A U-Net Framework for Detecting K-Wires in Fluoroscopic Images

Friede, J., Golec, J, and Mattioli, D.
joshua-freide@uiowa.edu, john-golec@uiowa.edu, dominik-mattioli@uiowa.edu
Industrial & Systems Engineering: 6380 - Deep Learning, The University of Iowa, USA

Objective

Train a model to detect and segment K-Wires in Fluoroscopy.

I. Background & Motivation

Many contemporary research efforts in Orthopedics aim to improve surgical residents' skills before these surgeons enter an Operating Room (OR) – commonly-understood to be both a high-risk and a high-cost environment.
To help reduce both costs and risks, an ongoing multi-disciplinary project at The University of Iowa is focusing on developing simulators for several surgeries. The aim of these simulators is to imitate various procedures that use fluoroscopy (fluoros) for tracking surgical progress. The simulator mimics these procedures and then scores the participant's effort.
Establishing a methodology for scoring a wire navigation should include an objective evaluation of the decisions made by the surgeon. Accomplishing this requires requires a reliable mechanism for tracking:

1. The wire's progression across consecutive images, and
2. The final location of the wire, relative to its destination.

Gaining this understanding will require analysis of many images across many surgeries. Currently, there is not readily available amount of annotated images for accomplishing this.

II. The Data

To compensate for a lack of real-life data to train on, this research looks to validate recently observed results showing that synthetic image data can be used to train a deep neural network object detection model [1].
The synthetic data for this model consists of 4,620 random augmentations of a base wire template, which are then overlaid onto 1,540 corresponding source images.
The source images consist of: 385 hip surgery images (provided in-house), 385 X-ray (randomly selected from MedPix[2]), and 770 random images (randomly selected from The Coco Dataset[3]). Augmentations of the wire include translation, shear, bend, duplication, reflection, and scale.

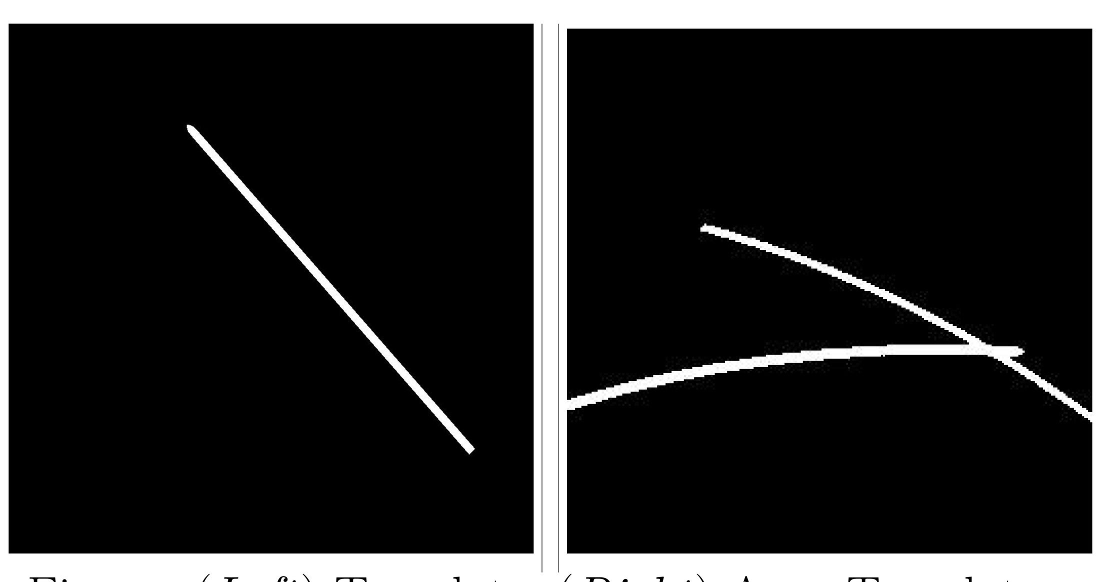

Figure: (*Left*) Template; (*Right*) Aug. Template.

