10.1148/ryai.2019180019

Appendix E1

Imaging Studies

Mount Sinai Health System (MSHS)

Imaging was performed on seven different scanners: Force, Definition AS, Definition AS + and Definition Flash of Somatom series by Siemens (Munich, Germany) and Discovery CT750 HD, LightSpeed VCT, and Revolution CT by General Electric (Chicago, Ill). Acquisition parameters varied across imaging studies and occasionally included pelvis and chest body regions in addition to the abdomen.

Liver Tumor Segmentation (LiTS) challenge

The freely available LiTS dataset sponsored by the Medical Image Computing and Computer Assisted Intervention (MICCAI) 2017 Annual Conference was incorporated into this study. The LiTS dataset contains 131 portal venous phase CTs that include the abdomen with paired whole liver segmentation masks. Data are collected from six different multiple institutions with heterogeneous acquisition parameters including in-plane resolution ranges from 0.60 mm to 0.98 mm and section spacing from 0.45 mm to 5.0 mm. The axial sections of all scans have an identical size of 512×512, but section numbers ranged between 42 and 1026. Chest and pelvis body regions were included in many studies. Tumor masks were provided in the challenge but not used as part of this study.

Natural language processing, preprocessing, and computer vision pipelines

Once scans were selected using aforementioned natural language processing (NLP) methods, corresponding study accession identifiers enabled transfer of DICOM data to a dedicated high-performance computer. Conversion of triphasic CT abdomen DICOM data into NIFTI format was performed using NiPype (25) and portal venous phase series were selected using DICOM metadata tags. All processing and training of data was performed using OpenSource Python and Tensorflow code adapted from Çiçek et al (26) (https://github.com/ellisdg/3DUnetCNN). Preprocessing was modified significantly given variable body region inclusion on our scans and differences in imaging modality, thus, we incorporated customized automatic centering, cropping, and rescaling. Whole images were resampled to an $128 \times 128 \times 128$ matrix array. A 3D U-Net identical to the one used by ÇiÇek et al was constructed for segmentation purposes (26). In brief, the 3D U-Net operates directly on the 4D volumes (three spatial dimensions and a fourth for sequences) to segment the underlying structures. The U-Net is similar to classical fully convolutional network architectures, but employs skip connections and an encoder-decoder structure that encodes the original volume into a latent space and then decodes the space into the final segmentation. Training hyperparameters included pool size $2 \times 2 \times 2$, batch size 1, 100 training epochs, decay learning rate every 10 epochs with an initial learning rate of 0.00001 and learning rate drop of 0.5. Average inference time for all trained models was 0.86 second per dataset. Calculations of Dice scoring are well described and were implemented in Python along with liver volume measurements.

Figure E3 shows effect per study with active learning training.

References

25. Gorgolewski K, Burns CD, Madison C, et al. Nipype: a flexible, lightweight and extensible neuroimaging data processing framework in python. Front Neuroinform 2011;5:13 .

26. Çiçek Ö, Abdulkadir A, Lienkamp SS, Brox T, Ronneberger O. 3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation. arXiv [cs.CV] (2016). https://arxiv.org/abs/1606.06650.