

# Segmentation of pathological liver tissue with Dilated Fully Convolutional Networks: A Preliminary Study

Jessica C. Delmoral<sup>1,2</sup> Durval C. Costa<sup>3</sup> Diogo Borges<sup>4</sup> João Manuel R. S. Tavares<sup>1,2</sup>

**Abstract**— Early detection of liver cancer, whether from primary occurrence or from metastization is highly important to establish informed treatment decisions. Accurate delineation of the liver tissues of interest facilitates quantitative assessment of the regions of interest, treatment application, and prognosis.

Segmentation of the liver in Computer Tomography (CT) images allows the extraction of the three-dimensional (3D) structure of the liver tissues in which the observation of their relative position to one another is particularly important in treatment scenarios of radiation therapy or interventional surgery planning. The adequate receptive field for the segmentation of such a big organ in CT images, from the remaining neighbouring organs was very successfully improved by the use of the state-of-the-art Convolutional Neural Networks (CNN) algorithms, however, certain issues still arise and are highly dependent of pre- or post- processing methods to refine the final segmentations. Here, the effects of Dilated Convolutional Networks is proposed, for the purpose of improving segmentation of liver tissues in CT. The introduction of a dilation module allowed the concatenation of feature maps with a richer contextual information. The hierarchical learning process given by different dilated convolutional layers is analysed quantitatively. Experiments on the MICCAI Lits challenge dataset are described achieving segmentations with a mean Dice coefficients of 95.57% and 59.36% for the liver and liver tumour, using a total number 30 CT test volumes.

## I. INTRODUCTION

Computed Tomography (CT) is the preferred imaging modality for assessment, detection, diagnosis and follow-up of liver diseases [1]. Automatic segmentation of different medically relevant liver tissues is an active research topic in medical image analysis. Liver segmentation in computed tomography (CT) challenges current computer-aided diagnosis (CAD) systems that aim to support clinicians and improve their diagnosis and treatment performance. Such segmentations can provide doctors with meaningful and reliable quantitative information of the structure of the liver, which subsequently enable the identification of abnormalities. Precise quantitative diagnosis of the liver structure becomes particularly relevant in individuals diagnosed with liver cancer. Liver cancer has an alarming prevalence in a global scale and is the second most lethal cancer worldwide, accountable for more than 788,000 deaths in 2015 [2]. Liver cancer is characterized by the development of abnormal cell accumulations, commonly referred as

lesions that will appear represented differently within the anatomy of the liver, in structural images such as CT. In this clinical scenario, an informed decision on the treatment course entails the assessment of the estimated prognosis and surgical or intra-operative treatment planning. Factors such as tumour precise and relative location, shape and size are taken into account when deciding the type of treatment adopted. Liver cancer treatment may include chemo- or radio- therapy, hepatectomy (liver resection) or in very specific cases transplantation. Manual segmentation, however, varies from practitioner to practitioner, must be performed slice-by-slice, and requires large amounts of time and labour. Performing liver segmentation on CT scans may take up to 90 minutes for a single patient [3]. There is thus a need for reliable automatic or semi-automatic segmentation tools. The segmentation of these structures is not trivial due to the great variability in liver and liver lesions shape, location and boundary definition. Several classes of segmentation methods are proposed in the literature to tackle this subject, such as deformable models, probabilistic atlases, intensity-based methods such as level-sets and region growing. However, the literature was highly enhanced when machine learning algorithms were introduced in segmentation tasks, and more recently the state-of-the-art results were achieved with the application of Fully Convolutional Neural Network (FCNs) methods. CNNs have become appealing to address this task in coming years since (i) they have achieved the most state-of-the-art results in several fields of computer vision and medical image analysis and (ii) they discover classification-suitable representations directly from the input data, unlike conventional machine-learning strategies that are dependent of hand-generated representations of the input data. However, the success of such methods is dependent firstly, of the availability of sufficient amounts of labelled data and secondly, of the generalization capacity and wide-domain capacity of the resulting models. Interest in the specific task of liver analysis in medical images was highly enhanced by the proposal of the 2007 MICCAI Sliver [4] and the 2017 MICCAI Lits [5] challenges, providing a well-established image dataset to benchmark different segmentation algorithms. In the literature 2D and 3D FCNs have been proposed. Ben-Cohen et al. (2016) proposed a VGG-16-based architecture adapted with only convolutional layers, trained on full CT images, using adjacent (above and below) images as input to the model, validated on 43 image slices [6]. Such methodology provides added 3D contextual information to the network. Christ et al. (2017) proposed also a fully convolutional (FCN) method, known as the U-net architecture. The authors propose the addition of a post-processing step using the widely established Conditional Random Field algorithm to further refine the CNN resulting segmentations [7], and validate their method on the 3DIRCAD public dataset. Under the scope of the 2017

