Introduction

- Introduced convolutional pose machines for anatomical landmark detection in simulated X-Ray images;
- Introduced holistically-nested networks for contour detection in simulated X-Ray images and transferred to real X-Ray images;
- Developed random region mask as data augmentation to simulate tool in the field;
- Perform thorough and detailed experiments to evaluate the models’ performance under different data.

Landmark Detection

- Convolutional pose machines consist a sequence of convolutional networks that repeatedly produce 2D belief map for each part.
- Belief maps at each stage represent the non-parametric encoding of the spatial uncertainty of the location for each part.
- The objective is to minimize the L2 distance defined by the belief map from each stage and the target belief map for each part.

\[ F = \sum_{t=1}^{T} \sum_{p=1}^{P} \sum_{x} ||b_{p}^{t}(x) - b_{gt}^{t}(x)||^2 \]

Contour Detection

- The architecture consists of a VGG convolutional net for multi-scale feature learning, and assign a side output at each stage, and a fusion layer aggregating all stages.
- The objective is to minimize class-weighted cross entropy defined at each side output and fusion output.

\[ L = \sum_{t=1}^{T} \left( -\beta \sum_{j \in I^t} \log(y_j = 1|x^t) - (1 - \beta) \sum_{j \in I^t} \log(y_j = 0|x^t) \right) \]

Data Augmentation

- randomly mask a region of arbitrary shape, arbitrary size, and arbitrary constant intensity to simulate tool in the field.

Outcomes and Results

- Achieved the minimum deliverables on landmark detection by using convolutional pose machines, and performed experiments to analyze the performance and limitations.
- Achieved the medium deliverables on tools in field of view by introducing random region mask as data augmentation
- Achieved the maximum deliverables on contour detection by utilizing HED network, and successfully transferred the model trained by soft tissue simulated data to real measured X-ray data.
- Not gathering more complex simulated data.

Support by and Acknowledgements

- Thank Robert Grupp for weekly mentoring and data preparation.
- Thank Professor Taylor for computational resources support.

Reference: