Dilated Convolutions for Brain Tumor Segmentation in MRI Scans

Marc Moreno Lopez[1] and Jonathan Ventura[2]

[1] Department of Computer Science
University of Colorado Colorado Springs
mmorenol@uccs.edu
[2] Department of Computer Science
University of Colorado Colorado Springs
jventura@uccs.edu

Abstract. In this work, we present a novel method to segment brain tumors using a dilated Fully Convolutional Network [7]. An accurate brain tumor segmentation is key for a patient to get the right treatment and for the doctor who must perform surgery. Due to the genetic differences that exist in different patients, even between the same kind of tumor, an accurate segmentation is crucial.To evaluate our algorithm we use the evaluation tool from the Brain Tumor Segmentation challenge, BraTS from 2017.

Keywords: Deep learning · Brain tumor segmentation · Dilated convolutions · Fully convolutional network

1 Introduction

Cancer is one of the leading causes of death in the world. In the US, cancer is the 2nd leading cause exceeded only by heart disease [1]. To put it in perspective, one out of every four deaths in the US is caused by cancer. Due to this high-death ratio, scientist all over the world have tried to find a cure for cancer. In this work, my intention is to find a faster and more efficient way to detect cancer in time [8, 11, 12, 13].

According to the National Cancer Institute [2], checking for cancer (or for conditions that may become cancer) in people who have no symptoms is called screening. Screening can help doctors find and treat several types of cancer early. Early detection is important because when abnormal tissue or cancer is found early, it may be easier to treat. By the time symptoms appear, cancer may have begun to spread and is harder to treat. Several screening tests have been shown to detect cancer early and to reduce the chance of dying from that cancer. But it is important to keep in mind that screening tests can have potential harms as well as benefits. Some screening tests may cause bleeding or other health problems.

Screening tests can have false-positive results – the test indicates that cancer may be present even though it is not. False-positive test results can cause

anxiety and are usually followed by additional tests and procedures that also have potential harms. Screening tests can also have false-negative resultsthe test indicates that cancer is not present even though it is. False-negative test results may provide false reassurance, leading to delays in diagnosis and possibly causing an individual to put off seeking medical care even if symptoms develop[14]. Finally, overdiagnosis is also possible. This happens when a screening test correctly shows that a person has cancer, but the cancer is slow growing and would not have harmed that person in his or her lifetime. Treatment of such cancers is called overtreatment.

In this work, we want to combine a screening method, such as Magnetic Resonance Imaging, MRI, with the latest technology on machine learning, deep learning. Deep Learning is a new area of Machine Learning research, which has been introduced with the objective of moving Machine Learning closer to one of its original goals: Artificial Intelligence. There have been some approaches using this method [9] [6], but they havent presented an approach that can beat human performance.

A correct segmentation is key for many reasons. The most important one is so that the patient can get the best possible treatment. An accurate tumor quantification is needed so that the patient gets the amount of treatment that he needs. A rightful segmentation is crucial too in life-threatening cases. These are cases where the tumor is next to or on top of one of the cerebellum, or to similar sensitive parts. Therefore, we need to do a correct segmentation, especially in the boundaries between tumor and edema. This last part is central when planning a brain tumor extraction. Doctors need to know what they are facing before performing any surgery.

One of the main factors for a correct image segmentation is having enough images to train the network [5]. When working with medical images this can become an issue.

We are going to work towards a correct segmentation using a Deep Convolutional Neural Network with dilated filters instead of pooling filters. Moreover, instead of training it with the whole image, we will use a patch-based training approach.

2 Methods

We apply a fully convolutional network approach in order to produce a per-pixel segmentation output. Our network is applied to each slice in a scan separately.

2.1 Dilation

An issue with traditional convolutional neural network architectures that use max pooling is that they downsample the image and thus produce a segmentation with resolution smaller than the input size.

In [15], Yu et al. develop a new convolutional network module that is specifically designed for dense prediction. They present a model that uses dilated

Layer	Type	Configuration	Dilation
1	Convolutional	3x3x1x32	1
2	Batch normalization		
3	ReLU		
4	Convolutional	3x3x32x32	1
5	Batch normalization		
6	ReLU		
7	Convolutional	3x3x32x32	2
8	Batch normalization		
9	ReLU		
10	Convolutional	3x3x32x32	4
11	Batch normalization		
12	ReLU		
13	Convolutional	3x3x32x32	8
14	Batch normalization		
15	ReLU		
16	Convolutional	3x3x32x32	16
17	Batch normalization		
18	ReLU		
19	Convolutional	3x3x32x32	1
20	Batch normalization		
21	ReLU		
22	Convolutional	1x1x32x4	1

Table 1. Configuration of the CNN

convolutions. This model is designed to systematically aggregate multi-scale contextual information without losing resolution. All their work is based on the fact that dilated convolutions support exponential expansion of the receptive field without losing resolution or coverage.

