Instruments Segmentation in X-ray Fluoroscopic Images for Endoscopic Retrograde Cholangio Pancreatography

Garance Martin, Saad El-Madafri, Aymeric Becq, Jérôme Szewczyk, Isabelle Bloch

Instruments Segmentation in X-ray Fluoroscopic Images for Endoscopic Retrograde Cholangio Pancreatography

Garance MARTIN[a,1], Saad EL-MADAFRI[a], Aymeric BECQ[b,c,2],
Jérôme SZEWCZYK[b] and Isabelle BLOCH[a]

[a] *Sorbonne Université, CNRS, LIP6, Paris, France*
[b] *Sorbonne Université, CNRS, INSERM, ISIR, Paris, France*
[c] *UPEC, Hôpital Henri Mondor, Service de Gastroentérologie, APHP, Créteil, France*

Abstract. In this work, we propose a method to segment endoscope and guidewire from 2D X-ray fluoroscopic images of an endoscopic retrograde cholangiopancreatography (ERCP). We used an improved U-Net model. We obtained a Dice score of 0.94±0.05 for endoscope segmentation and a Hausdorff distance of 24.26 pixels for the guidewire segmentation. These preliminary results pave the way for further applications aiming at aiding the medical procedure.

Keywords. ERCP, X-Ray Fluoroscopy, U-Net

1. Introduction

Endoscopic Retrograde CholangioPancreatography is a minimally invasive procedure allowing draining the bile ducts in the setting of stones or neoplastic structuring. The instruments and guidewire used are radio-opaque and can therefore be monitored within the bile ducts using 2D X-ray fluoroscopy. However, contrast agent should remain limited because of the risk of cholangitis. In light of these shortcomings, this preliminary work aims to segment instruments in 2D X-ray fluoroscopy images.

2. Methods

We used a well-known approach in biomedical imaging called U-Net [1] to segment the endoscope and the guidewire. To cope with the difficulties related to the topology of the guidewire which is long and thin, the low contrast and signal to noise ratio, we propose specific pre- and post-processing steps to enhance this deep learning method.
Endoscope segmentation. We followed the same steps of pre-processing as per the nnU-Net [2] approach, consisting in cropping to keep non-zero regions only, normalizing intensity and data augmentation. The image database contained 300 annotated images.

---

[1] Corresponding Author, Garance MARTIN; E-mail: garance.martin@lip6.fr
[2] Images were provided by the Gastroenterology and Endoscopy Division of Hôpital Saint-Antoine, Sorbonne Université, Paris, France

Data augmentation allowed for the creation of 3200 additional ones. We applied elastic transformations in a limited range to keep medically plausible images. A first dataset of 2100 images was used for training, a second of 700 images for validation, and a third of 700 images for test.

Training parameters. We trained a generic improved U-Net with a leaky ReLU, Batch normalization and a loss function combining Binary Cross-entropy and Dice loss.

Post-processing. We applied a post-processing to remove small artifacts from the network prediction. The post-processing consisted of a morphological opening with a structuring element of 10x10 pixels followed by a morphological closing by a structuring element of 70x70 pixels.

Guidewire segmentation. Some additional steps were added to the preprocessing. We enhanced the guidewire with a morphological erosion and dilated the binary mask (annotation) with a structuring element of 3x3 pixels. Then, we generated 1000 images from the 100 previously annotated images. Finally, we included a regularization term in the loss function to overcome large class imbalance. This term is based on the percentage of the guidewire in the images (less than 1%), as learned from the database, and constrains the model to prioritize guidewire pixels instead of background.

3. Results

The Dice score was used to evaluate the endoscope segmentation. After post- processing, an average Dice score of 0.94 0.05 was obtained. Furthermore, we applied segmentation activation map based on gradient [3] to visualize which pixels are used by the network during segmentation. The results clearly demonstrate that pixels from endoscope only are used for the prediction making. To evaluate the segmentation of the guidewire we used a stricter metric, the Hausdorff Distance (HD) which is more adapted to long, thin objects. The smaller the HD the better the segmentation. An average HD of 24.26 pixels for native images of size 1024x1024 pixels was obtained. Qualitative results are available at this link: https://github.com/garancemartin/mie_results.

4. Discussion & conclusion

The endoscope segmentation is very promising. More annotated data would certainly improve this result. On the other hand, although the proposed additional steps improve the performances on the guidewire segmentation, the results do not completely meet expectations for now.

In future work we plan to introduce hybrid AI, combining knowledge not only on the guidewire but also on the structure of the biliary ducts.

References

[1]   Ronneberger O, Fischer P, Brox T. U-Net: Convolutional networks for biomedical image segmentation. In: MICCAI; 2015. p. 234-41.

[2]   Isensee F, Jaeger PF, Kohl SAA, Petersen J, Maier-Hein KH. nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. Nature Methods;18(2):203-11.

[3]   Vinogradova K, Dibrov A, Myers G. Towards Interpretable Semantic Segmentation via Gradient-weighted Class Activation Mapping. In: AAAI; 2020.