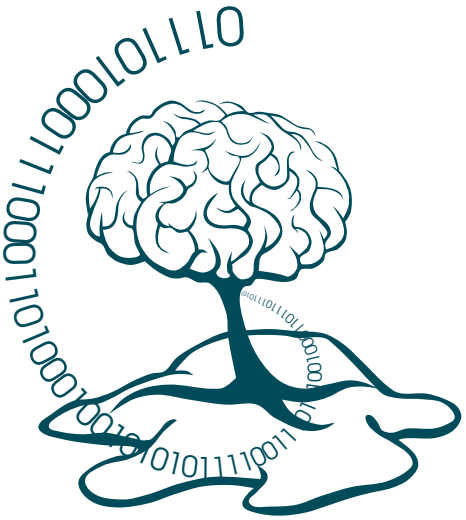

Ischemic Stroke Lesion Segmentation

www.isles-challenge.org

Proceedings

5th October 2015

Munich, Germany

UNIVERSITÄT ZU LÜBECK

Technische Universität München

UNIVERSITÄT BERN

MICCAI'15 MUNICH

Preface

Stroke is the second most frequent cause of death and a major cause of disability in industrial countries. In patients who survive, stroke is generally associated with high socioeconomic costs due to persistent disability. Its most frequent manifestation is the ischemic stroke, whose diagnosis often involves the acquisition of brain magnetic resonance (MR) scans to assess the stroke lesion's presence, location, extent, evolution and other factors. An automated method to locate, segment and quantify the lesion area would support clinicians and researchers alike, rendering their findings more robust and reproducible.

New methods for stroke segmentation are regularly proposed. But, more often than desirable, it is difficult to compare their fitness, as the reported results are obtained on private datasets. Challenges aim to overcome these shortcomings by providing (1) a public dataset that reflects the diversity of the problem and (2) a platform for a fair and direct comparison of methods with suitable evaluation measures. Thus, the scientific progress is promoted.

With ISLES, we provide such a challenge covering ischemic stroke lesion segmentation in multi-spectral MRI data. The task is backed by a well established clinical and research motivation and a large number of already existing methods. Each team may participate in either one or both of two sub-tasks:

SISS Automatic segmentation of ischemic stroke lesion volumes from multi-spectral MRI sequences acquired in the sub-acute stroke development stage.

SPES Automatic segmentation of acute ischemic stroke lesion volumes from multi-spectral MRI sequences for stroke outcome prediction.

The participants downloaded a set of training cases with associated expert segmentations of the stroke lesions to train and evaluate their approach, then submitted a short paper describing their method. After reviewing by the organizers, a total of 17 articles were accepted and compiled into this volume. At the day of the challenge, each teams' results as obtained on an independent test set of cases will be revealed and a ranking of methods established.

For the final ranking and more information, visit WWW.ISLES-CHALLENGE.ORG.

Oskar Maier, Universität zu Lübeck

Mauricio Reyes, University of Bern

Björn Menze, TU Munich

August 2015

Version: 2015-09-30

Organizers

Oskar Maier, Universität zu Lübeck, Germany

Mauricio Reyes, University of Bern, Switzerland

Björn Menze, TU Munich, Germany

Sponsoring Institutions

Institute of Medical Informatics, Universität zu Lübeck, Germany

Institute for Surgical Technology & Biomechanics, University of Bern, Switzerland

Computer Science, TU Munich, Germany

Version: 2015-09-30

A Convolutional Neural Network Approach to Brain Lesion Segmentation

Francis Dutil[1], Mohammad Havaei[1], Chris Pal[2], Hugo Larochelle[1], and Pierre-Marc Jodoin[1]

[1] Université de Sherbrooke, Sherbrooke, Qc, Canada
[2] École Polytechnique de Montréal, Canada

Abstract. Deep Neural Networks (DNNs) are often successful at solving problems for which useful high-level features are not obvious to design. This document presents how DNNs can be used for automatically segment brain lesions for the MICCAI Ischemic Stroke Lesion Segmentation (ISLES) challenge. We experimented several DNN architectures leveraging the recent advances in the field such as convolutional layers, max pooling, maxout units, dropout regularization, and various training strategies.
We present the results of our best performing network on the SISS and SPES training datasets. The results are obtained from the evaluation tool available on the Virtual Skeleton database. As of today, empirical results show that our approach is the most accurate one.

