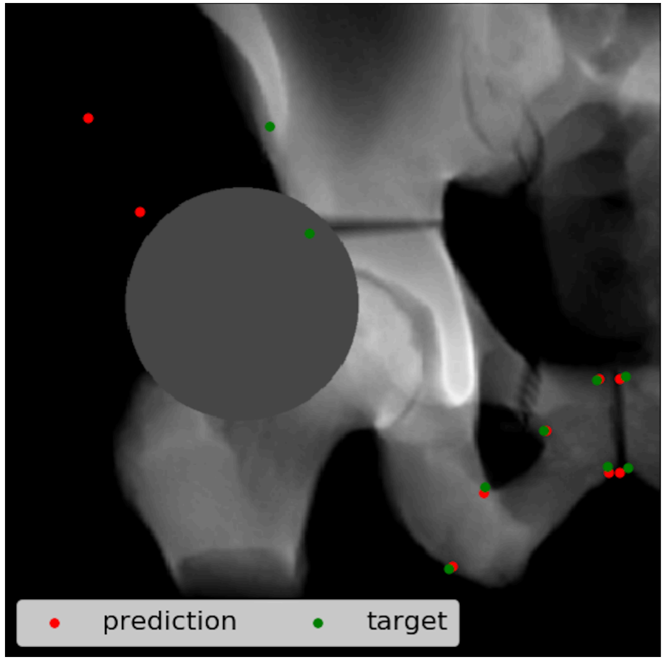
Spatial Error Plot



Landmark Spatial Error (**With Mask**)

