Residual Patch-based U-net for Brain Segmentation

Siti Nurbaya Yaakub, Eric Kerfoot, and James Clough

School of Biomedical Engineering and Imaging Sciences, King's College London
siti_nurbaya.yaakub@kcl.ac.uk

Our method for performing brain segmentation was based on a residual U-net [1, 2] deep learning neural network which was trained to predict segmentation maps for 3D patches. The network architecture, Figure 1, consists of blocks of three 3x3x3 convolutions, each followed by a PReLU [3] activation and batch normalisation [4]. In the encoding path, downsampling was done using convolutions with a stride of 2. In the decoding path, the upsampled input was concatenated with the output from the encoding path on the same level. The final convolution in the last layer used a softmax activation to obtain the probabilistic label predictions.

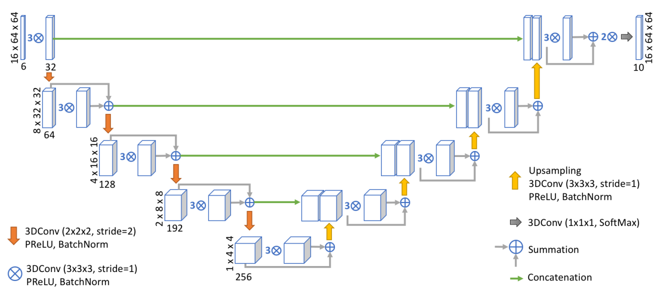

Fig. 1. Residual U-Net architecture.

We used the bias field corrected data for the 3 available imaging modalities (T1, T1-IR, T2-FLAIR) as 3 input features. We trained the network on minibatches containing 32 patches. Patches were chosen by randomly choosing a subject and a 16x64x64 patch from the original 48x240x240 volume. We did not apply weighting to these choices as the number of training examples was sufficiently low that even rare structures would be seen many times by the network during training.

Spatial Information

Similar looking brain regions may need to be classified differently depending on their spatial location in the brain. Our network receives only a small region of the total brain volume and so may struggle to determine the spatial location of a patch in the brain. We therefore also pass as an input the (X,Y,Z) image coordinates of each voxel in the patch as 3 extra input channels. This ensures that spatial context can be utilised to aid in identifying brain regions in the image.

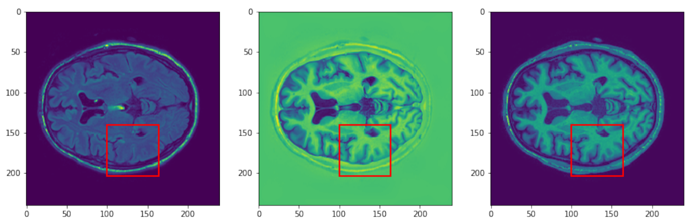

Fig. 2. A typical patch sampled from the input dataset for the three MR input channels.

Data Augmentation

We augmented the original 7 training examples by applying non-rigid free-form deformation [8] to each volume image a set number of times. This produces an expanded training set which includes the original images plus multiple deformed versions of each. Deformation distances are constrained to a prescribed range to ensure the resulting images are not so distorted they represent non-physiological structures nor represent a degradation in image fidelity. Each region category is deformed with the image independently to preserve boundary values between categories; attempting to naively deform a segmentation image will produce interpolation at the boundaries between categories which would represent an inappropriate category label. During training input patches are further augmented by applying a random selection on flipping, mirroring, and rotational transforms.

Training

The network was implemented using the Keras framework with the TensorFlow backend on an nVidia P6000 GPU. We trained our network on 25,000 mini-batches of 32 patches each of size 16x64x64. We used the Adam optimizer with Nesterov momentum [5, 6] with a learning rate of 10^{-5} and the multiclass dice coefficient [7] as the loss. Weights were initialised as in [3].

Inference

Since our network was trained on small 3D patches there were many possible approaches to using it to infer segmentation maps for full brain volumes. Our approach was to make predictions for a series of overlapping 3D patches, then taking a weighted average of the predicted probabilities for each category. The weighting on this average was chosen to favour predictions from the centre of the patch rather than those at the edges, as we found that predictions at the edge of the patch were less reliable.

References

1. Ronneberger, O., Fischer, P., Brox, T.: U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Medical Image Computing and Computer-Assisted Intervention MICCAI 2015, pp.234-241. Springer (2015)
2. He, K., Zhang, X., Ren, S., Sun, J.: Deep Residual Learning for Image Recognition. arXiv preprint arXiv:1512.03385 (2015)
3. He, K., Zhang, X., Ren, S., Sun, J.: Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 1026-1034 (2015)
4. Ioffe, S. Szegedy, C.: Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. arXiv preprint arXiv:1502.03167 (2015)
5. Dozat, T. Incorporating Nesterov Momentum into Adam. ICLR Workshop (2016)
6. Sutskever, I., Martens, J., Dahl, G., Hinton, G. On the importance of initialization and momentum in deep learning. In: Proceedings of the 30th International Conference on Machine Learning (ICML-13), pp. 1139-1147 (2013)
7. Dice, LR., Measures of the amount of ecologic association between species. Ecology. 1945;26:297302.
8. Simard, P. Y., Steinkraus, D., Platt, J. C., Best Practices for Convolutional Neural Networks Applied to Visual Document Analysis, In: Proceedings of the Seventh International Conference on Document Analysis and Recognition - Volume 2, 2003