1 Introduction

Brain lesions are abnormalities in the tissue of an organism, usually caused by disease or trauma. The delineation and quantification of brain lesions is critical to establishing patient prognosis, and for understanding the development of pathology over time. Typically, this is performed manually by a medical expert through investigation of several Magnetic Resonance Imaging (MRI) modalities. To alleviate the tedious, time consuming manual delineation, computerised methods can be very useful.

Recently, Convolutional Neural Networks (CNNs) have proven particularly successful in many computer vision applications. For instance, the so-called AlexNet architecture [7] was the first to establish CNNs as the *de facto* state-of-the-art methodology for object recognition in natural images. The main appeal of convolutional networks comes with their end-to-end training nature [6]. That is, their ability of learning low, medium, and high-level features (which involve linear and non-linear operators) as well as the classification function. The potential of CNNs for segmentation in medical imaging however is not well understood, and has only been the subject of preliminary investigations (see workshop publications [3, 10, 9]). In other work [5], alternative to the standard CNN framework have also been explored for more general image segmentation tasks, with the argument that CNN training is overly computationally intensive.

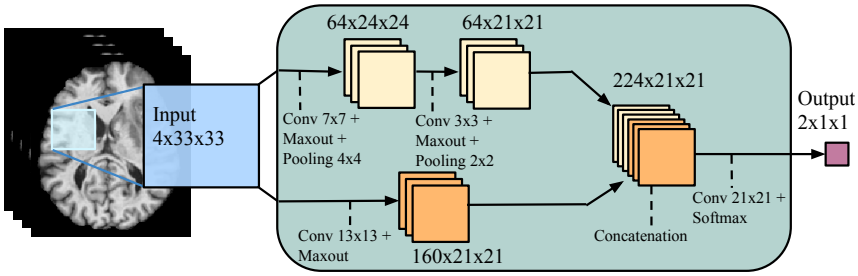

Fig. 1: Our CNN model. The input patch goes through two convolutional networks each comprising of a local and a global path. The feature maps in the local and global paths are shown in yellow and orange respectively.

In this document, we propose a successful, efficient, and automatic CNN architecture for brain lesion segmentation. Note that the proposed CNN is similar to the one used by our team on the 2015 MICCAI brain tumor segmentation (BRATS) challenge. We report results on SPES and SISS training datasets and confirm that our method is leading on both datasets.

2 Convolutional Neural Network Architecture

We approach the problem of brain lesion segmentation by solving it slice by slice, from the axial view. Let $\mathbf{X}$ be one such 2D image (slice), where each pixel is associated with multiple channels, one for each image modality. We treat the problem of segmentation as one of taking any patch it contains and predicting the label of the pixel at its center. The problem is thus converted into an image classification problem.

In the context of this work, we tested a large number of CNN architectures and the most effective one is shown in Figure 1. As can be seen, our method uses a two-pathway architecture in which each pathway is responsible for learning about either the local details or the larger context of tissue appearances (e.g. whether or not it is close to salient regions of the brain like the skull or the CSF). The pathways are joined by concatenating their feature maps immediately before the output layer.

Finally, a prediction of the class label is made by stacking a final output layer, which is fully convolutional to the last convolutional hidden layer. The number of feature maps in the last layer matches the number of class labels and prediction is made with the *softmax* non-linearity.

