

SUPPLEMENTARY MATERIALS

Development of a Deep Learning System

Development Dataset

The development dataset consisted of contrast agent-enhanced portal venous-phase liver computed tomography (CT) images from 813 patients. To develop a robust algorithm that performs well in various liver conditions, the development dataset included CT data from healthy livers, livers with various diseases, and post-hepatectomy livers. It was obtained in multiple steps from a large cohort of patients with CT data from a previous study (1). First, we included CT data from 45 subjects with a normal liver (i.e., living liver donors with no clinical and pathologic evidence of liver disease), 30 subjects with fatty liver disease (i.e., macro-vesicular steatosis $\geq 33\%$ at pathologic examination), 45 subjects with non-cirrhotic chronic liver disease (i.e., any chronic liver disease with a meta-analysis of histological data in viral hepatitis fibrosis stage of F2 or F3), 75 patients with pathologically proven cirrhosis, and 105 patients who underwent liver transplantation for decompensated cirrhosis. These patients were randomly selected from the development cohort of a previous study including 7461 patients who underwent liver biopsy, liver resection, or transplantation (1) and underwent liver CT within 3 months of liver pathologic examination from 2007 to 2016 at our institution. We additionally included postoperative CT data obtained within one month ($n = 24$) or from one to six months ($n = 24$) after liver resection from 48 patients who underwent liver resection that were randomly selected from the same study cohort as described above. To include CT data acquired for various types of liver diseases in the development dataset, we used the CT data of 465 patients who underwent liver biopsy for suspected or known chronic liver disease from 2007 to 2016 at our institution or from 2014 to 2017 at two tertiary referral hospitals (Inje University Baek Hospital and Hanyang University Hospital). These patients were derived from the test cohort of the previous study (1).

Deep Learning Algorithm

CT images were processed with the min-max normalization that linearly transforms a pixel value (X) to a normalized value (Y) as $Y = (X - \min) / (\max - \min)$ (2). We chose min and max values as 15 Hounsfield units (HU) and 1200 HU, respectively, based on the HU distribution of solid abdominal organs, vessels, bone, and fluid-filled structures on the portal venous phase CT images.

We designed the architecture of our deep learning algorithm (DLA) based on DeepLabV3+ (3), which is a deep convolutional neural network dedicated to a segmentation task using two-dimensional images. Although CT image-based segmentation of the liver and spleen requires three-dimensional spatial information, computing full three-dimensional image data requires a large number of computational resources. Therefore, we modified DeepLabV3+ by applying a 2.5-dimensional input set-up (4) to exploit three-dimensional spatial information on CT images in a time- and resource-efficient manner. The 2.5-dimensional input set-up utilizes the two-dimensional input axis for red, green, and blue (RGB) color images, which consists of a width \times height \times RGB channel and imports three consecutive sections of CT images (i.e., a CT image of interest as well as CT image sections above and below the CT image of interest) in the axis of the RGB channel. With this 2.5-dimensional input set-up, segmentation is performed on a CT image of interest utilizing the three-dimensional spatial information of three consecutive sections of CT images.

Our DLA consists of an encoder and decoder. The encoder part is based on a modified Xception model (5) that includes a series of downsampling layers and an Atrous Separable Spatial Pyramid Pooling (ASSPP) unit. We used the modified Xception model pre-trained with a PASCAL VOC image dataset (<https://github.com/bonlime/keras-deeplab-v3-plus/releases/tag/1.1>). The ASSPP unit performs 1×1 convolution, atrous depth-wise convolution with variable convolution rates, and global average pooling. The atrous depth-wise convolution allows a neural network to capture multi-scale information from each channel. Thus, the combination of a 2.5-dimensional input set-up and the atrous depth-wise convolution creates synergy in capturing three-dimensional spatial information from input images.

For training our deep learning system, we used three-class softmax cross-entropy loss as the main loss function. Since the cross-entropy loss evaluates the class predictions for each image pixel individually and then averages this over all pixels, it may not be robust enough for a segmentation task for images with class imbalance (i.e., the largest number of pixels for background and the smallest number of pixels for spleen). Therefore, to improve the segmentation performance of the algorithm at the boundary of each class, we additionally calculated a masked Dice coefficient loss (6) as the regularization loss for ten pixels from the border of segmented objects. The final DLA was trained for 125 epochs. Algorithm training took approximately 60 GPU hours on a computer equipped with one Nvidia Geforce GTX 1080 Ti (Nvidia, Santa Clara, CA, USA), two Intel Xeon E5-2695 (Intel, Santa Clara, CA, USA), and 256 gigabytes of RAM. The final DLA was implemented to a web-based Digital Imaging and Communications in Medicine viewer system (GoCDSS; SmartCareworks Inc., Seoul, Korea).

REFERENCES

1. Choi KJ, Jang JK, Lee SS, Sung YS, Shim WH, Kim HS, et al. Development and validation of a deep learning system for staging liver fibrosis by using contrast agent-enhanced CT images in the liver. *Radiology* 2018;289:688-697
2. Patro SGK, Sahu KK. Normalization: a preprocessing stage. Cornell University, 2015. Available at: <https://arxiv.org/abs/1503.06462>. Accessed January 31, 2020
3. Chen LC, Zhu Y, Papandreou G, Schroff F, Adam H. *Encoder-decoder with atrous separable convolution for semantic image segmentation*. The European conference on computer vision (ECCV);2018 September 8-14;Munich, Germany
4. Roth HR, Lu L, Seff A, Cherry KM, Hoffman J, Wang S, et al. A new 2.5D representation for lymph node detection using random sets of deep convolutional neural network observations. *Med Image Comput Comput Assist Interv* 2014;17:520-527
5. Chollet F. *Xception: deep learning with depthwise separable convolutions*. The IEEE conference on computer vision and pattern recognition (CVPR);2017 July 21-26;Honolulu, HI, USA
6. Milletari F, Navab N, Ahmadi SA. *V-net: fully convolutional neural networks for volumetric medical image segmentation*. 2016 fourth international conference on 3D vision (3DV);2016 October 25-28;Stanford, CA, USA