Title:

Multi-Planar U-Net

Authors and Institution:

Mathias Perslev, Akshay Pai, Christian Igel, Erik B. Dam
Department of Computer Science, University of Copenhagen

Correspondence email address:

map@di.ku.dk

Segmentation technique utilized (eg. Neural networks, active contour models, multi-atlas, etc):

Fully convolutional neural networks

Was training data augmented with additional data. If so, please describe images/contrasts used, image resolution, etc:

We used only the supplied dataset, but our 'multi-planar' technique involves extensive data augmentation by sampling many 2D image slices at random orientations through the 3D volume. In addition, we apply random elastic deformations to a sampled 2D images with probability 1/3 and likewise weight the loss contribution of such augmented image by 1/3. Please refer to the (Perslev, Dam, Pai, & Igel, 2019) paper for additional details.

Is your output a binary mask or probabilities? If probability, please specify the per-class threshold value:

For each voxel we output a probability for all classes and finally report the class with the highest probability.

Is your technique fully open sourced (code, model weights, etc). If so, please point to location:

Yes, please refer to https://github.com/perslev/MultiPlanarUNet

CNN Parameters

Network Architecture

We utilized a 2D U-Net model architecture described in (Louring Koch, Perslev, Igel, & Brand, 2019; Perslev et al., 2019). The base topology is similar to the original implementation of (Ronneberger, Fischer, & Brox, 2015) with added batch normalization layers intervening every down- or up-convolution block and using nearest neighbor up-sampling.

The 2D slices input to the model are sampled in a 'multi-planar' fashion. Specifically, we randomly select 6 sample planes of random orientation through the 3D image volume and sample 2D isotrophic images randomly across the 6 planes as input to a single channel of a single U-net model. At inference time, the model predicts the full image volume along each of the planes, and a small model (learned from the validation set) merges the 6 volumes into one in a simple, voxel-wise linear combination.

Network Dimensionality
2D

Input patch sizes
384 x 384

Network Depth
4 convolution blocks, each: 2D convolution → 2D convolution → batch norm → 2x2 max pool
Plus, 2x convolution with batch norm in the bottom after the last (4^{th}) convolution block.
The encoded signal then goes through 4 up-convolution blocks.

Feature Maps Per Depth
Encoder: 90 → 181 → 362 → 724 → 1448

Convolutional kernel and stride dimensions:
3x3 kernels, stride 1x1

Network Normalization Methods
Batch normalization, see (Ioffe & Szegedy, 2015).

Upsampling methods
Nearest neighbor up-sampling, see (Odena, Dumoulin, & Olah, 2016).

Image Pre-Processing Technique
Robust scaling (to median 0 and IQR 1) across the full image volume for each image individually.

Transfer Learning
No pretraining.

Weight Initialization
Glorot uniform.

Number of Training Epochs
Infinite, until early stopping condition, see below.

Batch size
16

Loss function
Multi-class cross entropy (no class balancing)

Optimizer utilized
Adam optimizer
Learning rate: 5e-5
Decay 0.0
Beta1: 0.9
Beta2: 0.999
Epsilon: 1.0e-8

Note: Learning rate is reduced by 10 % for every 2 non-improving epochs (as per validation dice performance).

Early stopping?
Patience: 20 epochs
Error delta: 0

Based on validation dice score performance.

Python random seed (if used)
N.A.

GPU model and numbers
1x Tesla V100

Note: Standard 8-12 GB memory GPUs can also run this model.

Multi-GPU training?
No.

Deep learning framework (Pytorch, TF, MATLAB, …)
TensorFlow 1.3

Data precision (FP32, FP16, etc)
FP32

Hyperparameter optimization strategy
None, the network and optimization parameters were applied with no modifications to this challenge. The original architecture and optimization parameters were developed for the 2018 MICCAI Medical Segmentation Decathlon based on experiments on a brain segmentation dataset. Please refer to (Perslev et al., 2019; Simpson et al., 2019).

Additional Information
- We train 7 models on 7 random splits of the combined training and validation data and submit a majority vote segmentation volume over the 14 (predict on both the original and a mirrored version of each knee MRI) produced outputs for each image.
- As we sample images on grids not aligned with the voxel grid, we may sample an infinite number of images, and thus the notion of an epoch is undefined. We define 1 epoch as 2500 sampled 2D image slices.

References

Ioffe, S., & Szegedy, C. (2015, February 1). Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. *Eprint ArXiv:1502.03167*. Retrieved from https://ui.adsabs.harvard.edu/abs/2015arXiv150203167I

Louring Koch, T., Perslev, M., Igel, C., & Brand, S. S. (2019). Accurate Segmentation of Dental Panoramic Radiographs with {U-Nets}. In *IEEE International Symposium on Biomedical Imaging (ISBI)*.

Odena, A., Dumoulin, V., & Olah, C. (2016). Deconvolution and Checkerboard Artifacts. *Distill*. https://doi.org/10.23915/distill.00003

Perslev, M., Dam, E., Pai, A., & Igel, C. (2019). One Network To Segment Them All: A General, Lightweight System for Accurate {3D} Medical Image Segmentation. In *Medical Image Computing and Computer Assisted Intervention (MICCAI)*.

Ronneberger, O., Fischer, P., & Brox, T. (2015, May 1). U-Net: Convolutional Networks for Biomedical Image Segmentation. *Eprint ArXiv:1505.04597*. Retrieved from https://ui.adsabs.harvard.edu/abs/2015arXiv150504597R

Simpson, A. L., Antonelli, M., Bakas, S., Bilello, M., Farahani, K., van Ginneken, B., … Cardoso, M. J. (2019, February 1). A large annotated medical image dataset for the development and evaluation of segmentation algorithms. *Eprint ArXiv:1902.09063*. Retrieved from https://ui.adsabs.harvard.edu/abs/2019arXiv190209063S