2.1 Efficient Two-Phase, Patch-Wise Training

By interpreting the output of our CNN as a model for the distribution over segmentation labels, a natural training criteria is to maximize the probability

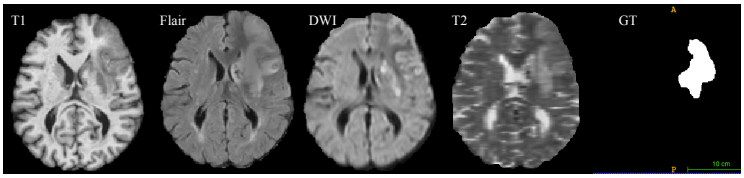

Fig. 2: SISS MRI modalities. The images show the MRI modalities used as input channels to the CNN model for SISS dataset.

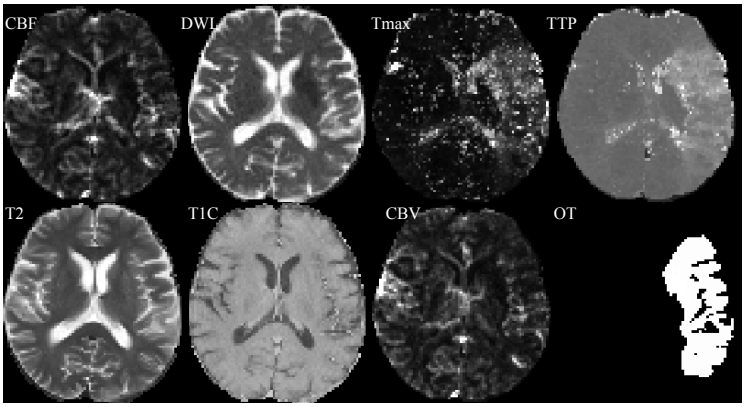

Fig. 3: SPES MRI modalities. The images show the MRI modalities used as input channels to the CNN model for SPES dataset.

of all labels in our training set or, equivalently, to minimize the negative log-probability $-\log p(\mathbf{Y}|\mathbf{X}) = \sum_{ij} -\log p(Y_{ij}|\mathbf{X})$ for each labeled brain. To do this, we follow a stochastic gradient descent approach by repeatedly selecting labels Y_{ij} at a random subset of positions (i.e. patches) within each brain, computing the average negative log-probabilities for this mini-batch of positions and performing a gradient descent step on the CNNs parameters.

Care must be taken however to ensure efficient training. Indeed, since the distribution of labels is very imbalanced (e.g. more than 98% of the brain is healthy), selecting patches from the true distribution would cause the model to be overwhelmed by healthy patches. It is well known that neural network training algorithms such as stochastic gradient descent perform poorly in cases of strong class imbalances. To avoid these issues, we initially construct our patches dataset such that all labels are equiprobable. This is what we call the *first* training phase. Then, in a *second* phase, we account for the unbalanced nature of the data and re-train only the output layer (i.e. keeping the kernels of all other layers fixed) with a more representative distribution over the labels. Using this approach, we were able to fully train CNNs in less than 6 hours.

3 Implementation details

Our implementation is based on the Pylearn2 which supports GPU's and can greatly accelerate the execution of deep learning algorithms [4].

To test the ability of CNNs to learn useful features from scratch, we employed only minimal preprocessing. We applied N4ITK bias correction [2] and clamp the 1% highest intensities to the maximum grayscale value of the 99% remaining pixels as done in [8]. These choices were found to work best in our experiments. The data was normalized within each input channel by subtracting the channel mean and dividing by its standard deviation.

The hyper-parameters of the model (kernel and pooling size for each layer) are illustrated in Figure 1. The learning rate α is decreased by a factor $\gamma = 10^{-1}$ at every epoch. The initial learning rate was set to $\alpha = 0.1$. A post processing method based on connected components was also implemented to remove flat blobs which might appear in the predictions due to bright corners of the brains close to the skull.

