**ABSTRACT**

The task of lumen segmentation is a fundamental part in the development of computer vision methods for surgical assistance, since this is the reference which marks the path the surgeons should follow with the endoscope in minimal invasive procedures. We study the implementation of different models of Deep Neural Network for the task of lumen segmentation. For the training of these networks, we analyze the use of two different color spaces: gray-scale and RGB data images and also converting the images to polar coordinates.

**Keywords:** deep learning, uroscopy, convolutional neural networks, lumen segmentation.

**MATERIALS & METHODS**

For this study 7 videos were collected and from them selected frames were manually annotated. The videos were acquired from the European Institute of Oncology (IEO) at Milan, Italy. The loss function used in this implementation was based on the dice similarity coefficient ($DSC$), and the $DSC$ as the performance metric to compare the models. The Kruskal-Wallis test on the $DSC$ was used to determine statistical significance between the different models trained.

**RESULTS 1**

In a second stage the model was compared with the Mask-RCNN architecture. Additionally, the images went through a polar transformation before being feed to the networks.

- In general it was observed that the proposed model performs better than Mask-RCNN
- In general the transformation of the frames to polar coordinates did not improve the performance of any of the networks tested.

**CONCLUSION**

- Residual U-Net has the best performance overall in $DSC$, $Pree$ and $Rec$ with average values of 0.73, 0.58, and 0.92 in comparison with the FCN-8 and the standard version of U-Net.
- Changing the image data to polar coordinates in general did not show improvements in the segmentation obtaining lower scores.
- The use of the Mask-RCNN in general was more complicated, so further research needs to be carried out in order to find the optimal hyperparameters before discarding the use of this architecture for lumen segmentation.

**REFERENCES**


**FUTURE RESEARCH**

Future work includes the exploitation of temporal features embedded in the collected videos along with the collection of more data. The exploration of other architectures to better tackle the segmentation of the lumen and specifically the border regions in which is hard to define where the lumen begins. There is also the possibility of adding extra annotated frames so temporal features could be exploited.

**CONTACT INFORMATION**

Email jorgefrancisco.lazo@polimi.it
Phone +39 334 2427 146