

Tool Tracking in Orthopedics Surgery

Seminar Paper Critical Review

Athira Jacob, Group 3

Background and Goals

K-wire insertion is a widely prevalent procedure in orthopedics surgery. K-wires are long, smooth stainless steel pins that are used for variety of tasks, from holding bones together to guiding screw insertion. Currently, K-wire insertion is done using numerous intra-operative X-ray images, which help the surgeon to mentally align the wire and bone. However, misplacement of the K-wire could cause severe damage to major structures [1]. As a result, this often requires multiple attempts [2]. This leads to multiple entry wounds on the patient and high X-ray exposure for the patient and clinicians, increasing overall OR time and staff frustration. Recently, camera augmented solutions have been proposed to help surgeons with mental alignment of the patient, the X-ray scan, and the tool [3,4].

We propose a convolutional neural network based solution to segment the K-wire in RGBD images to supplement augmented reality guidance systems. Traditional computer vision algorithms fail at the task of tracking K-wire due to reflections, specularities, occlusions etc. The first paper chosen relates to a pre-clinical study showing the benefits of mixed reality visualization systems for the same. The second paper describes a possible CNN architecture that can be used for the segmentation.

Paper 1: Fischer et. al (CAMP, JHU) , 'Preclinical usability study of multiple augmented reality concepts for K-wire placement', Int J CARS (2016)

Background

The main challenge during percutaneous K-wire placement and screw fixation is the mental alignment of patient, medical instruments, and the intra-operative X-ray images, which also requires frequent repositioning of the C-arm. The standard treatment procedure for undisplaced superior pubic ramus fractures requires several K-wire placements and subsequent screw insertions. For each K-wire, the surgeon first locates the entry point location and performs a skin incision at the lateral side of the hip, which requires several intra-operative X-ray images from various perspectives to confirm the exact tool orientation. Computer-aided surgical navigation systems have been introduced to assist the placement of K-wires and screws. However, the benefits of navigation systems are controversial [5].

Another line of research follows using mixed reality visualization systems to provide visual feedback to the surgeon [3,4]. This paper presents a preclinical usability study to provide a more comprehensive understanding of whether enhanced C-arm systems provide a clinically relevant benefit.

Methods

The paper compares three imaging techniques for guidance during K-wire insertion. Seven surgeons perform the procedure with a phantom made to mimic tissue and bone. The phantom was created out of methylene bisphenyl diisocyanate (MDI) foam, which is stiff, lightweight, and not radiopaque. The bone phantom was created out of a thin aluminum mesh filled with MDI. The beginning and end of the bone were marked with a rubber radiopaque ring. Therefore, the bone phantom is very similar to the superior pubic ramus in terms of haptic feedback during K-wire placement, as the K-wire will easily exit the bone without significant resistance.

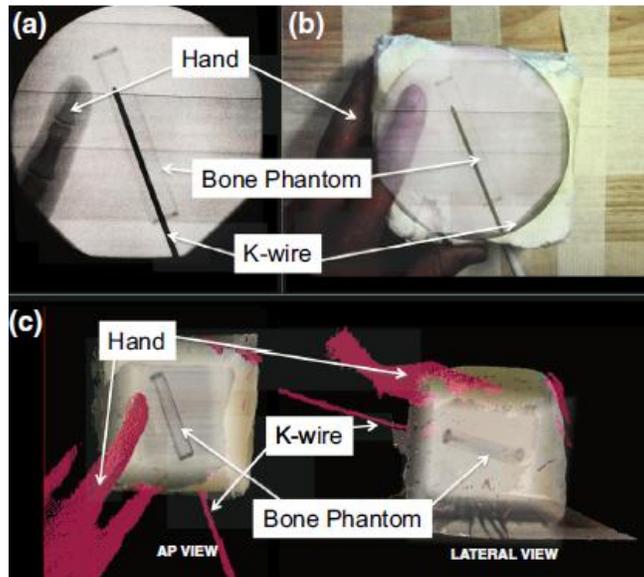


Figure 1 Same stage in the K-wire placement has been recreated using the different image-guidance systems. a) Conventional intra-operative X-ray imaging b) 2D RGB video and X-ray visualization c) 3D RGBD and DRR via CBCT visualization

The three imaging techniques compared are:

a) Conventional intra-operative X-ray imaging

This imaging method using a standard C-arm provides the baseline performance as it is the most commonly used system to perform image-guided K-wire placement. (Fig 1a)

b) 2D RGB video and X-ray visualization

X-rays are augmented onto 2D camera feed on a screen. This involves a RGB camera rigidly attached to the C-arm. The alignment registration of optical and X-ray images is performed using a single plane phantom with radiopaque markers that are also visible in the optical view. A problem with this kind of visualization is that a new X-ray image has to be taken each time the surgeon wants a new view. (Fig 1b)

c) 3D RGBD and DRR via CBCT visualization

This imaging system consists of 2D Digitally Reconstructed Radiographs (DRR) augmented onto a 3D surface model of the object, obtained from an RGBD camera. A CBCT volume image of the phantom is taken before the procedure. An RGBD camera rigidly attached to the C-arm continuously provides 3D information about the phantom. The two are aligned by surface matching. As a result, any arbitrary view of the phantom can be displayed, with the corresponding X-ray sliced from the CBCT, augmented onto it. (Fig 1c)

The metrics on which these systems are compared are given below:

1. Duration of each K-wire placement
2. Number of X-ray images
3. Cumulative dose
4. Error in placement
 - Average distance from the center line of bone phantom
5. Surgical task load
 - Surgical Task Load Index questionnaire (SURG-TLX)

Results

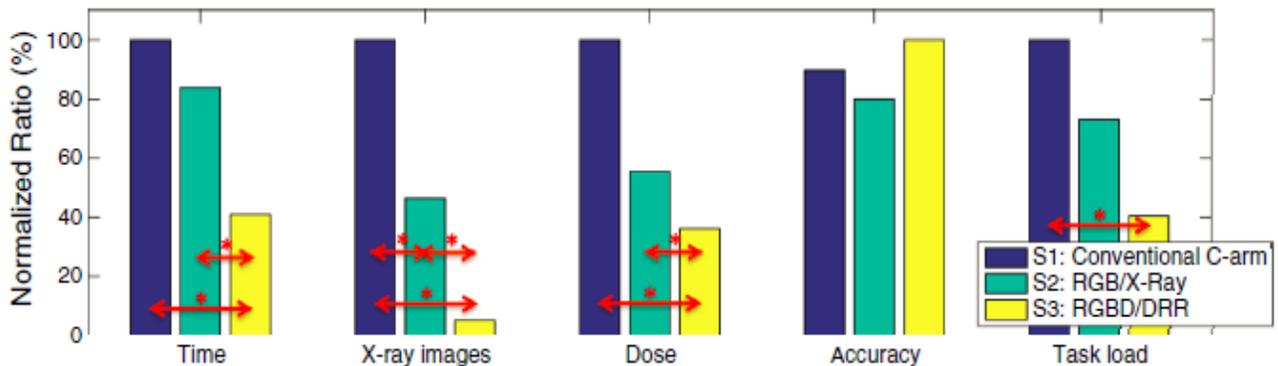


Figure 2 Each bar shows the accumulated values using one of the systems (conventional X-ray, RGB/X-ray fusion, or RGBD/DRR). Each measure is normalized relative to the maximum value observed. The * symbols indicate significant differences the systems (conventional X-ray, RGB/X-ray fusion, or RGBD/DRR).

The results (Fig 2) show statistically significant benefits for using 3D augmentation of DRR on RGBD data for all metrics, except for accuracy. The authors justify this by saying that since the surgeons are experienced, trained surgeons, no apparent differences in accuracy are expected.

Conclusions

The paper does a thorough job of investigating the benefits of the new guidance systems they propose. The authors use appropriate statistical testing to take into account the small sample size and also follow established registration methods to align their different views. The results for every surgeon are reported in the paper. Thus mixed reality visualizations are a promising alternative to conventional guidance systems, as demonstrated by this study. This is despite the fact that currently there is no system to track the tool. A tool tracking method, like the one proposed by our project, could potentially make these benefits even more evident.

Paper 2: Ronneberger et. al (University of Freiburg, Germany), Convolutional Networks for Biomedical Image Segmentation, MICCAI 2015

Background

Deep learning has shown remarkable successes in the recent past [6], mainly due to deeper networks, larger datasets, better optimization techniques etc. Segmentation is a typical task that deep learning algorithms generally excel at. Traditionally, pixel wise segmentation is done through a 'patch' based approach, where the image is divided into many patches, each surrounding a pixel. This method, however, is slow and inefficient, as we need to perform a forward pass for each patch and typically a single image will have thousands of over-lapping patches. In addition, there is a tradeoff between context and localization accuracy. Fully convolutional networks offer to solve some of these problems [7]. In such networks, the fully connected layers are replaced by convolutional layers, hence retaining spatial context. In addition, this allows end-to-end training, with any input size. The contracting path is supplemented by up-sampling path, that up-samples the images to the size required. This gives an efficient, fast way of training a segmentation network.