For each augmentation, the wire appears as a random gray-intensity value between 128 (gray) and 0 (black).

II. The Data - Cont.

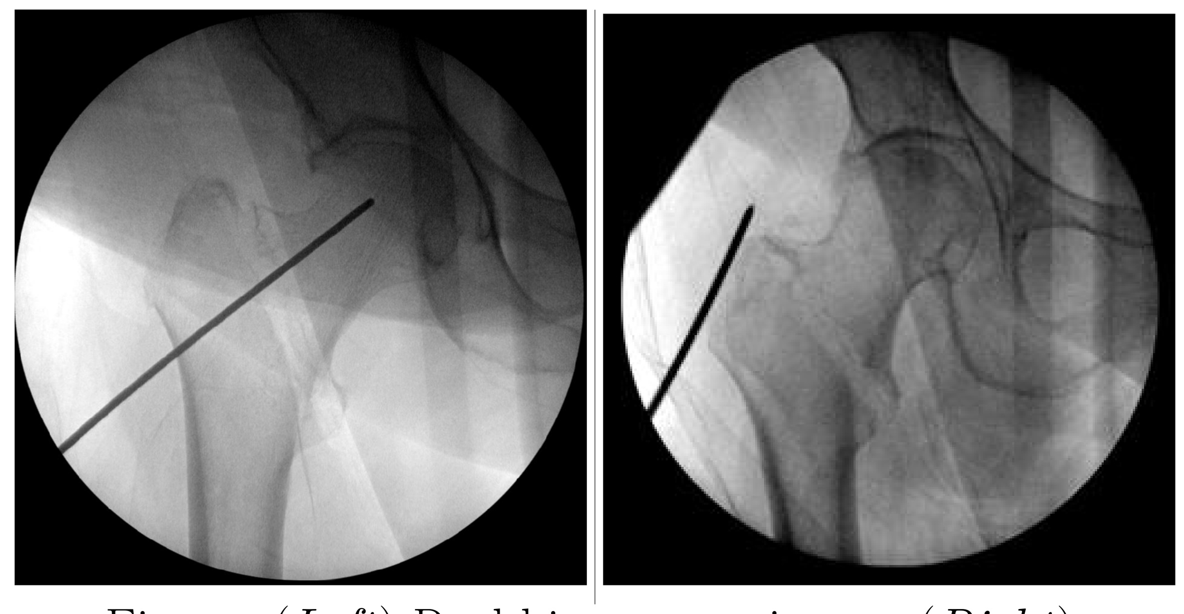

Figure: (*Left*) Real hip surgery image; (*Right*) Synthetic image with an overlaid, augmented wire template.

III. Methods

Segmentation of the wire in each synthetic image is accomplished via U-Net model architecture, which has shown to provide high accuracy in biomedical imaging applications[4]. U-Net is a fully convolutional network (FCN), which is comprised one initial gaussian noise layer following the input layer with a standard deviation of 0.1, a *contraction path* and an *expansion path*, as shown in the figure below.