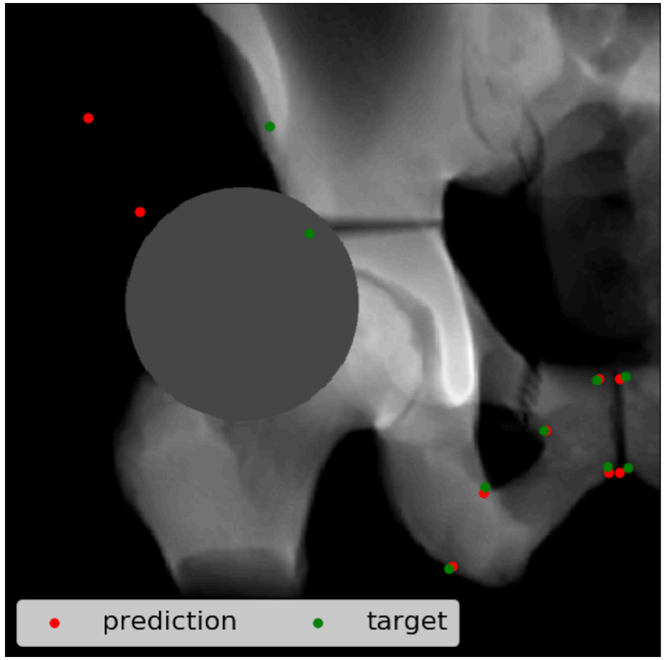
Evaluation and Visualization

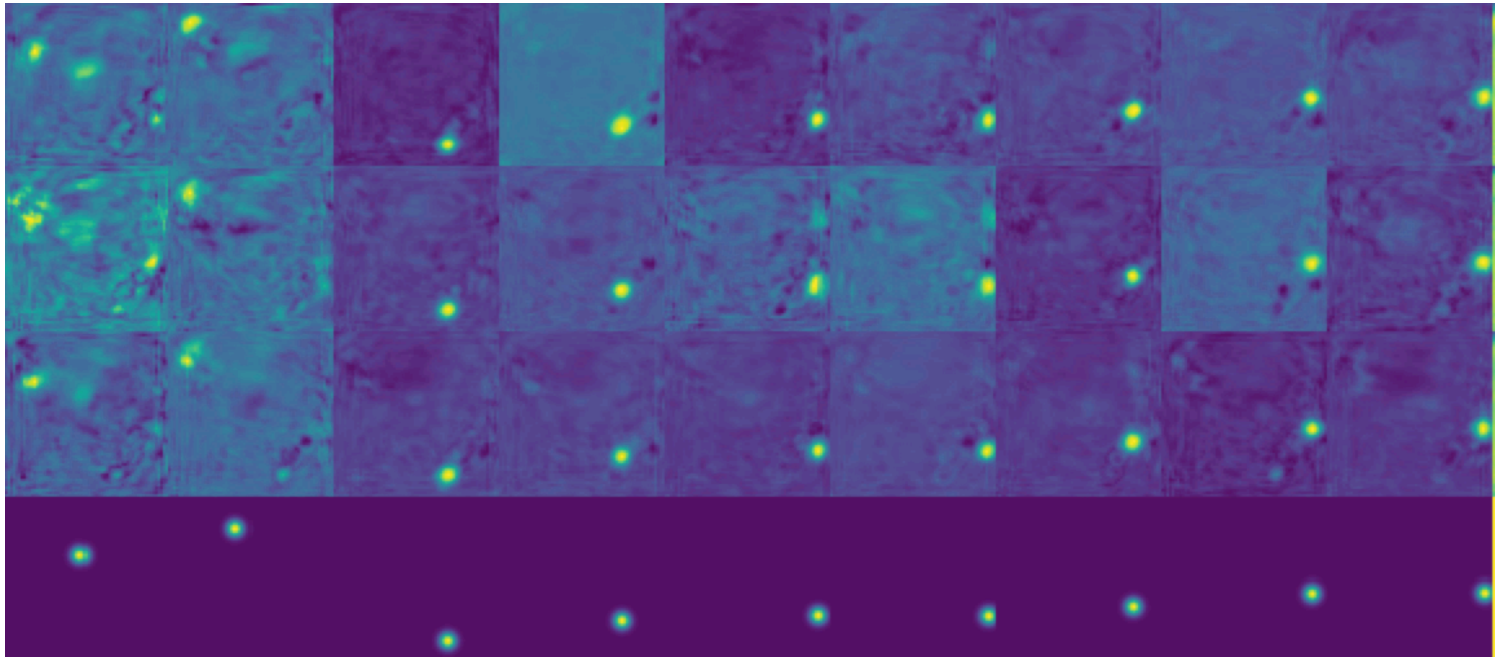
- Multistage Visualization - Intuition
- Spatial Error Plot - Quantitative Analysis
- Good Prediction / Bad Prediction - Case Study



Evaluation and Visualization

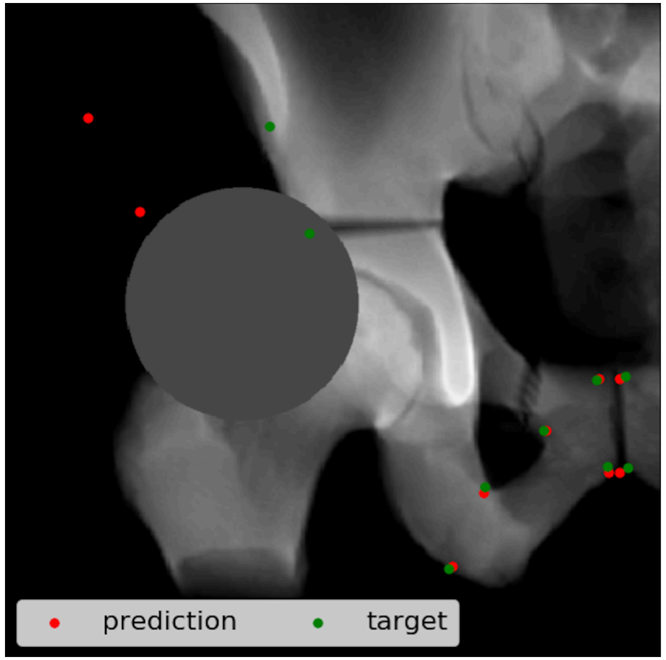
- Multistage Visualization - Intuition
- Spatial Error Plot - Quantitative Analysis
- Good Prediction / Bad Prediction - Case Study