U-Net is a modification of the FCN that has given state of the art results in the domain of biomedical image segmentation. This task faces some challenges similar to our task and thus makes it a potential candidate for solving our problem of segmenting the K-wire.

Network Architecture

Some of the novel features that the authors introduce are:

- a. Extensive data augmentation

Labelled data in the field of cell tracking is expensive, as it requires expert hand-labelling. As a result the network had around 30 images as the training set, which is much lesser than what a typical neural network requires. However the authors propose a novel augmentation method of using smooth affine deformations to artificially expand the dataset.

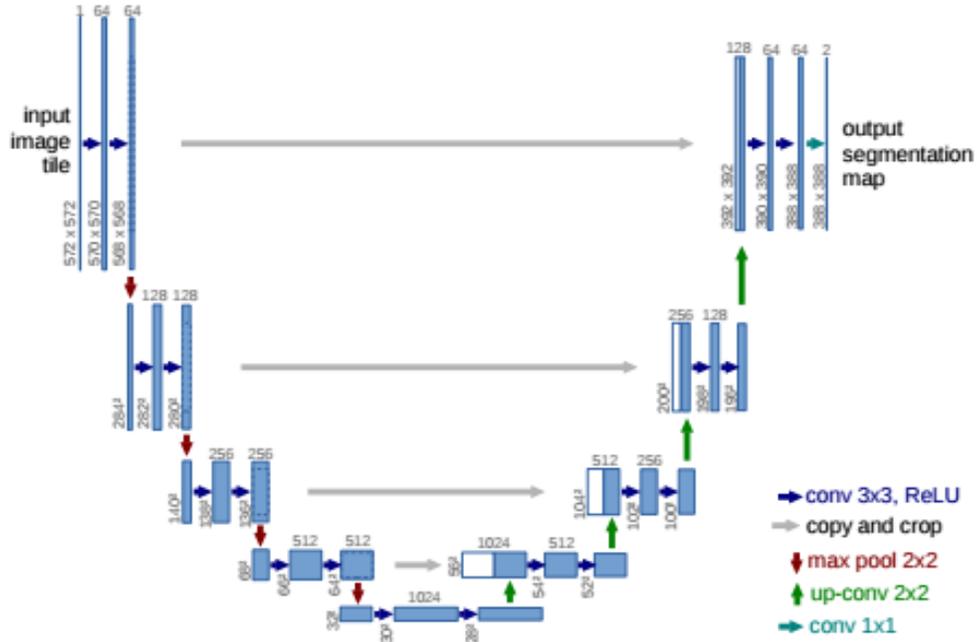


Figure 3 U-Net architecture

b. Custom weight balancing

Class balancing is a typical reason to use weighted loss functions. However, the authors use an additional weight term that gives high weights to background pixels separating cells. This forces closely located cells to be segmented separately and not merge into one.

$$w(\mathbf{x}) = w_c(\mathbf{x}) + w_0 \exp\left(-\frac{(d_1(\mathbf{x}) + d_2(\mathbf{x}))^2}{2\sigma^2}\right)$$

$w_c : \Omega \rightarrow \mathbb{R}$: weight map to balance the class frequencies,
 $d_1 : \Omega \rightarrow \mathbb{R}$: distance to the border of the nearest cell
 $d_2 : \Omega \rightarrow \mathbb{R}$: distance to the border of the second nearest cell

Results

The performance of the network is tested on three datasets (DIC Hela cells, PhC-U373 and EM neuronal structures) and U-Net shows state of the art results on all three.

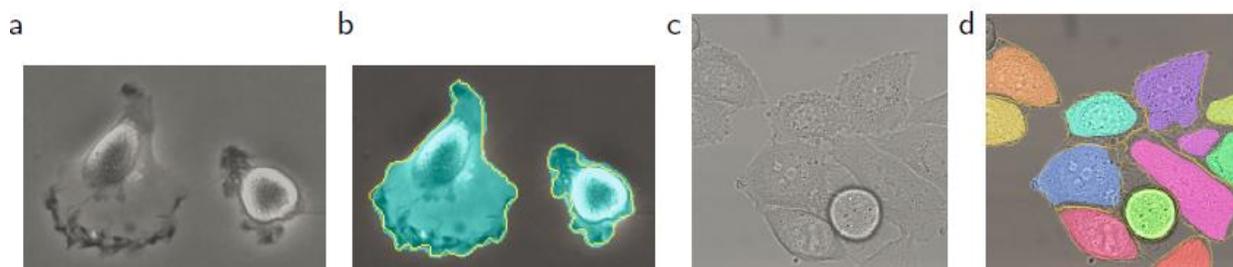


Figure 4 (a) part of an input image of the “PhC-U373” data set. (b) Segmentation result (cyan mask) with manual ground truth (yellow border) (c) input image of the “DIC-HeLa” data set. (d) Segmentation result (random colored masks) with manual ground truth (yellow border).

