Applying PyramidLSTM to ATLAS Dataset
CS230-Winter 2018

Kevin Pham*
Department of Electrical Engineering
Stanford University
kpham123@stanford.edu

Abstract

The ATLAS dataset of MRI scans is one of the few publicly available datasets of MRI data and was recently released in February 2018. The release of this dataset presented a good opportunity to research and apply a novel deep learning architecture called Pyramid-LSTM [1] instead of the conventional CNN. The Pyramid-LSTM divides the 3-D image space into 6 tetrahedron subvolumes and then merges the results of each, which are fed into fully-connected layers. However, the high compute power needed to run this LSTM on a large dataset prevented training from being completed in time.

1 Introduction

Healthcare is an emerging class of problems in machine learning but the number of datasets in the public domain is low due to patient confidentiality. ATLAS (Anatomical Tracings of Lesions After Stroke) was released by the USC Stevens Neuroimaging and Informatics Institute to train and test lesion segmentation algorithms and provides a standardized dataset for comparing the performance of different segmentation methods. The input to this algorithm are 189 different 2D grayscale images of size 232x196 which represents a 3D MRI scan of a patient's brain, whereas the training data is labeled with pixels comprising a lesion. On test data, we then use a trained and customized LSTM to output segmented pixels which represent where the lesions in unlabeled data are predicted.

2 Related work

Lesion segmentation is typically done by manually by neuroscientists over several hours but there is growing interest in applying computing power to automate this lengthy process. Towards this end deep learning has only recently penetrated the field of computer vision. Initial approaches used a technique called patch classification. The modern and typical approach to semantic segmentation involves convolutional max-pooled networks after the seminal paper "Fully Convolutional Networks (FCN)" by Long et al. from Berkeley.

Patch classification
Deep Neural Networks by Ciresan et al. [2]

Encoder-decoder
U-Net by Ronneberger et al. [3]

*Use footnote for providing further information about author (webpage, alternative address)—*not* for acknowledging funding agencies.

RefineNet by Lin et al. [4]
Large Kernel Matters by Peng et al. [5]

Dilated/atrous convolutions
Multi-Scale Context Aggregation by Yu and Koltun [6]

3 Dataset and Features

The original dataset comprised 229 separate MRI scans of 220 patients. Since the images came in 189 different 2-d image slices of 232x196, they were combined to form a single 3-D volume of {height=232,width=196,depth=189}. To allow the dataset to be split into pyramids comprising a cube, this volume was padded to a cube of the largest dimension. It is first padded by 1 voxel to allow the pyramid to culminate in a single pixel. It is padded by another 2 voxels to allow a conv2d output layer of 3x3.

4 Methods

I found interest in an alternative model known as the multi-dimensional LSTM, which will process 3D images as an RNN along each of its dimensions. The drawback of multi-dimensional LSTM is that the classical implementation cannot take advantage of GPU parallelism; however a novel implementation known as PyramidLSTM[2] processes the volume along a simplex (e.g. pyramid in the 3-D space) to overcome this drawback and has managed to achieve notable results in recent benchmarks[3]. The model was implemented using tensorflow. [7]

5 Experiments/Results/Discussion

The current model needs to be trained and code is available at: https://github.com/kpham123/atlas.

6 Conclusion/Future Work

7 Contributions

References

[1] M. F. Stollenga, W. Byeon, M. Liwicki, and J. Schmidhuber, "Parallel multi-dimensional lstm, with application to fast biomedical volumetric image segmentation," in *Advances in neural information processing systems*, pp. 2998–3006, 2015.

[2] D. Ciresan, A. Giusti, L. M. Gambardella, and J. Schmidhuber, "Deep neural networks segment neuronal membranes in electron microscopy images," in *Advances in neural information processing systems*, pp. 2843–2851, 2012.

[3] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*, pp. 234–241, Springer, 2015.

[4] G. Lin, A. Milan, C. Shen, and I. Reid, "Refinenet: Multi-path refinement networks with identity mappings for high-resolution semantic segmentation," *arXiv preprint arXiv:1611.06612*, 2016.

[5] C. Peng, X. Zhang, G. Yu, G. Luo, and J. Sun, "Large kernel matters–improve semantic segmentation by global convolutional network," *arXiv preprint arXiv:1703.02719*, 2017.

[6] F. Yu and V. Koltun, "Multi-scale context aggregation by dilated convolutions," *arXiv preprint arXiv:1511.07122*, 2015.

[7] Google, "Tensorflow." `https://github.com/tensorflow/tensorflow`. Accessed on 2018-06-10.