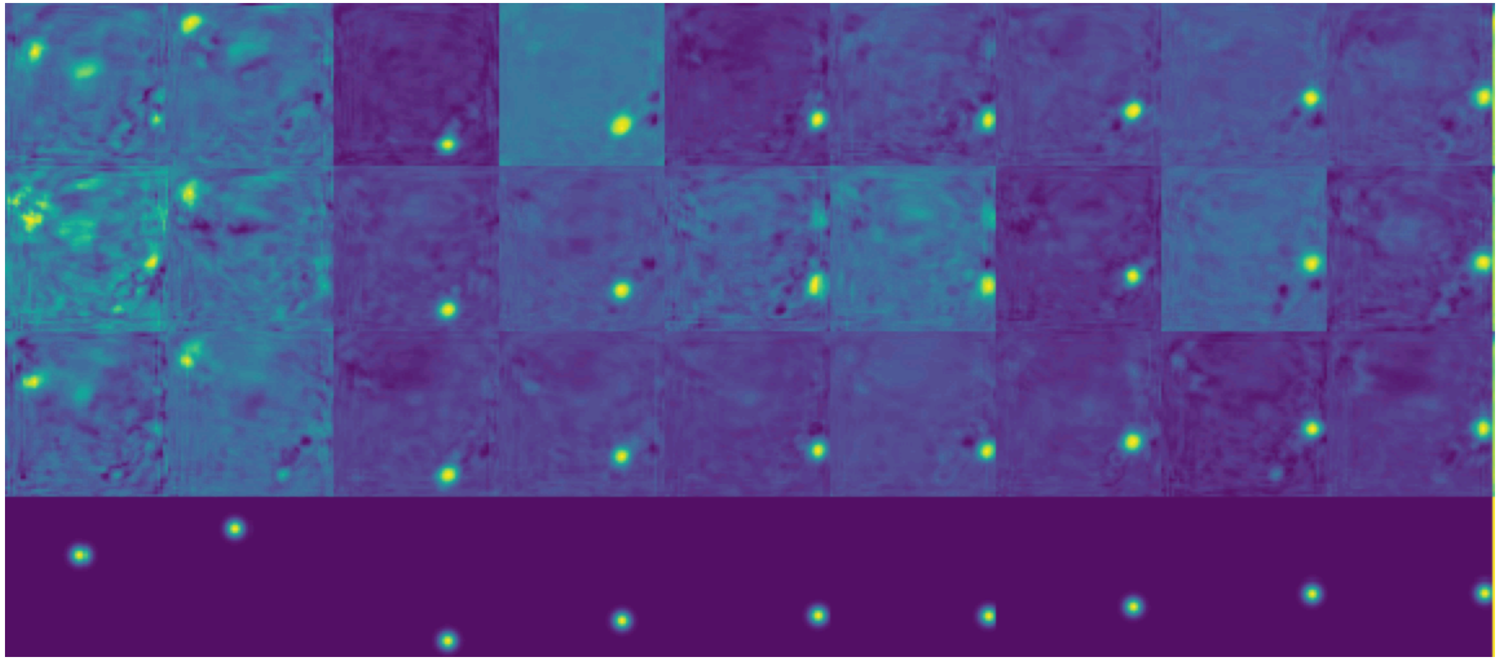




Why regress heatmap instead of position

- Heatmap gives **uncertainty** measure about the prediction
- Heatmap generation is essentially the **prior about the annotation**





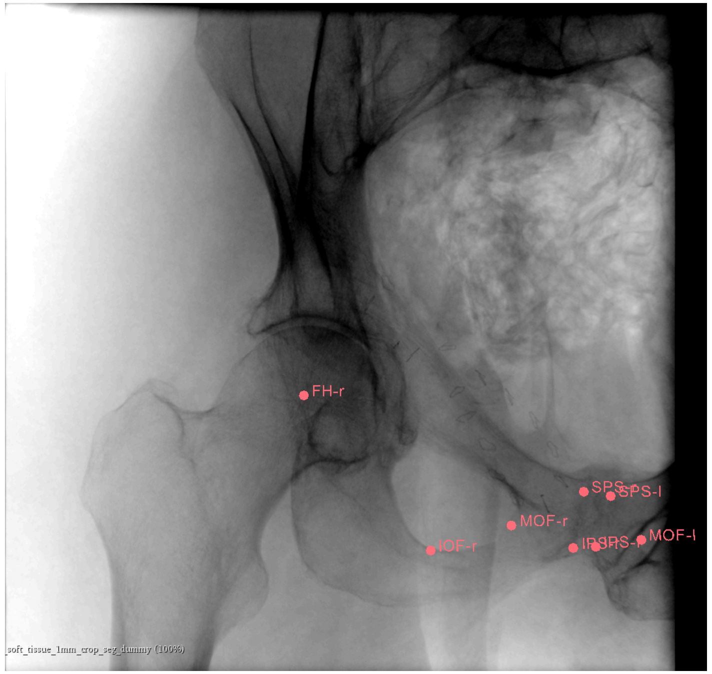
Why regress heatmap instead of position

- Heatmap gives **uncertainty** measure about the prediction
- Heatmap generation is essentially the **prior about the annotation**
- Could be used for detecting outliers
- 2D-3D registration by utilizing the distribution of 2D prediction
  - Let the prediction of landmark is given by $\mathcal{N}(\mu; \Sigma)$
  - The objective function could be derived by Mahalanobis distance

$$\underset{\theta \in SE(3)}{\arg\min} \sum_i \frac{1}{2} [P_{2D}^i - \mathcal{P}(P_{3D}^i; \theta)]^T \Sigma^{-1} [P_{2D}^i - \mathcal{P}(P_{3D}^i; \theta)]$$

