Deep Learning in Fluoroscopic Feature Detection

#14 Project Proposal, EN.601.656, Spring 2018
Liujiang Yan, Robert Grupp(Mentor)
lyan12@jhu.edu

1. Topics and Goal

   Intro-operation 2D-3D registration assists surgeon with spatial localization. However, current method is performed by manually identifying corresponding landmarks in 2D X-ray images, which is time consuming. The goal of this project is to utilize deep neural network with annotated simulated X-ray data to perform landmarks detection.

2. Relevance

   The motivation for using learning based method, specifically deep neural network for landmark detection in 2D fluoroscopic image comes from the intra-operative 2D-3D registration that is essential for surgeons to be aware of the spatial position of the operating fragment.

   Current feasible 2D-3D registration process consists of two steps. First of all, the corresponding anatomical landmarks are annotated both in 2D intra-operative X-ray image and 3D pre-operative model. Then, through the correspondence of landmarks, we could solve the registration problem by minimizing the projective error.

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

   However, while the landmarks are well defined in 3D model with anatomical meaning, it is not trivial to pick up such landmarks in 2D image views and needs experts to pick up them manually. Therefore, the landmarks annotation process is time consuming and cannot be performed in real-time manner especially while operating.

   This project tries to realize the real-time landmarks detection process by utilizing deep neural network with pre-annotated training data set. An obvious limitation is that real X-ray data is rare and we will first of all validate the model by simulated data. Furthermore, we will extend the landmark detection to contour detection, which might be a better representation feature to perform registration.

3. Technical Summery

   This project consists of the following components:

   a. Development of landmarks detection pipeline
   b. Deep neural network architecture design for landmarks detection
   c. Accuracy evaluation on simulated / real test data set

Intuition

Intuitively, landmark detection problem could be represented as position regression problem. Therefore, we could formulate the problem as, given training data set of images with corresponding landmarks positions, train model such that given a test image of same size, regress exact number of landmarks positions. Based on this problem formulation, we could derive various network architectures.

Landmark Detection Pipeline

General framework for deep learning problem includes several relatively individual components. The first component is to read dataset and transform the data into appropriate form. The second component is to define the network architecture and specify settings and parameters. Then we need components for training and testing the data. We might also implement component for visualization the results, including position error and registration difference.

Deep Neural Network Design

As a start, here we will present two deep neural network architecture designs for the landmark detection.

The first architecture is to couple convolutional layers and fully connected layers for position regression. The convolutional layers could efficiently extract feature map from the input image, and the following fully connected layer could perform the position regression with feature map extracted by convolutional layers. Here, the loss function here could be the sum of square error of the positions.

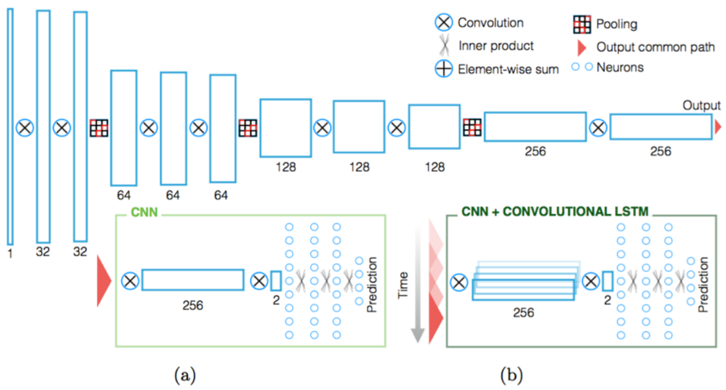


Figure 1 Convolutional Layers Followed by Fully Connected Layers

The second architecture treats landmark positions in a heatmap. Therefore, the architecture generally takes image as input and infer a heatmap as output. One of the advantage of this specific architecture is that we could also address the visibility of landmarks (if a landmark does not exist in the field of view, the corresponding heatmap is empty), while the disadvantage is that it might lead to inaccurate position due to the downsampling and discrete localization.

