Title:
HEad and neCK TumOR segmentation and outcome prediction using AI: lessons from three consecutive years of the HECKTOR challenge

Authors:
Vincent Andrearczyk, Valentin Oreiller, Mario Jreige, Sarah Boughdad, Catherine Cheze-Le Rest, John O. Prior, Mathieu Hatt, Adrien Depeursinge

Background and aim
Automatic segmentation and radiomics on 2-[^{18}F]fluoro-2-deoxyglucose Positron Emission Tomography (FDG-PET) / Computed Tomography (CT) images and clinical data can contribute to optimizing patient management for head and neck cancer. We present the main outcomes after three editions (2020/2021/2022) of the HECKTOR-MICCAI challenge and discuss the future goals.

Patients/materials and Methods
We collected FDG-PET/CT images with clinical data from 5/6/9 centers in 2020/2021/2022, totaling 254/325/883 cases. Primary tumors and metastatic lymph nodes were delineated by experts according to established guidelines. The HECKTOR challenge was organized at the MICCAI 2020, 2021 and 2022 conferences to evaluate algorithms on tumor/lymph node segmentation as well as prediction of outcome Recurrence- or Progression-Free Survival (RFS/PFS).

Results
Simple, well-designed 3D U-Nets obtained the best segmentation results (Dice coefficients of 0.76/0.78 in 2020/2021 and aggregated Dice coefficient of 0.79 in 2022). PET images contained the most predictive information with the tumor metabolic uptake acting as tumor detection. In association with CT providing additional tissue characteristics, multi-modal models provided the best results. Some nodal metastases were hard to differentiate from primary tumors. A high inter-algorithm agreement was obtained, with some cases incorrectly segmented by all algorithms. For the RFS/PFS prediction, approaches relying on deep learning and/or standard radiomics were proposed. In 2021, best performance (C-index of 0.72) was achieved without using the expert contours.

Conclusion and Discussion
The growing interest in the challenge is matched by the growth of the consortium and dataset. We conducted various post-challenge analyses, leading the path for new challenge designs and tasks. Next directions include the use of other image modalities and refinement of patient subpopulations in terms of Human Papilloma Virus (HPV) status as well as studying the influence of imaging protocol diversity on algorithms' performances.