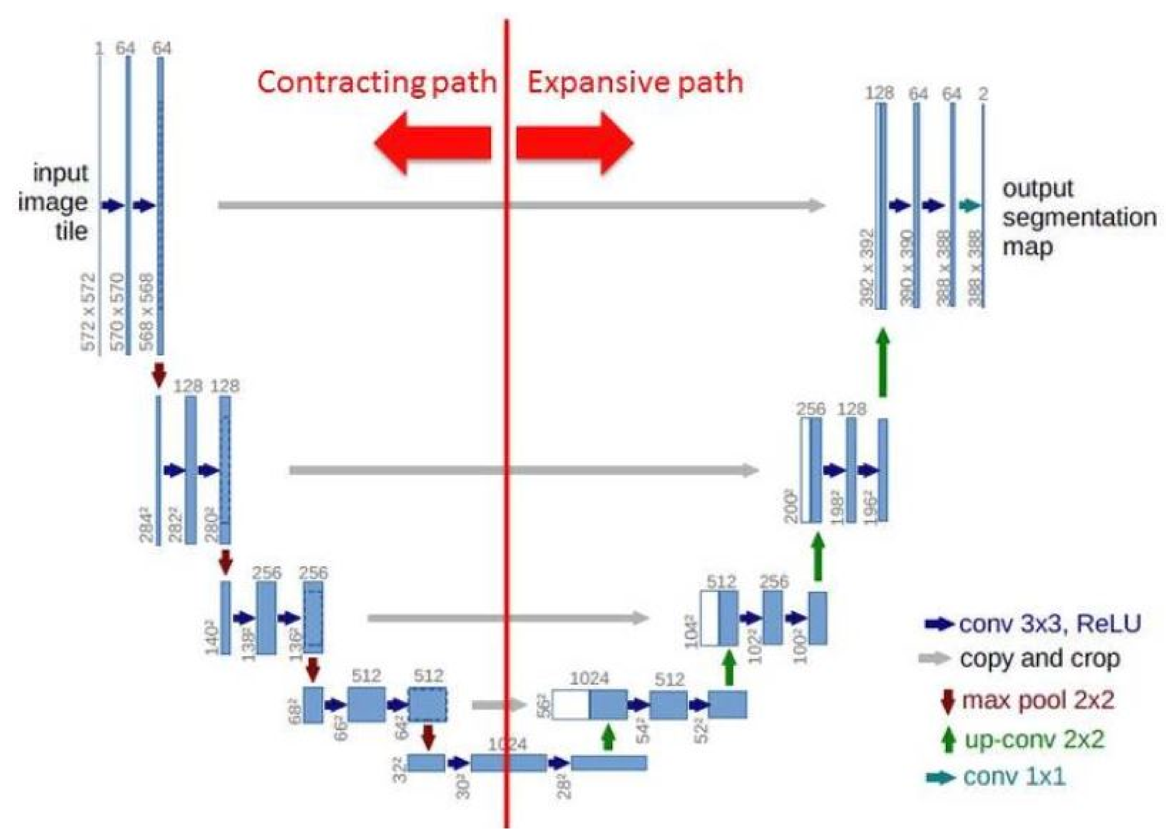

Figure 1: U-Net Architecture

The *contraction path* consists of a sequence of two 3x3 convolution layers and a 2x2 max pooling layer, while the *expansion path* consists of a sequence of a 2x2 up convolution layer and two 3x3 convolution layers. The contraction path also contains a dropout layer of 30% after each max pooling layer. In addition, *expansion path* contains a feature map concatenation from the contraction layer after each up convolutional layer. The dropout layer in the U-Net architecture is commonly used as a regularization method to prevent over-fitting.
This model was trained with checkpoints for 50 epochs with a validation split of 0.20 and a batch size of 32, and uses the SøRensen-Dice coefficient (DSC) for evaluation and training loss. DSC is a similarity measurement of two discrete sets of data (X and Y).

