NerveNet: Convolutional Neural Network Architectures for Medical Image Segmentation

Christopher Goldsworthy, c.goldsworthy@mail.utoronto.ca
Karan Grewal, karanraj.grewal@mail.utoronto.ca

1 Introduction

1.1 Background

Advances in medical imaging are aiding in the identification of nerve structures in a patient's X-ray images. Accurately identifying the appropriate nerve structures is imperative to proceeding with pain management catheters. We explore ultrasound images of the neck and attempt to segment out nerve structures from a large online dataset.

1.2 Convolutional Neural Networks

We used Convolutional Neural Networks for the task of segmenting out nerve structures. In computer vision research, CNNs are used extensively for image classification. The ImageNet database [1] contains over 14 million arbitrary labelled images used by researchers at Stanford University to train CNNs. However, for our image segmentation task, we constructed networks to perform regression as opposed to classification.

We followed a technique led by researchers at the University of California, Berkeley to transform CNNs designed for image classification into *fully convolutional networks* [2] which were able to segment medical images.

Architecture	Conv. Layers	Learning Rate
U-Net [3]	18	1.0×10^{-5}
FCN-AlexNet [4]	7	1.0×10^{-3}
FCN-VGG16 [5]	16	1.0×10^{-4}
FCN-LeNet5 [6]	4	1.0×10^{-3}

Table 1. The various network architectures that we used.

All our CNNs were implemented using Keras [7], a python machine learning library.

2 Experimentation

2.1 Training

We used a g2.2xlarge GPU instance hosted on Amazon Web Services to train our CNNs.

Each CNN was trained using stochastic gradient descent with momentum 0.9 and the specified learning rate from **Table 1**. We used *Cross Validation* to test our model against 10% of the training images after each epoch. In total, each CNN was trained for 50 epochs on a batch size of 20 images.

The training loss was measured by the *Sørenson-Dice Coefficient*, commonly used in statistics to measure the similarity between two samples of data. In this experiment, the Dice Coefficient recorded the similarity in Run-length encoding between the network's prediction on the validation data and the desired segmented output image.

2.2 Results

U-Net was the only CNN used in this experiment that was specifically designed for the purpose of segmenting medical images. Each of the other three networks that we used were classifcation networks that we mutated for this experiment.

Architecture	Dice Coeff.
U-Net	0.57
FCN-AlexNet	0.55
FCN-VGG16	0.37
FCN-LeNet5	0.53

Table 2. The best scores achieved by each CNN over 50 epochs as measured by the Dice Coefficient.

U-Net, FCN-AlexNet and FCN-LeNet5 all had scores within a tight range, however FCN-VGG16 performed far worse.

2.3 Conclusion

Based on our results in **Table 2**, we discover an extremely weak correlation between the number of convolutional layers in a network and its prediction accuracy. Instead, we take away that the network architecture plays a large role in properly segmenting out nerve structures.

References

[1] J. Deng, W. Dong, R. Socher, L. Li, K. Li and L. Fei-Fei. ImageNet: A Large-Scale Hierarchical Image Database. In *CVPR*, 2009.

[2] J. Long, E. Shelhamer and T. Darrell. Fully Convolutional Networks for Semantic Segmentation. In *CVPR*, 2015.

[3] O. Ronneberger, P. Fischer and T. Brox. U-Net: Convolutional Networks for Biomedical Image Segmentation. arXiv:1505.04597v1 [cs.CV], 2015.

[4] A. Krizhevsky, I. Sutskever and G.E. Hinton. ImageNet Classification with Deep Convolutional Neural Networks. In *NIPS*, 2012.

[5] K. Simonyan and A. Zisserman. Very Deep Convolutional Networks for Large-Scale Image Recognition. In *ICLR*, 2015.

[6] Y. LeCun, L. Bottou, Y. Bengio and P. Hafner. Gradient-Based Learning Applied to Document Recognition. In *Proc. of the IEEE*, 1998.

[7] F. Chollet, Keras. Software available at: `https://keras.io/`.