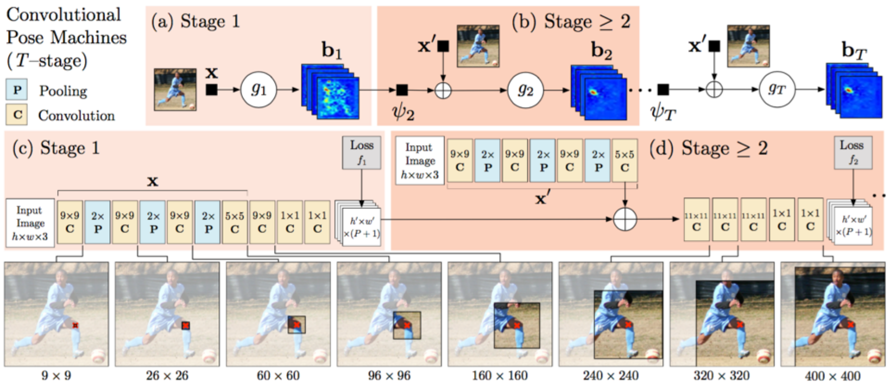


Figure 2 Heat Map Regression for Landmark detection

<u>Accuracy Evaluation</u>

Once we have trained the network by training data set and validated and tuned by validation data set. We could evaluate the accuracy by test data set, which may consist of simulated and / or real X-ray data set. To evaluate the generalization, we will use a separate data set generating from a separated patient's model. As for the evaluation criteria, we will use two stage evaluation, firstly by mean square error of landmarks position, and secondly by the difference of led 2D-3D registration error.

4. Deliverables

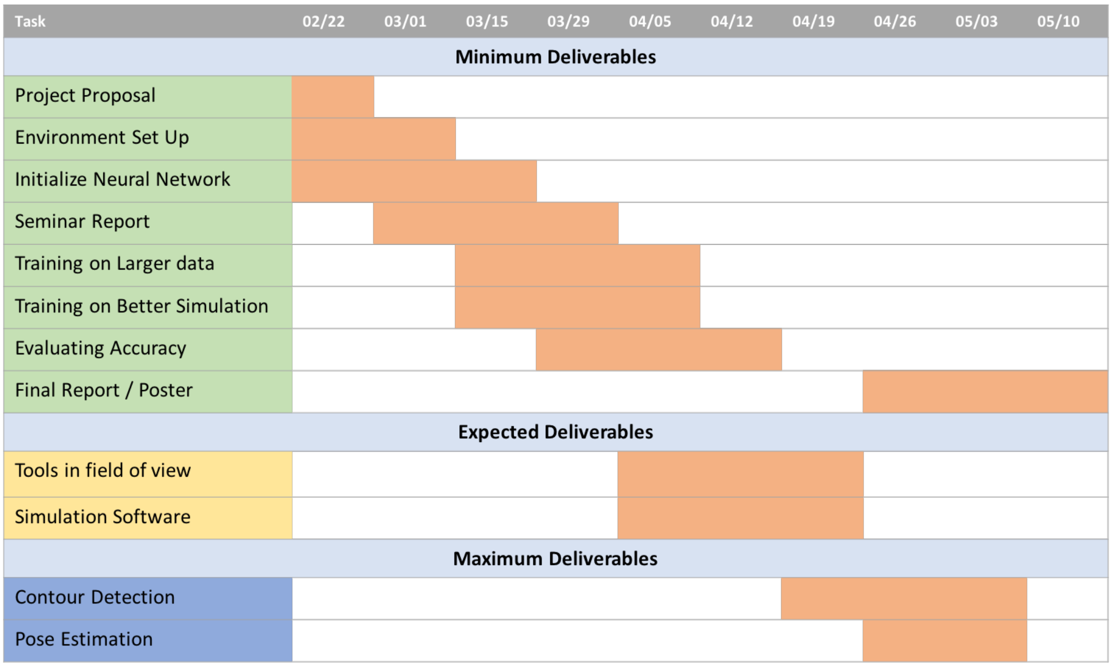
The deliverables consist of three parts: minimum, expected and maximum. Each specific deliverable comes with the details and criteria.

- Minimum Deliverables
  - Environment setup and data preparation.
    - ✓ Data loading, pre-processing, and post-processing and evaluation framework are implemented.
  - Initial network architecture validated on small data set.
    - ✓ Perform position regression with designed network architecture and validate the accuracy on small data set.
  - Network training on larger / refined data.
    - ✓ Deploy the designed network with larger data set. Perform cross validation to validate the accuracy. Fine tune the hyper parameters as well.
  - Accuracy Report by evaluated on simulated / real data.
    - ✓ Evaluate the trained model by real data set (from Robert Grupp) and / or separate test data set (from simulation).
- Expected Deliverables
  - Tools segmentation from field of view.
  - Better simulation software involved.
    - ✓ Utilize better simulation tools for soft tissue and changes in intrinsics parameters.
- Maximum Deliverables
  - Edge / contour detection.