4 Experiments and Results

We conducted our experiments on the SISS and SPES datasets. The SISS dataset contains 28 brains with four modalities namely: T1, Flair, Diffusion Weighted Image (DWI) and T1. SPES dataset contains 30 brains with 7 modalities namely: CBF, CBV, DWI, T1c, T2, Tmax and TTP. Both datasets provide pixel-accurate level ground truth of the abnormal areas. Although the ground truth for SPES dataset contains three classes (healthy, stroke, and edema), according to the challenge website the evaluation is done by merging the two unhealthy classes. Figure 3 and Figure 2 show examples from the SPES and SISS datasets.

The virtualskeleton webpage provides a quantitative evaluation of the model [1]. It reports the dice, precision and recall coefficient, as well as the average symmetric surface distance (ASSD) and the Hausdorff distance (HD).

Table 1 and Table 2 show the results obtained from the virtualskeleton webpage on both SISS and SPES datasets and how we compare with other methods applied on these datasets. As one can see, our method (dutif1) is well in front the other methods. Our approach provides the best score on 4 of the 5 metrics on the SISS dataset, and on 3 of the 5 metrics on the SPES dataset. Also, each time our method is not rank first, it is ranked second. Let us underline the fact that since the Hausdorff distances of our method (31.75 and 23.28) is significantly lower than the ones obtained by the other methods, we may conclude that our approach is less prone at detection outliers in the brain.

Figure 4 shows visual segmentation maps produced by our model on both datasets. The first two rows show segmentation results on SPES dataset and the two bottom rows show segmentation results on SISS dataset. It takes on average 25 seconds to produce a segmentation result. The larger receptive field in the two-pathway method allows the model to have more contextual information of the lesion. At the same time, the smaller receptive field make model flexible enough to recognize the fine details of the lesion as opposed to making very smooth segmentation as in the one path method. By allowing for a second phase training and learning from the true class distribution, the model corrects most of the misclassifications produced in the first phase.

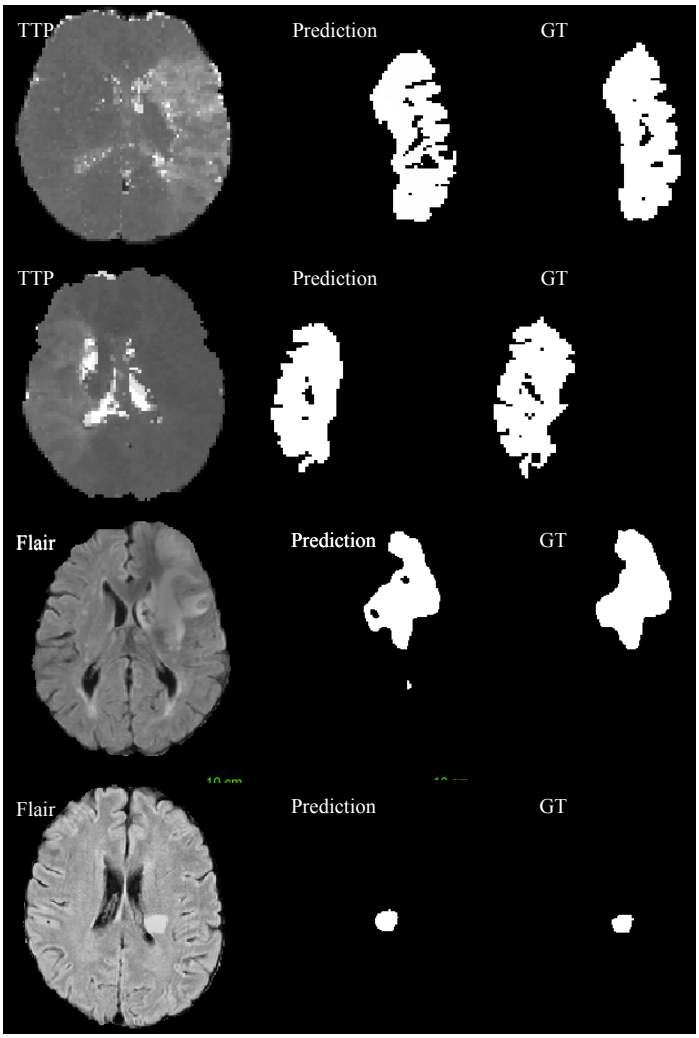

Fig. 4: Results obtained by our model on SPES (top row) and SISS (bottom row) datasets.