Name	PhC-U373	DIC-HeLa
IMCB-SG (2014)	0.2669	0.2935
KTH-SE (2014)	0.7953	0.4607
HOUS-US (2014)	0.5323	-
second-best 2015	0.83	0.46
u-net (2015)	0.9203	0.7756

Figure 5 Segmentation results (IOU) on the ISBI cell tracking challenge 2015

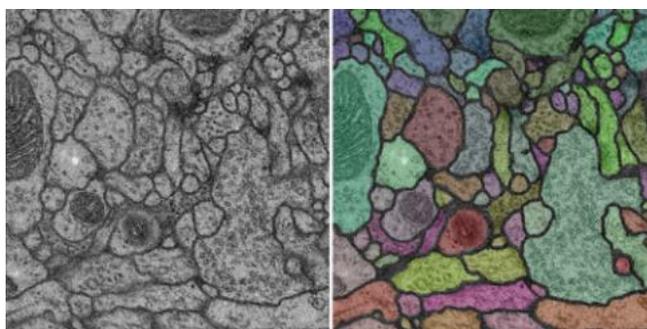


Figure 6 ISBI EM Segmentation challenge, Sample image and corresponding ground truth

Rank	Group name	Warping Error	Rand Error	Pixel Error
	** human values **	0.000005	0.0021	0.0010
1.	u-net	0.000353	0.0382	0.0611
2.	DIVE-SCI	0.000355	0.0305	0.0584
3.	IDSIA [2]	0.000420	0.0504	0.0613
4.	DIVE	0.000430	0.0545	0.0582
	⋮			
10.	IDSIA-SCI	0.000653	0.0189	0.1027

Figure 7 Ranking on the EM segmentation challenge [14] (march 6th, 2015), sorted by warping error.

Conclusions

The authors provide their implementation of the U-Net on their website. All details of implementation are described, between the paper and their website. However it is not clear why they used the metric 'Warping Error' for the EM dataset, as it is not one of the two metrics suggested by the competition organizers. It is to be noted that U-Net does not give the best results on these two metrics, as seen in the table above, though it is mentioned that this better performance could be due to extensive post-processing done by other participants. The authors note that U-Net gives these performances without any pre/post processing. In conclusion, U-Net does seem to be a useful variant of the FCN, with many potential uses in various segmentation tasks.

References

- [1] R. Rohilla, R. Singh, N. K. Magu, A. Devgan, R. Siwach, and S. S. Sangwan, "Simultaneous Use of Cannulated Reamer and Schanz Screw for Closed Intramedullary Femoral Nailing," *ISRN Surg.*, vol. 2011, 2011
- [2] B. Diotte et al., "Radiation-Free Drill Guidance in Interlocking of Intramedullary Nails," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2012*, 2012, pp. 18–25.
- [3] Navab, Nassir, Sandro-Michael Heining, and Joerg Traub. "Camera augmented mobile C-arm (CAMC): calibration, accuracy study, and clinical applications." *IEEE transactions on medical imaging* 29.7 (2010): 1412-1423.
- [4] Habert, Séverine, et al. "Rgbdx: First design and experimental validation of a mirror-based rgbd x-ray imaging system." *Mixed and Augmented Reality (ISMAR), 2015 IEEE International Symposium on*. IEEE, 2015.
- [5] Kraus M, Weiskopf J, Dreyhaupt J, Krischak G, Gebhard F (2015) Computer-aided surgery does not increase the accuracy of dorsal pedicle screw placement in the thoracic and lumbar spine: a retrospective analysis of 2,003 pedicle screws in a level I trauma center. *Global Spine J* 5(2):93
- [6] LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton. "Deep learning." *Nature* 521.7553 (2015): 436-444.
- [7] Long, Jonathan, Evan Shelhamer, and Trevor Darrell. "Fully convolutional networks for semantic segmentation." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2015.