Abstract: Multi-dataset Approach to Medical Image Segmentation
MultiTalent

Constantin Ulrich[1,4,5], Fabian Isensee[1,2], Tassilo Wald[1,2], Maximilian Zenk[1,5], Michael Baumgartner[1,2,6], Klaus H. Maier-Hein[1,3]

[1]Division of Medical Image Computing, German Cancer Research Center (DKFZ), Heidelberg, Germany
[2]Helmholtz Imaging, DKFZ, Heidelberg, Germany
[3]Pattern Analysis and Learning Group, Department of Radiation Oncology, Heidelberg University Hospital, Heidelberg, Germany
[4]National Center for Tumor Diseases (NCT), NCT Heidelberg, A partnership between DKFZ and University Medical Center Heidelberg
[5]Medical Faculty Heidelberg, University of Heidelberg, Heidelberg, Germany
[6]Faculty of Mathematics and Computer Science, Heidelberg University, Germany
constantin.ulrich@dkfz-heidelberg.de

The medical imaging community generates a wealth of data-sets, many of which are openly accessible and annotated for specific diseases and tasks such as multi-organ or lesion segmentation. Current practices continue to limit model training and supervised pre-training to one or a few similar datasets, neglecting the synergistic potential of other available annotated data. We propose MultiTalent, a method that leverages multiple CT datasets with diverse and conflicting class definitions to train a single model for a comprehensive structure segmentation [1]. Our results demonstrate improved segmentation performance compared to previous related approaches, systematically, also compared to single-dataset training using state-of-the-art methods, especially for lesion segmentation and other challenging structures. We show that MultiTalent also represents a powerful foundation model that offers a superior pre-training for various segmentation tasks compared to commonly used supervised or unsupervised pre-training baselines. Our findings offer a new direction for the medical imaging community to effectively utilize the wealth of available data for improved segmentation performance. The code and model weights are publicly available: `https://github.com/MIC-DKFZ/MultiTalent`.

References

1. Ulrich C, Isensee F, Wald T, Zenk M, Baumgartner M, Maier-Hein KH. MultiTalent: a multi-dataset approach to medical image segmentation. Proc MICCAI. 2023.