        - ✓ Train network for contour detection with annotated data set, and validate the performance both by accuracy and led registration accuracy.
    - Pose estimation.

5. Dependencies
    - Access to simulated training data set
        - We have an initial data of 13 patients with both left and right side sub data set and each side has varying fragments and directions of projection. The data set is generated by simulation, with landmarks annotated.
        - We will use more realistic simulation tools for generating more data, both for training and validating the proposed network architecture.
    - Access to High Performance Computer
        - Training deep neural network consumes huge computational resources. We have permission to work on thin6. If the server does not meet the needs, we might use online cloud platform for training.
    - Access to Mentors
        - We have scheduled weekly meeting with Robert Grupp at each Thursday afternoon.

6. Management Plan
    - Basics
        - Source code and version will be maintained via GitHub private repository.
        - Document will be kept and updated timely via CIS2 course wiki webpage.
        - Weekly meeting has been scheduled with Robert Grupp to keep updating the development. Additional meetings and paper discussions will be held if necessary.
    - Key Milestones
        - February 25: All simulated data obtained and validated.
        - March 15: Environment and initial training network Set up.
        - March 29: Deploy neural network for complete training data set.
        - April 12: Evaluate accuracy in (simulated / real) test data set.
            - ✓ Minimum Deliverables Achieved.
        - April 19: Utilize advanced software for simulated data.
        - April 19: Training network for segment tools in field of view.
            - ✓ Expected Deliverables Achieved.
        - May 03: Training network for contour detection.
            - ✓ Maximum Deliverables Achieved.
        - May 11: Final Report / Poster Session.



7. Reading List

Deep Learning Background

- LeCun, Yann, et al. "Gradient-based learning applied to document recognition." Proceedings of the IEEE 86.11 (1998): 2278-2324.
- LeCun, Yann, et al. "Backpropagation applied to handwritten zip code recognition." Neural computation 1.4 (1989): 541-551.
- Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "Imagenet classification with deep convolutional neural networks." Advances in neural information processing systems. 2012.
- LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton. "Deep learning." nature 521.7553 (2015): 436.
- Goodfellow, Ian, et al. Deep learning. Vol. 1. Cambridge: MIT press, 2016.

Feature / Landmark Detection

- Wang, Ching-Wei, et al. "Evaluation and comparison of anatomical landmark detection methods for cephalometric x-ray images: a grand challenge." IEEE transactions on medical imaging 34.9 (2015): 1890-1900.
- Payer, Christian, et al. "Regressing heatmaps for multiple landmark localization using CNNs." International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, Cham, 2016.
- Urschler, Martin, Thomas Ebner, and Darko Štern. "Integrating geometric configuration and appearance information into a unified framework for anatomical landmark localization." Medical image analysis 43 (2018): 23-36.

- Sofka, Michal, et al. "Fully convolutional regression network for accurate detection of measurement points." Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. Springer, Cham, 2017. 258-266.
- Ronneberger, Olaf, Philipp Fischer, and Thomas Brox. "U-net: Convolutional networks for biomedical image segmentation." International Conference on Medical image computing and computer-assisted intervention. Springer, Cham, 2015.
- Bulat, Adrian, and Georgios Tzimiropoulos. "Human pose estimation via convolutional part heatmap regression." European Conference on Computer Vision. Springer, Cham, 2016.

X-Ray Simulation

- Badal, Andreu, and Aldo Badano. "Accelerating Monte Carlo simulations of photon transport in a voxelized geometry using a massively parallel graphics processing unit." Medical physics 36.11 (2009): 4878-4880.
- Badal, Andreu, and Aldo Badano. "Monte Carlo simulation of X-ray imaging using a graphics processing unit." Nuclear Science Symposium Conference Record (NSS/MIC), 2009 IEEE. IEEE, 2009.
- Bert, Julien, et al. "Geant4-based Monte Carlo simulations on GPU for medical applications." Physics in Medicine & Biology58.16 (2013): 5593.
- Jia, Xun, et al. "A GPU tool for efficient, accurate, and realistic simulation of cone beam CT projections." Medical physics39.12 (2012): 7368-7378.

8. References

- Wei, Shih-En, et al. "Convolutional pose machines." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016.
- Sofka, Michal, et al. "Fully convolutional regression network for accurate detection of measurement points." Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. Springer, Cham, 2017. 258-266.