Dependencies

- Simulated dataset with soft tissue / Real dataset
- Simulated dataset with contour annotated
- 2D-3D registration methods



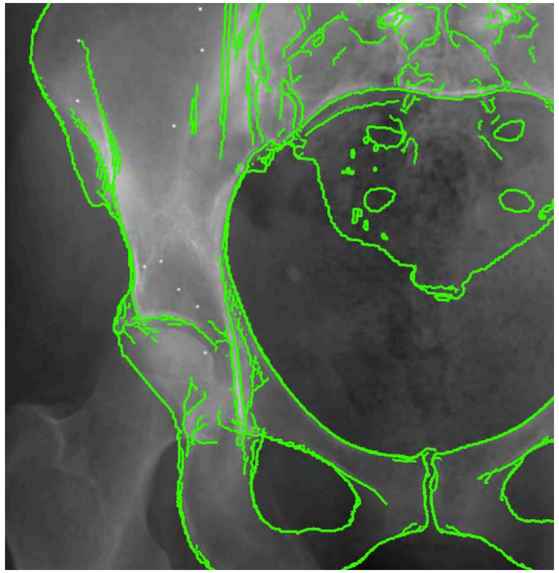
Real X-Ray Data

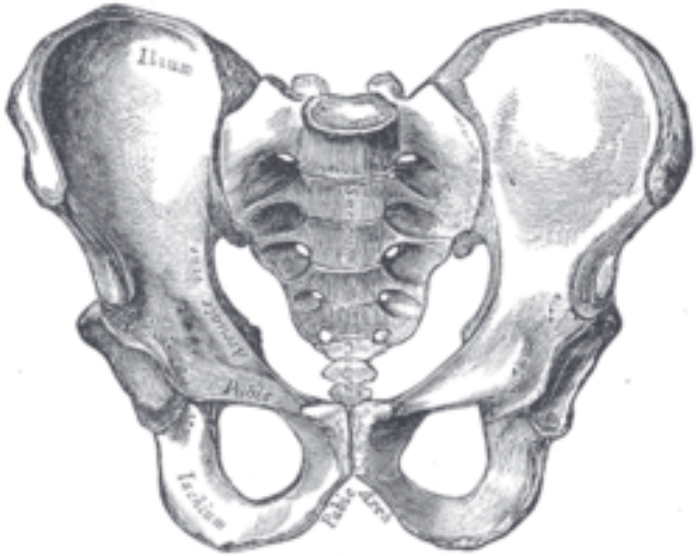
Simulated X-Ray Data

JOHNS HOPKINS
WHITING SCHOOL
of ENGINEERING

Dependencies

- Simulated dataset with soft tissue / Real dataset
- Simulated dataset with contour annotated
- 2D-3D registration methods





[1] Wikipedia contributors. (2018, March 17). Pelvis. In Wikipedia, The Free Encyclopedia. Retrieved 17:09, April 9, 2018.

Dependencies

- Simulated dataset with soft tissue / Real dataset
- Simulated dataset with contour annotated
- 2D-3D registration methods
  - Evaluate performance by comparing the led registration result.

$$\arg\min_{\theta \in SE(3)} \sum_i \frac{1}{2} \left\| \mathbf{p}_{2D}^{(i)} - \mathcal{P}(\mathbf{p}_{3D}^{(i)}; \theta) \right\|_2^2$$

JOHNS HOPKINS
WHITING SCHOOL of ENGINEERING

Conclusion

- Minimum deliverable achieved; Ongoing works on medium / maximum;
- Developed landmark detection pipeline using PyTorch;
- Tuned convolutional pose machine architecture;
- Perform data augmentation to address tool in the view problem;
- Evaluated and visualized results for further analysis;
- Need more training data (landmark, contour) for further experiment;
- Need registration method for evaluation.

Thank you for listening.