Let $F : \mathbb{Z}^2$ be a discrete function. Let $\Omega_r = [-r, r]^2\mathbb{Z}^2$ and let $k : \Omega_r \to \mathbb{R}$ be a discrete filter of size $(2r+1)^2$. The discrete convolution opreator can be defined as

$$(F * k)(p) = \sum_{s+t=p} F(s)k(t) \tag{1}$$

We now generalize this operator. Let l be a dilation factor and let $*_l$ be defined as

$$(F *_l k)(p) = \sum_{s+lt=p} F(s)k(t) \tag{2}$$

We will refer to $*_l$ as a dilated convolution or and l-dilated convolution.

2.2 Patch-based training

The dataset exhibits severe class imbalance, i.e. the tumor pixels are vastly outnumbered by the non-tumor pixels. This poses a problem when training the network, because the non-tumor pixels influence the total loss function much more strongly than the tumor pixels.

To address this issue, we adopt a patch-based training approach. During training, we randomly sample patches from the images to form a batch. Each patch is exactly the size of receptive field of the network.

We sample patches using a uniform distribution over the classes. In other words, we ensure that each batch has the same number of examples of each class. This effectively remedies the class imbalance that would be caused by simply randomly sampling the patches. Moreover, we don't sample patches from pixels with zero intensity.

Because the network is fully convolutional, we can use the same network trained on patches to test on images. At test time, we use whole images as input to the network to produce the full-resolution output.

3 Results

3.1 Implementation details

To develop the experiment, we used Keras for Python 2.7. All training was done in one machine with Ubuntu and Nvidia Titan X card with 12 GB of memory.

To train the network, we used a batch size of 120 patches, 1000 batches per epoch and 1000 epochs. We used an Adagrad optimizer [4] with a learning rate of 0.01. One training epoch takes 146 seconds.

3.2 Evaluation on validation dataset

To evaluate our algorithm, we used the BRATS evaluation tool. To measure the performance of the algorithms, we have to evaluate recall (or sensitivity), specificity and the dice score (also called F1) [3]. Recall (also known as sensitivity) is the fraction of relevant instances that are retrieved, therefore it measures the proportion of positives that are correctly identified as such. Specificity measures the proportion of negatives that are correctly identified as such. The dice score or F1, is a is a statistic used for comparing the similarity of two samples. It is calculated using the precision and recall parameters.

DICE	Sensitivity	Specificity	Hausdorff95
0.52788	0.68864	0.99495	32.01849

Table 2. Mean results for Enhanced Tumor

DICE	Sensitivity	Specificity	Hausdorff95
0.63685	0.71717	0.90964	36.46434

Table 3. Mean results for Tumor Core

DICE	Sensitivity	Specificity	Hausdorff95
0.73717	0.86329	0.97142	43.38166

Table 4. Mean results for Whole Tumor

In Figures 1, 2 and 3 we have included some examples of segmented MRI. Figure 1 and 2 are good examples of how the segmentation should look like. In FIgure 3 there are some issues with the edema detection that we will discuss later.

4 Discussion

Based on the results that we have obtained, we can say that the balancing of data is key to avoid getting too many pixels labels as background or normal tissue. In Figure 1 and Figure 2, we can see that the segmentation is correct, but we still need to address some issues. We have a balanced algorithm, but we still have room for improvement until we reach human performance.

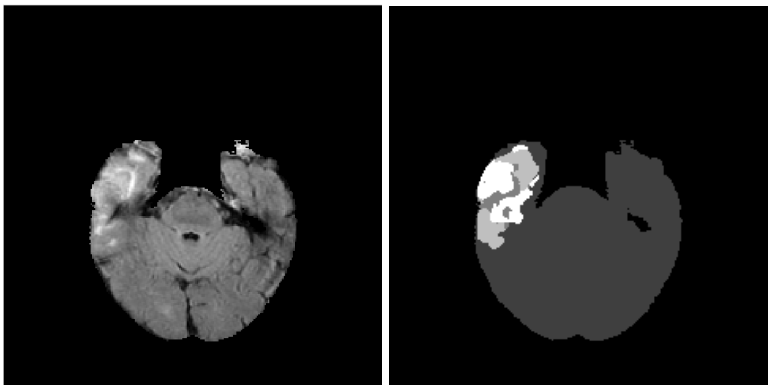

Fig. 1. MRI with its corresponding segmentation

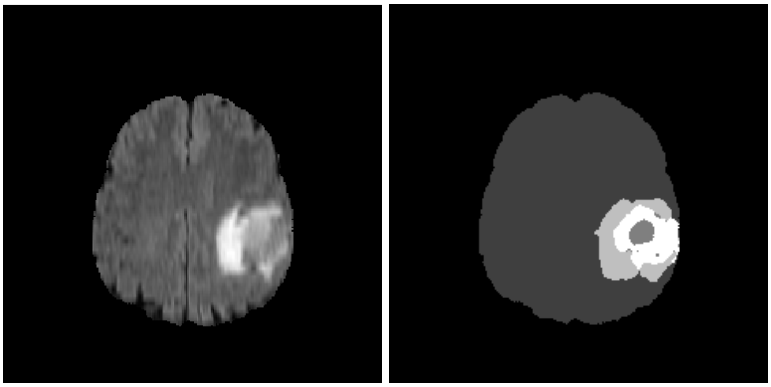

Fig. 2. MRI with its corresponding segmentation

One of our main areas of improvement is edema detection. There are some issues with the detection of edema, since too many pixels are detected as edema and this lowers the precision of the algorithm. This is caused by the similarity of some of the edema pixels with adjacent brain pixels. In Figure 3 we can see a good example of this problem.

As for the tumor pixels, we can see a problem with boundaries or edges. The edges between edema and tumor core are not segmented well enough. Many tumor pixels are labeled as edema. And the same thing happens with enhanced tumor and tumor pixels.

As for enhanced tumor, the algorithm behaves similarly for most of the cases, and it has a regular behavior in most of the images. However, there are some issues with enhanced tumor detection, where the dice score is low. We think that this is due to the low number of enhanced tumor pixels in some of the MRI. In these cases, the MRI contains few Enhanced Tumor pixels (10-20) and our algorithm misses all or most of them.

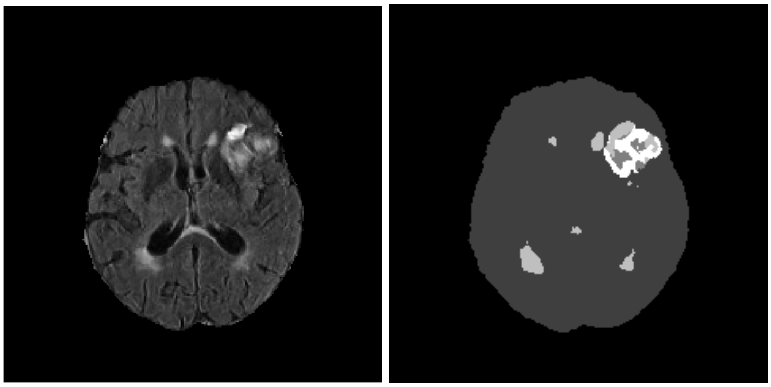

Fig. 3. MRI with its corresponding segmentation

5 Conclusions

Results are encouraging, especially for whole tumor and lead us to believe that, even though we still have work to do until we reach our objective, we can keep working in this direction. The lack of medical data is one of the main problems that we have encountered in this project. To publish medical data, the three parts involved, doctor, institution and patient, must agree to publish it. Therefore, if you arent working with a medical institution, its difficult to obtain lots of medical data or even trustworthy data. As a future work we would like to implement a data augmentation model using elastic image deformation, to overcome this scarcity of data.

Finding the right network configuration to work with such a small amount of data was an enormous challenge. We went through many different networks and different configurations to try to find the most suitable network for our needs. The dilated convolution is effective for brain tumor segmentation to introduce context without losing output spatial resolution and it is an interesting direction to explore in the future. One of the main ideas we want to apply is combining well-known structures like U-Net [10] with the dilation model. Another interesting approach due to its effectiveness for segmentation, is the one proposed by Yu et al. in [16], where they use Dilated Residual Networks for image segmentation.

To increase the accuracy when detecting Enhanced Tumor, we could address this by changing the balance of classes during patch sampling and/or adding class weights.

Brain tumor segmentation isn't an easy task. Due to the genetics of cancer, it remains being a task for which the doctors help is needed. However, they dont have to do all by themselves, since with tools like the one that we have designed in this work, we can help them and we can contribute in the fight against cancer.

References

[1] http://www.who.int/mediacentre/factsheets/fs297/en/.

[2] https://www.cancer.gov/.

[3] Jesse Davis and Mark Goadrich. "The Relationship Between Precision-Recall and ROC Curves". In: *Proceedings of the 23rd International Conference on Machine Learning*. ICML '06. Pittsburgh, Pennsylvania, USA: ACM, 2006, pp. 233–240. ISBN: 1-59593-383-2. DOI: 10.1145/1143844.1143874. URL: http://doi.acm.org/10.1145/1143844.1143874.

[4] John Duchi. "Adaptive Subgradient Methods for Online Learning and Stochastic Optimization". In: *Journal of Machine Learning Research* 12 (2011), pp. 2121–2159.

[5] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. http://www.deeplearningbook.org. MIT Press, 2016.

[6] Mohammad Havaei, Axel Davy, David Warde-Farley, Antoine Biard, Aaron Courville, Yoshua Bengio, Chris Pal, Pierre-Marc Jodoin, and Hugo Larochelle. "Brain tumor segmentation with deep neural networks". In: *Medical Image Analysis* 35 (2017), pp. 18–31. arXiv: arXiv:1505.03540.

[7] Jonathan Long, Evan Shelhamer, and Trevor Darrell. "Fully Convolutional Networks for Semantic Segmentation". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)* (2015). arXiv: arXiv:1605.06211.

[8] Bjoern Menze, Andras Jakab, Stefan Bauer, Jayashree Kalpathy-Cramer, Keyvan Farahani, Justin Kirby, Yuliya Burren, Nicole Porz, Johannes Slotboom, Roland Wiest, Levente Lanczi, Elisabeth Gerstner, Marc-Andre Weber, Tal Arbel, Brian Avants, Nicholas Ayache, Patricia Buendia, Louis Collins, Nicolas Cordier, Jason Corso, Antonio Criminisi, Tilak Das, Herve Delingette, Cagatay Demiralp, Christopher Durst, Michel Dojat, Senan Doyle, Joana Festa, Florence Forbes, Ezequiel Geremia, Ben Glocker, Polina Golland, Xiaotao Guo, Andac Hamamci, Khan Iftekharuddin, Raj Jena, Nigel John, Ender Konukoglu, Danial Lashkari, Jose Antonio Mariz, Raphael Meier, Sergio Pereira, Doina Precup, S. J. Price, Tammy Riklin-Raviv, Syed Reza, Michael Ryan, Lawrence Schwartz, Hoo-Chang Shin, Jamie Shotton, Carlos Silva, Nuno Sousa, Nagesh Subbanna, Gabor Szekely, Thomas Taylor, Owen Thomas, Nicholas Tustison, Gozde Unal, Flor Vasseur, Max Wintermark, Dong Hye Ye, Liang Zhao, Binsheng Zhao, Darko Zikic, Marcel Prastawa, Mauricio Reyes, and Koen Van Leemput. "The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS)". In: *IEEE Transactions on Medical Imaging* 34.10 (2015), pp. 1993–2024.

[9] Sergio Pereira, Adriano Pinto, Victor Alves, and Carlos A Silva. "Brain Tumor Segmentation Using Convolutional Neural Networks in MRI Images". In: *IEEE Transactions on Medical Imaging* 35.5 (2016), pp. 1240–1251.

[10] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. "U-Net: Convolutional Networks for Biomedical Image Segmentation". In: *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015* (2015),

pp. 234–241. ISSN: 16113349. DOI: 10.1007/978-3-319-24574-4_28. arXiv: 1505.04597.

[11] Bakas S, Akbari H, Sotiras A, Rozycki M Bilello M, Kirby J, Freymann J, Farahani K, and Davatzikos C. "Advancing The Cancer Genome Atlas glioma MRI collections with expert segmentation labels and radiomic features". In: *Nature Scientific Data* (2017). [In press].

[12] Bakas S, Akbari H, Sotiras A, Rozycki M Bilello M, Kirby J, Freymann J, Farahani K, and Davatzikos C. "Segmentation Labels and Radiomic Features for the Pre-operative Scans of the TCGA-GBM collection". In: *The Cancer Imaging Archive* (2017). DOI: 10.7937/K9/TCIA.2017.KLXWJJ1Q.

[13] Bakas S, Akbari H, Sotiras A, Rozycki M Bilello M, Kirby J, Freymann J, Farahani K, and Davatzikos C. "Segmentation Labels and Radiomic Features for the Pre-operative Scans of the TCGA-LGG collection". In: *The Cancer Imaging Archive* (2017). DOI: 10.7937/K9/TCIA.2017.GJQ7R0EF.

[14] R. Weinberg. *The Biology of Cancer, Second Edition*. Taylor & Francis Group, 2013. ISBN: 9781317963462. URL: https://books.google.com/books?id=MzMmAgAAQBAJ.

[15] Fisher Yu and Vladen Koltun. "Multi-Scale Context Aggregation by Dilated Convolutions". In: *International Conference on Learning Representations (ICLR)* (2016). arXiv: arXiv:1511.07122v3.

[16] Fisher Yu, Vladlen Koltun, and Thomas Funkhouser. "Dilated Residual Networks". In: *Computer Vision and Pattern Recognition (CVPR)* (2017). arXiv: arXiv:1705.09914.