Method	ASSD		Dice		Hausdorff Distance		Precision		Recall	
	average	std	average	std	average	std	average	std	average	std
dutif1	8.92	19.23	0.69	0.30	31.75	28.52	0.72	0.31	0.67	0.31
halmh1	6.77	13.17	0.63	0.23	36.16	36.46	0.68	0.24	0.64	0.26
jessa1	11.59	18.34	0.45	0.24	39.23	30.70	0.52	0.26	0.51	0.31
mahmq2	10.30	11.11	0.54	0.26	82.78	23.95	0.67	0.33	0.50	0.25
maieo1	12.36	12.30	0.36	0.25	56.94	40.98	0.65	0.41	0.35	0.21
muscj1	56.77	79.90	0.48	0.38	76.88	81.77	0.57	0.43	0.44	0.37
pinta1	12.18	22.59	0.50	0.31	43.21	30.50	0.61	0.34	0.55	0.33
robbd1	9.36	13.85	0.57	0.28	53.88	34.58	0.58	0.33	0.68	0.21

Table 1: Results on the SISS training dataset showing how our method compares with other methods.

Method	ASSD		Dice		Hausdorff Distance		Precision		Recall	
	average	std	average	std	average	std	average	std	average	std
dutif1	1.76	0.94	0.85	0.08	23.28	14.13	0.83	0.11	0.88	0.08
haect1	3.51	2.13	0.78	0.08	46.31	25.17	0.78	0.11	0.80	0.12
mckir1	1.42	1.01	0.85	0.06	30.71	18.91	0.84	0.10	0.87	0.07
robbd1	2.03	1.35	0.82	0.07	44.29	27.59	0.81	0.14	0.85	0.07

Table 2: Results on the SPES training dataset showing how our method compares with other methods.

5 Conclusion

In this document, we proposed a brain lesion segmentation method based on deep convolutional neural networks. Results on the SISS and SPES datasets reveal that our method is clearly the most accurate one. The high performance is achieved with the help of a novel two-pathway architecture which can model both the local details and global context. Note that the proposed CNN is close to the one used by our team on the 2015 MICCAI brain tumor segmentation (BRATS) challenge. Since there also our approach produced the most accurate results, we are inclined to believe that CNN is a promising technology for brain segmentation applications.

References

1. Virtual skeleton database. http://www.virtualskeleton.ch/
2. Avants, B.B., et al.: Advanced normalization tools (ants). Insight J (2009)
3. Davy, A., et al.: Brain tumor segmentation with deep neural networks. proc of BRATS-MICCAI (2014)
4. Goodfellow, I., et al.: Pylearn2: a machine learning research library. arXiv preprint arXiv:1308.4214 (2013)
5. Huang, G.B., Jain, V.: Deep and wide multiscale recursive networks for robust image labeling. ICLR, arXiv:1310.0354 (2014)
6. Jaderberg, M., Vedaldi, A., Zisserman, A.: Speeding up convolutional neural networks with low rank expansions. In: in proc of BMVC (2014)
7. Krizhevsky, A., et al.: ImageNet classification with deep convolutional neural networks. In: NIPS (2012)
8. Menze, B., et al: The multimodal brain tumor image segmentation benchmark (brats). Medical Imaging (2014)
9. Urban, G., et al.: Multi-modal brain tumor segmentation using deep convolutional neural networks. proc of BRATS-MICCAI (2014)
10. Zikic, D., et al.: Segmentation of brain tumor tissues with convolutional neural networks. proc of BRATS-MICCAI (2014)