Oral presentation

Comparison between traditional and deep learning-based semi-automatic segmentation methods for metastatic breast cancer lesions monitoring

Noémie Moreau[1,2], Caroline Rousseau[3,4] MD PhD, Ludovic Ferrer[3,4] PhD, Mario Campone[3,4] MD PhD, Mathilde Colombié[4] MD, Nicolas Normand[1] PhD and Mathieu Rubeaux[2] PhD
[1] *LS2N, Centrale Nantes, Nantes, France*
[2] *Keosys, Saint Herblain, France*
[3] *University of Nantes, CRCINA, INSERM UMR1232, CNRS-ERL6001, Nantes, France*
[4] *ICO Gauducheau Cancer Center, Saint Herblain, France*
noemie.moreau@keosys.co

Acknowledgments. This work is partially financed through" Programme opérationnel regional FEDER-FSE Pays de la Loire 2014-2020" noPL0015129 (EPICURE) and by the SIRIC ILIAD Nantes-Angers-INCA-DGOS-Inserm 12558 grant.

Hypothesis
In the context of metastatic breast cancer disease management, the lesion sites can be monitored on PET/CT images using manual or semi-automatic segmentation to initiate more precise characterizations, which is time consuming and error prone. In a first step towards automating segmentation, we compare the performances of semi-automatic traditional and deep learning-based segmentation methods.
Methods
308 metastatic lesions from 16 patients of the EPICURE$_{seinmeta}$ study were manually delineated by an ICO nuclear medicine physician using the Keosys viewer. The SUV$_{max}$ seed point was extracted from these delineations to mimic a one click user initialization, and was employed to automatically build a bounding box around each lesion.
5 traditional region growing-based algorithms (Nestle, 41% SUV$_{max}$, Daisne, Black and a STAPLE of the 4 previous methods) were compared against a 3-fold cross validated deep learning segmentation method based on U-Net. The bounding boxes defined earlier were used i) to contain the spread of the traditional methods and ii) to define patches during the training of the deep learning algorithms. All methods were applied to the SUV-converted PET data.
Segmentation performances were evaluated calculating the Dice score between the semi-automatic segmentations and ground-truth expert manual segmentations.
Results
The 5 traditional methods give very heterogeneous results depending on the location and contrast of the lesions, with a Dice score of 0.33 ± 0.24, 0.12 ± 0.07, 0.20 ± 0.15, 0.19 ± 0.13, 0.16 ± 0.14 for Nestle, Daisne, Black, 41% SUV$_{max}$ and STAPLE methods, respectively. On the other hand, the deep learning approach obtained a Dice score of 0.48 ± 0.19.
Conclusion
Using a deep learning approach to segment cancerous lesions semi-automatically improves the performances against traditional methods. These first results in a difficult clinical context are promising, since they could be improved, particularly by considering the anatomic site of the different lesions.