

# Brain-vascular segmentation for SEEG planning via a 3D fully-convolutional neural network.

Sara El Hadji<sup>1</sup>, Sara Moccia<sup>2</sup>, Davide Scorza<sup>1,3</sup>, Michele Rizzi<sup>4</sup>, Francesco Cardinale<sup>4</sup>, Giuseppe Baselli<sup>1</sup> and Elena De Momi<sup>1</sup>

**Abstract**—Three dimensional visualization of vascular structures can assist clinicians in preoperative planning, intra-operative guidance, and post-operative decision-making. The goal of this work is to provide an automatic, accurate and fast method for brain vessels segmentation in Contrast Enhanced Cone Beam Computed Tomography (CE-CBCT) dataset based on a residual Fully Convolutional Neural Network (FCNN). The proposed NN embeds in an encoder-decoder architecture residual elements which decreases the vanishing effect due to deep architecture while accelerating the convergence. Moreover, a two-stage training has been proposed as a countermeasure for the unbalanced nature of the dataset. The FCNN training was performed on 20 CE-CBCT volumes exploiting mini-batch gradient descent and the Adam optimizer. Binary cross-entropy was used as loss function. Performance evaluation was conducted considering 5 datasets. A median value of Dice, Precision and Recall of 0.79, 0.8 and 0.69 were obtained with respect to manual annotations.

## I. INTRODUCTION

Stereo-Electro-encephalography (SEEG) involves the implantation of multiple intracranial electrodes to map spatio-temporal dynamics of epileptic seizures to identify the epileptogenic foci. In SEEG procedures, brain vessels are among the most critical landmarks that need to be accurately localized for reducing surgical risks, such as intracranial hemorrhage. Thus, a careful preoperative planning is required to identify safe electrode trajectories that preserve vascular structures [1].

Several approaches have been proposed for vessel segmentation and a recent review can be found in [2]. Focusing on SEEG planning, the work in [3] and [4] applies Frangi's 3D multi-scale vesselness on Magnetic Resonance Imaging (MRI) volumes. Similarly, the work in [5] proposes a multi-scale, multi-modal tensor voting algorithm. In [6] a Gaussian Mixture Model (GMM) based on Markov Random Fields (MRFs) combined to manual thresholding is used to segment vessels from Maximum Intensity Projection (MIP) volumes.

Despite these approaches being promising, (i) the limits of currently involved imaging techniques and (ii) the com-

plex morphology of cerebrovascular structures hinders their translation into the actual clinical practice [7].

To partially overcome limits in vessel imaging, a new imaging procedure called Cone Beam Computed Tomography (CBCT) Three-Dimensional Digital Subtraction Angiography (3D DSA) has been proposed [8]. This procedure consists in the acquisition of two CBCT volumes, before and during the injection of a contrast medium, namely the bone-mask (BM) and CE-CBCT volumes. The two volumes are organized to obtain the 3D Digital Subtracted Angiography (3D-DSA) volume, in which vessel and bone contrast are enhanced and attenuated, respectively.

3D-DSA allows enhanced vessel visualization, however, it requires further post-processing in order to obtain an accurate vascular mask for automatic SEEG planning. Moreover, 3D-DSA can suffer from registration error: despite the subtraction of the BM from the CE, residual bone structure and noise affect the quality of 3D DSA for which a manual thresholding is not sufficient. This could be done using deep-learning strategies, such as Fully Convolutional Neural Networks (FCNN). FCNNs have achieved remarkable results in vessel segmentation by automatically learning complex image features and combining them into hierarchical representations for prediction and classification [2]. Most approaches are however developed to process 2D images, while medical data involved in SEEG planning consist of 3D volumes.

In this work, we propose a novel 3D brain vasculature segmentation that exploits the power and effectiveness of an end-to-end convolutional neural network architecture. Our contribution is two folds. First, instead of processing the input volumes in a 2D slice-wise fashion, we propose to apply directly 3D convolutional operations to exploit the intrinsic tridimensional information encoded in the CE-CBCT volume. Second, we combine a two-stage training specifically designed for medical image segmentation to leverage class imbalance.

This paper is organized as follows: in Sec. II the proposed FCNN is described, along with our two-phase training strategy to account for vessel/background imbalance. In Sec. III, dataset specifications and performance metrics are reported. In Sec. IV, the achieved results are reported. Finally, Sec. V concludes this paper summarizing the main achievements and discussing some future research directions.