$$DSC = \frac{2|X \cap Y|}{|X| + |Y|}$$

IV. Results

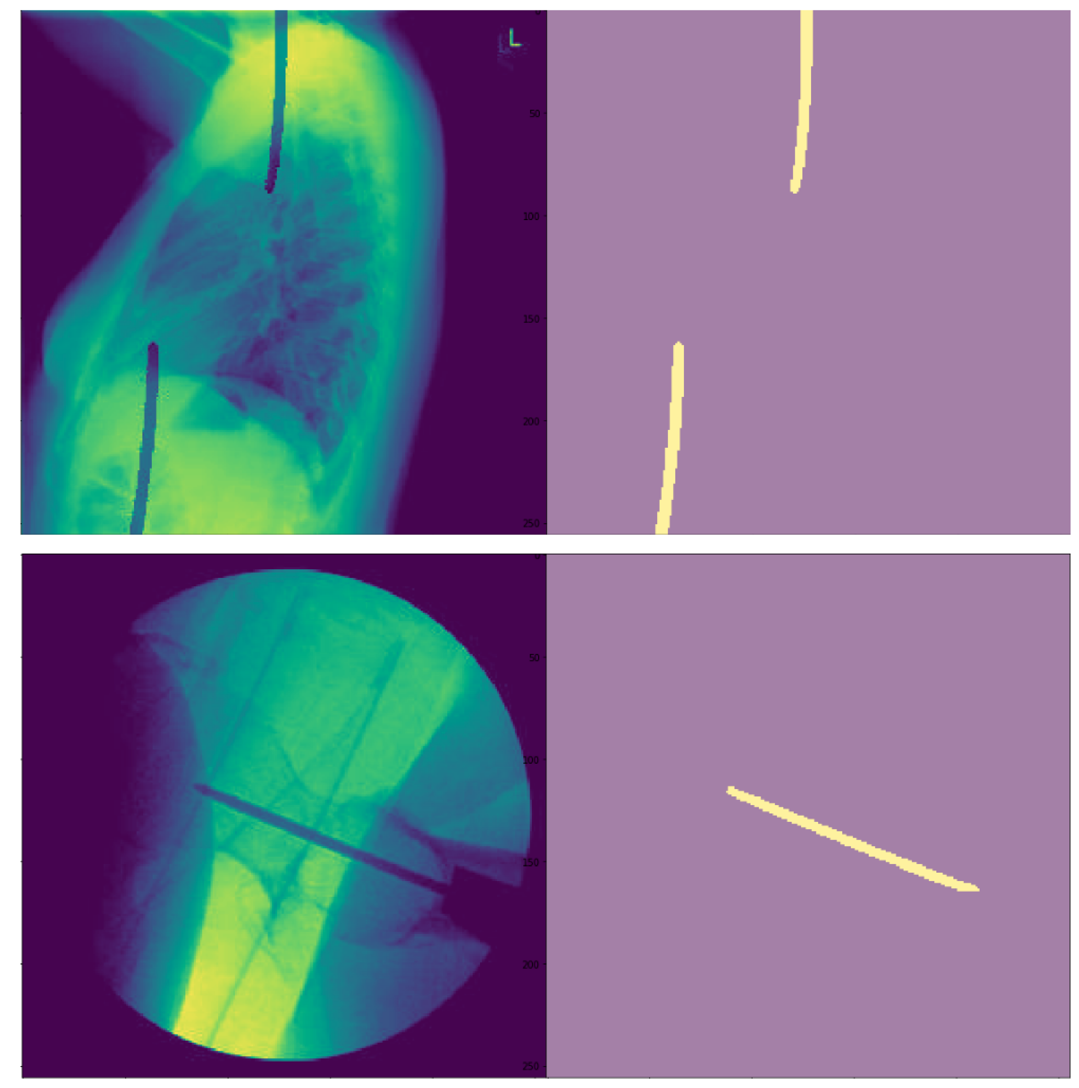

Figure: Prediction (90% probability): (*Top*) synthetic; (*Bottom*) validation.

V. Future Work

Future iterations of this research should implement a more meticulous methodology for generating the synthetic images, perhaps on even larger data sets and with better-balanced ratios of source images.
The preliminary success of this model indicates that additional objects could be trained for detection. Detecting other anatomical features would give a more precise quantification of the wire's relative location, and allow for more robust scoring.

References

[1] Jonathan Tremblay, Aayush Prakash, David Acuna, Mark Brophy, Varun Jampani, Cem Anil, Thang To, Eric Cameracci, Shaad Boochoon, and Stan Birchfield.
Training deep networks with synthetic data: Bridging the reality gap by domain randomization.
CoRR, abs/1804.06516, 2018.

[2] 000 images. MedPix-Medical (radiological) image database with more than 20.
X-ray image dataset; exclude graphics, 2019.

[3] Cocodataset.org.
2017 val images [5k/1gb], 2017.

[4] Olaf Ronneberger, Philipp Fischer, and Thomas Brox.
U-net: Convolutional networks for biomedical image segmentation.
CoRR, abs/1505.04597, 2015.

Acknowledgements

This research was made possible by Dr. Stephen Baek.

THE UNIVERSITY OF IOWA