<sup>1</sup>Instituto de Ciência e Inovação em Engenharia Mecânica e Industrial, R. Dr. Roberto Frias 400, 4200- Porto, Portugal, email: tavares@fe.up.pt

<sup>2</sup>Faculdade de Engenharia, Universidade do Porto, R. Dr. Roberto Frias, s/n, 4200- Porto, Portugal

<sup>3</sup>Champalimaud Centre for the Unknown, Fundação Champalimaud,

<sup>4</sup>School Of Health Sciences - University of Aveiro, Portugal

MICCAI LiTS challenge, Bi et al. (2017) presented the performance of the deep Residual Network (ResNet) for liver segmentation. Such architectures provide residual skip connections with the objective of avoiding the occurrence of vanishing gradients, especially in very deep architectures [8]. Recognizing the drawbacks of previous methods used for liver segmentation and the great ability of CNNs in learning richer features, we designed our approach to medical image segmentation that leverages the power of convolutional neural networks, trained end-to-end, to process CT volumes. We propose the usage of a Dilated Spatial Pooling module as the encoding portion of the explored CNN on CT slices, in a method that segments liver and lesions sequentially, leading to significantly higher segmentation quality, as demonstrated on a public challenge dataset.

## II. MATERIALS AND METHODS

### A. Dataset

The presented approach was validated in the 2017 MICCAI Lits dataset, consisting of 130 publicly available contrast-enhanced CT scans. The image dataset contains scans acquired from different scanner models, and acquisition protocols, and preserves a common image-resolution of 512x512 pixels, comprising a varying resolution in the axial plane that varies from 94 to 910 image slices, with slice thicknesses ranging from 0.7 to 5 mm. The dataset was divided into a training, validation and test non-overlapping groups composed of 92, 8 and 30 scans, respectively.

### B. Pre-Processing

Image contrast and intensity normalization are fundamental techniques when it comes to the analysis of the intensity content present in abdominal CT. Under this context, abdominal CT images of a patient depict several organs such as liver, spleen, gal bladder, etc. CT images acquisition outputs a quantification of X-rays by tissues at a pixel wise level, which is outputted according to the known scale of Hounsfield units (HU), proportional to the degree of tissue attenuation suffered. Although different HU intervals characterize different organs, these values often overlap, making the discrimination of the present tissues, particularly difficult. To eliminate the noise effect of other ranges of HU

values, a technique named CT windowing is often applied. Thus, all CT slices were thresholded with a window range of  $[-200,250]$  recommended for the liver, removing the irrelevant tissue intensities. We enhanced the contrast of images at a certain level by performing histogram stretching of the obtained pixel values. Ground truth delineations of the liver and liver tumor are available for each exam, used to develop the CNN methods proposed in a supervised manner, and against which the performances of the developed methods were compared.

### C. Segmentation method

This section describes a two-step segmentation framework with no need of user interaction. The hierarchical segmentation approach has two main stages: whole liver binary segmentation, followed by intra hepatic liver tumor segmentation. The approach is cascaded since the first step aims at identifying the complete liver anatomy, which is used in the following step as region of interest. The implemented FCN is based on the well-known U-net architecture, proposed by Ronneberger et al. [7], hence it is comprised of a contracting and expanding paths. However, certain changes in this structure were studied, to analyse the resulting effects in the segmentation results. In the classical U-net implementation, several feature maps are extracted in the contracting portion capturing increasingly more intra-slice contextual information, by pooling of the features extracted. The network is characterized by addition-based connections between the shape-corresponding feature maps of the contracting to the expanding paths. This architecture allows the transfer of finer intensity and boundary details into the deeper operations of the network, which were roughly lost by the previous downsampling of the contracting path.

- Dilated Convolution

Dilated convolutions are proposed in the literature as operations that intend to effectively expand receptive field without losing resolution. Such effect occurs by the usage of standard pooling operations in the encoding path of U-net, as proposed by the authors [7]. However, the pooling operation makes the model invariant to local image transform, and

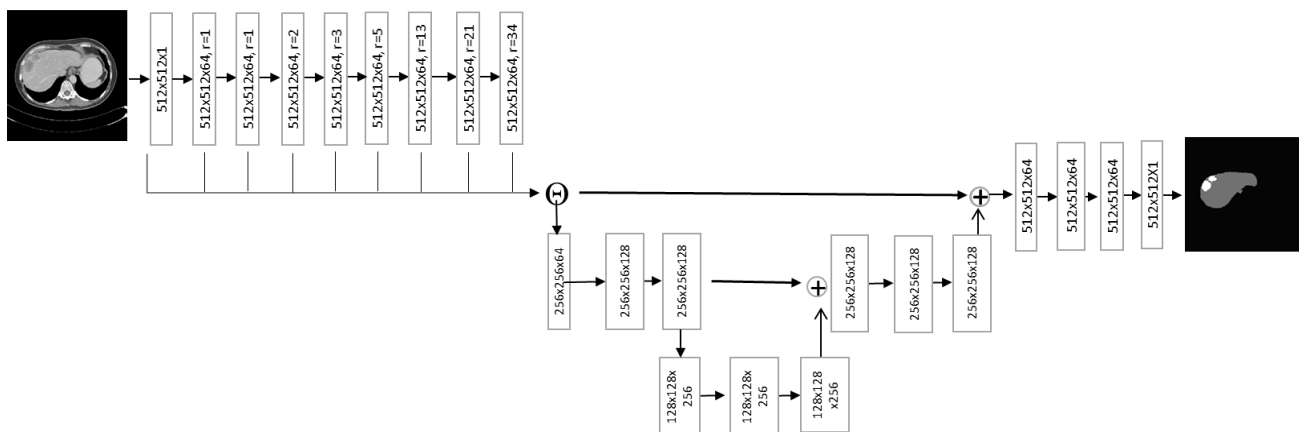


Figure 1. Architecture implemented. Dilated convolutions are indicated with reference to the dilation rate implemented.  $\oplus$  indicates concatenation and  $\oplus$  indicates addition of feature maps.

downsizes the pictures as well as losing voxel-wisely resolution. To tackle such loss of resolution, in the proposed architecture we introduce a first operation of each block to systematically aggregate multi-scale contextual information. Very recently, Yu and Koltun [8] adopted this operation for semantic segmentation to increase the receptive field of deep CNNs, as alternative to down-sampling feature maps. The concept is based on the use of convolution kernels that have been dilated by inserting zero holes between the non-zero positions of a kernel. Given the signal  $x[i]$ , the  $y[i]$  output of a dilated convolution with the dilation rate  $r$  and a filter  $w[s]$  with size  $S$  is formulated as:

$$y[i] = S \sum_{s=0}^S x[i + r \cdot s]w[s]$$

Such type of convolutional kernel is also rotation invariant. The dilations can be mentally conceptualized as the introduction of discrete intervals of pixel that are used for the convolution kernel, and are dictated by the dilation rate  $r$ .

In this context the proposed architecture, depicted in Figure 1, makes use of dilated kernels in the contracting path of the network, with a total of 8 dilated convolutional layers. Moreover, in the present study we intended to explore the effects of the dilations sequence in an exponential manner (here after referred to as Method 1) and non-exponential

(here after referred to as Method 2), given that these may have an effect in the results that may result in gridding of the output probability maps as reported in previous studies of the literature [9].

As non-linear activation, we use Rectifier Linear Units (ReLU) in the convolutional operations. The contracting path of the network intends to extract higher order feature representations of the images. The expanding path follows, and upsamples the information retrieved in the feature maps towards the original resolution. Input images and the corresponding liver segmentation masks provided by human experts were used to train the network. Examples of ground truth masks are latter presented in the Results section. To learn the whole liver supervised features an FCN model was trained. Such model was formulated taking into account the several sizes of receptive fields that can allow the network to learn the most discriminative feature maps. Such methods require also the kernelized image context to correctly identify the liver voxels. The size of receptive field roughly indicates the amount of context information that is used in each feature map. The networks were trained iteratively, using the Adam optimizer [10], a learning rate of 0.0001 and the minimization of the class weighted cross-entropy. The hyperparameters were tuned so as to give best performance on validation set. The neural networks were implemented using Keras [11] with Tensorflow backend and trained using one NVIDIA Titan Xp GPU. The dataset was augmented using rotation and horizontal flipping to increase generalizability of the model.

### III. RESULTS AND DISCUSSION

The described method was used to segment the 2017 MICCAI Lits Challenge dataset. The proposed architecture was analysed with different dilation rate sequences to analyse certain issues and uncertainties introduced by this operation. All the presented variants of the method used that are being investigated in this paper were evaluated in the same test set. Training took from 9 to 12 hours for each model. The performance results are reported in terms of quantitative segmentation accuracy metrics. Quantitative evaluation of the developed methods is assessed by the Dice Score Coefficient (DSC), and sensitivity between the segmentation by one method and the ground truth. The obtained results are shown in Table I, along with a comparison towards other top performing methods presented in the literature. As baseline, we started by training the U-net network, as defined by [14], thus comprising convolutional, pooling, and upsampling operations along with short and long skip connections. The approach proposed attempts to overcome the drawbacks of data imbalance and boundary clarity. Results were highly enhanced by the combination of a CNN and the and the dilation-based encoding step implemented in the architecture. Moreover, the positive difference in performance derived from the analysis of the non-exponential dilation pyramid sequence was highly enhanced in particular for the segmentation of such small objects such as the liver lesions. To the best of our knowledge, no previous method taking advantage of the positive performance aspects of dilated convolutions was

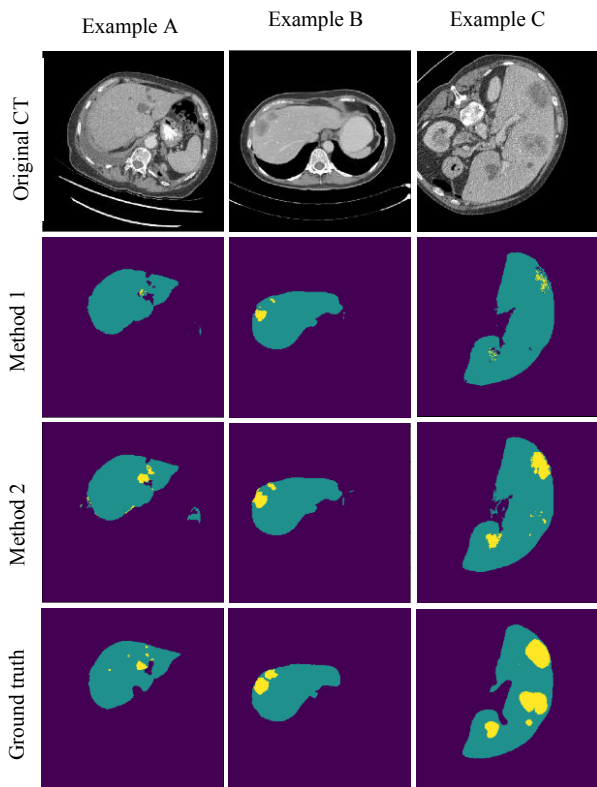


Figure 2. Examples of segmentations obtained on 30 CT scans of the 2017MICCAI Lits dataset. From top to bottom are shown the original CT, the liver segmentation (green) and tumour segmentation (yellow), of each method variant tested, to be compared against ground truth segmentations, represented in the last row.

previously proposed for the task of liver segmentation in abdominal CT images.

The proposed method achieves similar performance scores relative to the three top performing methods, as shown in Table I (example segmentation in Figure 2). We trained the model on slice-wise CT data for both liver tissues segmentation. A total 15051 image slices, corresponding to a total of 30 3D CT scans, not previously used for model training, were used to test the performance of the proposed model. Results obtained in this work, particularly regarding the segmentation of liver lesions, were also enhanced from the non truncation of HU intensities of the CT images after liver ROI selection from the first segmentation step of the cascaded approach used. For the specific task of lesion segmentation the truncation according to liver-specific HU values did not apply, and could potentially create the loss of local texture differences. Qualitatively, the segmentation results can be evaluated through Figure 2. In the two example results, the complex and heterogeneous structures of the liver were detected in the shown images. Overall, the model predictions were accurate in the classification of true positives. However, from the analysis of the entire dataset, the fuzziness of the liver boundaries in some scans leaked to the neighboring tissues, depicted in similar intensities. This is observable in Figure 2, in the example C.

In this work, we devise a simple, but efficient and end-to-end method that achieves state-of-the-art results in both quantitative metrics when compared to the three top performing methods of the literature.

#### IV. CONCLUSION

Segmentation of the liver in Computer Tomography (CT) images allows the extraction of the three-dimensional (3D) structure of the liver. The adequate receptive field for the segmentation of such a big organ in CT images, from the remaining neighboring organs was very successfully improved by the use of the state-of-the-art Convolutional Neural Networks (CNN) algorithms, however, certain issues still arise and are highly dependent of pre- or post-processing methods to refine the final segmentations. Here, a Fully Convolutional Network is investigated as a potential method to improve the segmentations results of liver tissues in CT. The introduction of a dilation module has allowed the concatenation of feature maps with a richer contextual information. The hierarchical learning process of such feature maps allows the decoder module of the model to have an improved capacity to analyze more internal pixel areas of the liver, with additional contextual information, given by different dilation convolutional layers.

TABLE I. RESULTS OBTAINED FOR THE SEGMENTATION USING THE DIFFERENT CNN VARIANTS EXPLORED (DSC – DICE SCORE COEFFICIENT, SE – SENSITIVITY).

Method	Liver DSC	Tumour DSC	Liver SE	Tumour SE
[12]	95.90	50.01	-	-
[13]	96.3	65.7	-	-
[14]	94.3	72.0	-	-
U-net	94.33	54.41	0.92	0.47
Method 1	95.02	57.22	0.94	0.49
Method 2	95.57	59.36	0.96	0.50

The obtained results are preliminary, of a deeper study of the role of dilated convolutions and the adequate aggregation of context for the purpose of object segmentation in future work. Specifically, further exploration of the advantages of the usage of dilated FCN approaches for the segmentation of small object, such as liver tumours will be investigated.

#### ACKNOWLEDGMENT

The authors gratefully acknowledge the funding from Project NORTE-01-0145-FEDER- 000022 - SciTech - Science and Technology for Competitive and Sustainable Industries, cofinanced by "Programa Operacional Regional do Norte (NORTE2020), through "Fundo Europeu de Desenvolvimento Regional (FEDER). The authors also kindly thank Nvidia, for the contribution with one Nvidia Titan XP GPU, that was used in this work.

#### REFERENCES

- [1] A. Gotra, Sivakumaran, L., Chartrand, G., Vu, K. N., Vandenbroucke- Menu, F., Kauffmann, C., Tang, A. (2017). Liver segmentation: indications, techniques and future directions. *Insights into imaging*,8(4), 377-392.
- [2] World Health Organization. Fact sheet: Cancer, 2015.URL: <http://www.who.int/mediacentre/factsheets/fs297/en/>.
- [3] Dawant, B. M., Li, R., Lennon, B., Li, S. (2007). Semi-automatic segmentation of the liver and its evaluation on the MICCAI 2007 grand challenge data set. *3D Segmentation in The Clinic: A Grand Challenge*, 215-221.
- [4] Lits - liver tumor segmentation challenge, 2017.
- [5] A. Ben-Cohen, I. Diamant, E. Klang, M. Amitai, and H. Greenspan. Fully convolutional network for liver segmentation and lesions detection. In G et. al Carneiro, editor, *Deep Learning and Data Labeling for Medical Applications*, pages 7785, Cham, 2016. Springer International Publishing. ISBN 978-3-319-46976-8.
- [6] P. F. Christ et al. (2017) Automatic liver and tumor segmentation of ct and mri volumes using cascaded fully convolutional neural networks. arXiv preprint arXiv:1702.05970.
- [7] Ronneberger O, et al. U-Net: Convolutional Networks for Biomedical Image Segmentation. In: *MICCAI 2015*. vol. 9351 of LNCS. Springer; 2015. p. 234241.
- [8] F. Yu and V. Koltun. (2015) Multi-scale context aggregation by dilated convolutions, arXiv preprint arXiv: 1511.07122.
- [9] Wang, P., Chen, O., Yuan, Y., Liu, D., Huang, Z., Hou, X., Cottrell, G. Understanding convolution for semantic segmentation. *WACV 2017*, 2018, 1451-1460.
- [10] Diederik P Kingma and Jimmy Ba. (2014) Adam: A method for stochastic optimization, arXiv preprint arXiv:1412.6980.
- [11] Chollet, F. URL: <https://github.com/keras-team/keras>
- [12] Lei Bi, Jinman Kim, Ashnil Kumar, and Dagan Feng. (2017) Automatic liver lesion detection using cascaded deep residual networks. URL <http://arxiv.org/abs/1704.02703>.
- [13] Yading Yuan. (2017) Hierarchical convolutional-deconvolutional neural net-works for automatic liver and tumor segmentation. URL- <http://arxiv.org/abs/1704.02703>.
- [14] Chlebus, G., Schenk, A., Moltz, J. H., van Ginneken, B., Hahn, H. K., Meine, H. (2018) Automatic liver tumor segmentation in CT with fully convolutional neural networks and object-based postprocessing. *Scientific reports*, 8(1), 15497.