\*This work has been funded by Medtronic External Research Program.

<sup>1</sup>S. El Hadji, G. Baselli, E. De Momi Department of Electronics, Information and Bioengineering, Politecnico di Milano, Milan (Italy)sara.elhadji@polimi.it

<sup>2</sup>S. Moccia Department of Information Engineering, Università Politecnica delle Marche, Ancona (Italy), Department of Advanced Robotics, Istituto Italiano di Tecnologia, Genoa (Italy)

<sup>3</sup>D. Scorza e-Health and Biomedical Applications Department, Vicomtech, Parque Tecnológico, San Sebastián (Spain), Department of Electronics, Information and Bioengineering, Politecnico di Milano, Milan

<sup>4</sup>M. Rizzi, F. Cardinale Claudio Munari Centre for Epilepsy and Parkinson Surgery, Niguarda Ca Granda Hospital, Milan (Italy)

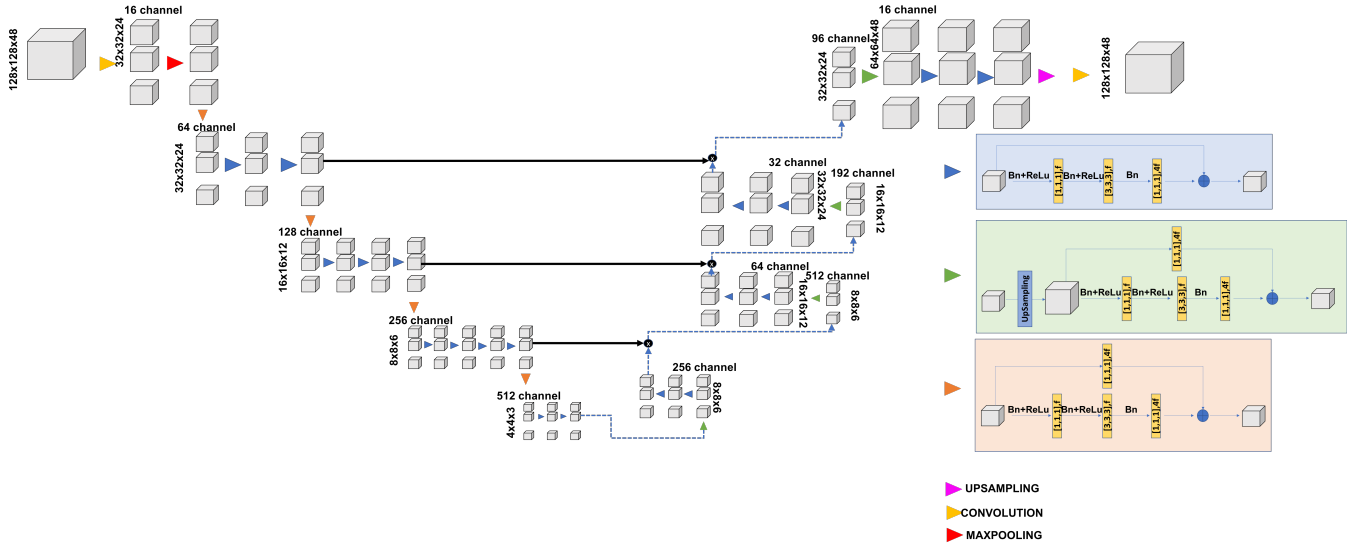


Fig. 1. The scheme of the proposed Fully Convolutional Neural Network (FCNN) architecture: a U-Net architecture [9] is integrated with the residual blocks introduced in [10] (designed according to a bottleneck setting firstly proposed in [11]). Blue, green and orange boxes on the right report the architecture of the employed shortcuts.

## II. METHODS

### A. Architecture

Figure 1 illustrates the proposed 3D-FCNN architecture, inspired by [9], that demonstrated outstanding performance in medical-image segmentation. Compared to [9], our 3D FCNN also embedded residual blocks, firstly introduced in [10] which allows to overcome the vanishing gradient effect typical of FCNN training and speed up the training convergence.

Our FCNN architecture is composed of eight stages, excluding the input and output blocks, arranged as a downsampling-upsampling style encoder-decoder network with four resolution steps. Each stage contains a series of residual unit blocks characterized by two types of shortcut connections, namely identity and projection. The shortcuts are structured with a bottleneck design, which consisted of a  $1 \times 1 \times 1$  layer, a  $3 \times 3 \times 3$  layer, and a  $1 \times 1 \times 1$  layer where the first and last convolutional layers are used for reducing and restoring image dimension, respectively. This technique is used to decrease the number of parameters while not affecting the performance of the network [12]. Moreover, in the identity mapping residual block, each convolutional layer with unit stride is preceded by batch normalization and ReLU activation according to the full pre-activation setting described in [11], which has been found to be more powerful than the baseline ResNet residual block.

In the first four stages, the number of convolutional filters is increased of a factor of two, allowing a wider reception field, as well as in the remaining four stages decreased of a factor of two. Skip connections from layers of equal resolution allows to recover spatial information by merging features from same resolution levels between the contracting and expanding path.

The input block, was equivalent to the first 3 layers of

ResNet50: a convolutional layer with  $7 \times 7 \times 7$  and strides = (2,2,1) followed by BN + ReLU activation and maxPooling operation with a stride of 2 for downsampling. The output block was a fully connected layer followed by a binary sigmoid classifier to produce class probabilities for each voxel. The architecture had a total of 20760481 parameters and 279 layers.

### B. Training

As described in Sec. III-A, each of our CE-CBCT volume had size of  $512 \times 512 \times 192$  pixels. To balance the huge memory requirements of a 3D-FCNN training without suffering neither from loss of resolution due to volume resampling or efficiency issues due to dense patch-wise training, we opted for a “large” patch-extraction approach. Thus, we extracted from each volume 64 3D patches of dimensions  $128 \times 128 \times 48$ , i.e. the maximum dimension allowed by current graphic card. This on one hand increased the training-set size and on the other one offered a good alternative to capture both local and global contextual information.

1) *Two-phase training*: Our brain-vessel segmentation problem was highly class-imbalanced (background voxels accounted for the 99% of total voxels). A random patch selection would have caused the FCNN training to be overexposed to no-vascular patches, negatively affecting the FCNN predictive performance. To overcome this problem, we initially built our training-patch dataset such that the two classes (i.e. vessel, non-vessels) were equiprobable. Thus, we firstly trained our model on this dataset. Then, in a second phase, we re-built the patch dataset accounting for the unbalanced nature of our dataset and re-trained only the FCNN output layer (i.e., keeping fixed the FCNN first-layer weights).

2) *Training parameters*: During training, the FCNN model was regularized with a weight decay of  $1e-4$  and all

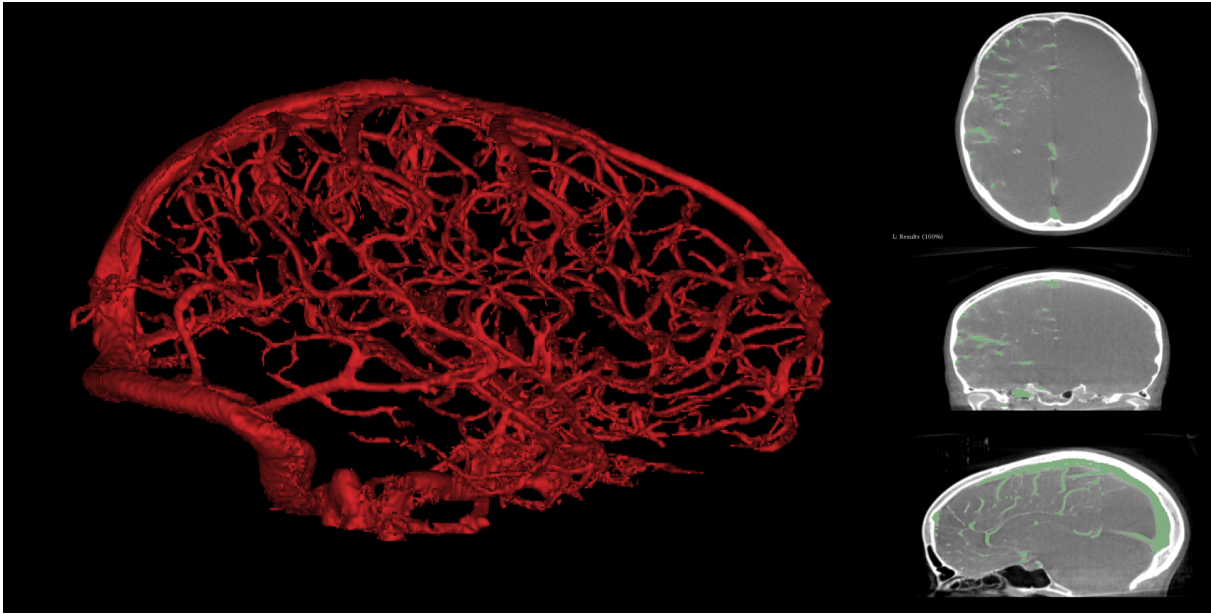


Fig. 2. On the left the volume rendering of a cerebrovasculature mask. On the right the axial, sagittal and coronal views of a Contrast Enhanced (CE) Cone Beam Computed Tomography (CBCT) volume with the vascular mask in green.

the weights were initialized according to [13]. To update the FCNN weights during training, the ADAM [14] optimizer was used with an initial learning rate of  $3e^{-4}$  and a scheduled reduction of 10% at each loss function plateau. Mini-batch gradient descent was used to minimize the binary cross-entropy, used as loss function, with a batch size equal to 6.

To monitor the training process, we randomly selected the 25% of the training data as a validation set. The validation set was used to early-stop the training process based on the Dice similarity Coefficient ( $DSC$ ), with patience of 25 epochs:

$$DSC = 2 \cdot \frac{2Tp}{Fp + Fn + 2Tp} \quad (1)$$

where  $Tp$  is the number of vessel voxels correctly classified, and  $Fp$  and  $Fn$  are the voxels misclassified as vessels and background, respectively.

The network was trained using Keras and Tensorflow framework<sup>1</sup>. All training procedures and experiments were conducted on a workstation equipped with one single GPU NVIDIA GEFORCE GTX-1080 Ti with 12 GB of RAM on a Windows 10 Professional 64-bit operating system, AMD Ryzen 7 1700X eight-core processor @3.40GHz with 16 GB of RAM.

### III. EXPERIMENTAL SETUP

#### A. Dataset

The proposed FCNN was retrospectively applied on 25 patient volumes acquired in 2017 at the Claudio Munari Centre for Epilepsy and Parkinson Surgery, Niguarda Ca Granda Hospital, Milan (Italy). All procedures were in accordance with the ethical standards of the institutional research committee and with the 1964 Helsinki Declaration

<sup>1</sup><https://keras.io/>

and Ethical approval was sought from the Niguarda Hospital Ethical Committee. Individual patient consent was obtained for use of anonymized preoperative imaging. For this study, formal consent was not required. At the time of the acquisition, all patients did not present any vasculature pathologies and underwent the standard acquisition protocol for SEEG pre-operative planning.

The CE-CBCT acquisition was performed by means of the O-arm<sup>TM</sup> System (Medtronic; Minneapolis, Minnesota, USA) over 12s after the start of iopamidole infusion (300 mg/ml) in the ICA depicting the left/right hemisphere's vasculature tree. Each volume had a voxel resolution of  $0.415 \times 0.415$  mm in axial plane and 0.833 mm in z-direction, accounting for 192 slices, each of size  $512 \times 512$  pixels.

Each volume came with a single-label cerebro-vascular mask, which was retrieved from the 3D-DSA. The labelling was performed by a neurosurgeon, in a slice-by-slice fashion, by means of manual thresholding with the assistance of 3D visualization tools<sup>2</sup> to take into account the global 3D vascular structure. The binary mask was further manually refined using ITKsnap<sup>3</sup>.

#### B. Evaluation

The dataset was divided in three folds: 20 volumes were used for training and 5 others for testing purpose only. To quantitatively evaluate the performances of the proposed 3D FCNN, we computed the  $DSC$ , precision ( $Prec$ ) and recall ( $Rec$ ), which are defined as follows:

$$Prec = \frac{Tp}{Tp + Fp} \quad (2)$$

<sup>2</sup><https://www.slicer.org/>

<sup>3</sup><http://www.itksnap.org>

TABLE I

MEDIAN (INTER-QUARTILE RANGE) OF PERFORMANCE METRICS FOR OUR 3D FCNN COMPARED TO STATE OF THE ART ARCHITECTURES.

DICE SIMILARITY COEFFICIENT ( $DSC$ ), PRECISION ( $Prec$ ) AND RECALL ( $Rec$ ) ARE REPORTED. TRAINING TIME (TIME) IS REPORTED, TOO.

Methods	$DSC$	$Prec$	$Rec$	Time
Otsu's thresholding	0.49 (0.04)	0.41 (0.07)	0.39 (0.08)	—
Frangi Filter [15]	0.54 (0.04)	0.43 (0.07)	0.41 (0.08)	—
U-net [9]	0.59 (0.1)	0.51 (0.09)	0.42 (0.13)	22 hours
V-net [16]	0.65 (0.12)	0.59 (0.24)	0.57 (0.19)	28 hours
Our method	0.79 (0.13)	0.8 (0.09)	0.69 (0.12)	14 hours

$$Rec = \frac{T_p}{T_p + F_n} \quad (3)$$

where  $T_n$  represents voxels that were correctly identified as belonging to the background.

#### IV. RESULTS & DISCUSSION

Table 1 reports the performance in terms of median (inter-quartile range (IQR)  $Prec$ ,  $Rec$ , and  $DSC$ ). The obtained results showed a good agreement between the segmentation and manually-annotated data. Our approach resulted to be robust against beam hardening effect caused by the contrast mean, which increased the intensity of the surrounding area nearby the vessels (Fig. 2). Computational training time (14 hours) could be acceptable, considering both the lack of code optimization for our preliminary version and the standard available hardware. Moreover, it is worth noticing that the problem of SEEG surgical planning is typically offline.

A limitation of this study could be seen in the lack of a large clinical dataset for training and evaluation. This initial work has to be intended as a proof of concept for the described methodology, and we are currently working with our clinical partners on expanding our dataset.

#### V. CONCLUSION

In this paper, we introduced an automatic brain-vessel segmentation method in CE-CBCT that uses 3D deep convolutional neural networks. Experimental results demonstrated that the proposed architecture managed to improve with respect to the current state-of-the-art architectures. Such performance was obtained with the application of the full pre-activation setting of the residual blocks, which increased both the model convergence rate and the performance. Moreover, training the model in two phases offered an effective and efficient procedure in presence of highly unbalanced label distribution. As it regards the computational time for the segmentation, for a volume, excluding the patches extraction and reassembling time, it takes 145 seconds for the segmentation. Despite being slightly more computationally expensive against the Otsu's thresholding, this approach guarantees a better performance.

#### REFERENCES

- [1] Jeffrey P Mullin, Michael Shriver, Soha Alomar, Imad Najm, Juan Bulacio, Patrick Chauvel, and Jorge Gonzalez-Martinez. Is SEEG safe? A systematic review and meta-analysis of stereo-electroencephalography-related complications. *Epilepsia*, 57(3):386–401, 2016.
- [2] Sara Moccia, Elena De Momi, Sara El Hadji, and Leonardo S Mattos. Blood vessel segmentation algorithms—Review of methods, datasets and evaluation metrics. *Computer Methods and Programs in Biomedicine*, 158:71–91, 2018.
- [3] Rina Zelmann, Silvain Bériault, MM Marinho, Kelvin Mok, Jeffery A Hall, Nicolas Guizard, Claire Haegelen, Andre Olivier, G Bruce Pike, and D Louis Collins. Improving recorded volume in mesial temporal lobe by optimizing stereotactic intracranial electrode implantation planning. *International Journal of Computer Assisted Radiology and Surgery*, 10(10):1599–1615, 2015.
- [4] Xiaofei Du, Hui Ding, Wenjing Zhou, Guangming Zhang, and Guangzhi Wang. Cerebrovascular segmentation and planning of depth electrode insertion for epilepsy surgery. *International Journal of Computer Assisted Radiology and Surgery*, 8(6):905–916, 2013.
- [5] Maria A Zuluaga, Roman Rodionov, Mark Nowell, Sufyan Achhala, Gergely Zombori, Alex F Mendelson, M Jorge Cardoso, Anna Miorocchi, Andrew W McEvoy, John S Duncan, and S. Ourselin. