Automated PET lesion segmentation

Zacharia Mesbah[1,2,4], Romain Modzelewski[2], and Sebastien Thureau[3]

[1] INSA Rouen Normandie, Univ Rouen Normandie, Université Le Havre Normandie, Normandie Univ, LITIS UR 4108, F-76000 Rouen, France
[2] Nuclear Medicine Department, Henri Becquerel Cancer Center, Rouen, France
[3] Radiotherapy Department, Henri Becquerel Cancer Center, Rouen, France
[4] Siemens Healthineers
zacharia.mesbah@insa-rouen.fr

Abstract. Automated Positon Emission Tomography (PET) scan segmentation is a key component for medical images analysis in oncology. Obtaining a tool capable to achieve whole body PET lesion segmentation would be of significant importance in the nuclear medicine workflow.
The AUTOPET II competition is a follow up to last year's AutoPET competition. The goal is to train an algorithm to delineate cancerous lesions in PET/CT imaging. This should be done in the paradigm of dense semantic segmentation. Our method makes use of state-of-the-art Deep Learning framework nnUNet[1] to train a deep neural network from the UNet family to perform semantic segmentation.

Keywords: PET · Functional Imaging · Segmentation · Deep Learning

1 Introduction

^{18}F-FDG (Fluoro-DesoxyGlucose) PET imaging is the key medical imaging modality in cancer detection, as its usage has been growing for the past decades. So much so that it became a must have for oncology and cancer research centers [2]. It is performed at the same time as a Computed Tomography (CT), an anatomical imaging modality used for attenuation correction among other things.

Due to the increasing availability of such functional imaging, quantitative analysis of these volumes rapidly came into discussion. As it is a representation of the body's activity, algorithms were developed to extract information that could serve diagnostics, prognostic and patient treatment in oncology.

The first challenge for these algorithms is to differentiate between normal and abnormal FDG uptake in the body. Indeed, organ such as as the brain, kidneys or the liver have naturally high glucose consumption, which is also the case for cancerous lesions. Organs such as the bladder can also appear as having high FDG uptake since the radio-tracer is evacuated from the body through it.

Once we are capable of isolating cancer-related high uptake areas in the volume, we can study the relationship between the characteristics of these lesions and patient outcome,through the use of radiomics for example[3].

1.1 Competition Task

The AutoPET II competition aims, just as its predecessor did, to build an algorithm capable of delineating cancerous lesions in 18F-FDG PET whole body scans. To achieve this, participants are given 1014 PET/CT studies, gathered from two German medical centers:

- University Hospital of Tübingen
- University Hospital of the LMU in Munich)

Only 900 patients are present in this training set, meaning that some patients might appear on multiple studies.

For each of these studies, two modalities are available[5]:

- The Computed Tomography, resampled towards the PET spacing
- The PET, for which voxel values have been converted to Standardized Uptake Value.

Studies selected were from patients with lung cancer, lymphoma, melanoma or healthy patients. Each of the studies were first screened for cancerous lesions, then if FDG avid lesions existed, they were delineated by two radiologists, one from each hospitals.

Dataset The dataset is split in three parts:

- The training set, that can be used by all contestants to build and train their algorithm. It corresponds to the 1014 studies presented above.
- The preliminary test set, containing 5 sample studies. Considering the statistical non-significance of this set, it should be used to make sure no catastrophic failure happens in the algorithm execution.
- The final testing set, with 120 sample studies. This is the scoring set which will be used for the competition ranking.

	Training set	**Valid**	**Test**
No. of Scans	1014	5	120
No. of patients	900	-	-

Table 1: Dataset Distribution: train / test split

[5] The original, not pre-processed images were also made available by the organizers

Evaluation Dice Score will be used for geometrical comparison, as well as False Positive Value (FPV) and False Negative Value (FPN).

Standard deviation will also be taken into account, thus assessing the robustness and stability off the model. Due to the epsilon values in the dice score, null examples (studies wih no segmentation) will yield a Dice Score of zero (0).

1.2 No-New-Unet (nn-Unet)

nnUNet is a deep neural network training framework, allowing anyone to train out-of-the-box segmentation networks, in two and three dimensions, with dense or cascade resolutions. Networks trained wit this network are considered state-of-the-art today, often doing better even than transformers based segmentation solutions.

It is considered a must have baseline for any research conducted on semantic segmentation, in the medical imaging domain mainly. It is also the technique used in most winning submissions to MICCAI and other challenges based around segmentation.

2 Method

Our method simply consists in adapting the hyper parameters initially obtained by nnUNet's dataset fingerprinting and planning functionality

We start by using nnUNet's *plan_and_preproces* functionnality on the AutoPET II dataset, formatted to the Medical Segmentation Dectathlon format (i.e.: nnUNet input format). This gives us plan files, containing hyper parameters for:

- The data pre-processing
- The neural network configuration
- The network training stage.

We then modified two hyperparameters in the training configuration:

- The number of iterations per epoch: initially set to 250, this would mean the model sees only 500 3D patches (because batch size is 2) each epoch. We set it to 1250, rising the number of examples per epoch to 2500
- The number of validation iterations per epoch, rising it from 50 to 250 iterations per epoch, for similar reason: having a more meaningful validation loss // metric throughout the training.

We then trained a Residual Encoder UNet in the 3D full resolution setting,, using both the CT and PET volumes as channel at indices 0 and 1 respectively.

3 Results

The results displayed below are the ones obtained on the 5 studies dataset preliminary test set.

As mentioned above, the Dice Score is dragged down by healthy studies. The Dice Scores obtained on the three non healthy patients were: **.908, .898** and **.937**. If this level of performance remains on the final testing set, we could consider the model as reliable for this segmentation task.

	Dice Score	**FNV**	**FPV**	**Dice std.**	**FPV std.**	**FNV std.**
Ours	0.5488	0.9256	0.4529	0.5012	0.9919	1.3438

Table 2: Results of our model on the 5-studies preliminary test set

Our model obtained results that situated us middle of the pack on the preliminary test set, which does give us an dea of what our final result will be.

4 Conclusion

The key takeaway from this challenge is that a dataset of such dimensions makes the training of highly reliable PET lesions segmentation deep learning based tools accessible to researchers in cancer research centers. We will conduct future work, as well as review the others participants' methods to attempt developing a robust PET segmentation model.

Such a model would take (part of) the burden of PET segmentation off the shoulders of physicians, in turn freeing their time to conduct more impactful work.

5 Acknowledgement

We would like to express our gratitude to Fabian Isensee for his assistance with the Docker submission using nnunetv2.

References

1. Fabian Isensee, Jens Petersen, Andre Klein, David Zimmerer, Paul F. Jaeger, Simon Kohl, Jakob Wasserthal, Gregor Koehler, Tobias Norajitra, Sebastian Wirkert, and Klaus H. Maier-Hein. nnu-net: Self-adapting framework for u-net-based medical image segmentation, 2018.
2. Rachel Bar-Shalom, Ana Y. Valdivia, and M. Donald Blaufox. Pet imaging in oncology. *Seminars in Nuclear Medicine*, 30(3):150–185, 2000. Entering a New Millennium.
3. Gary J.R. Cook, Gurdip Azad, Kasia Owczarczyk, Musib Siddique, and Vicky Goh. Challenges and promises of pet radiomics. *International Journal of Radiation Oncology*Biology*Physics*, 102(4):1083–1089, 2018. Imaging in Radiation Oncology.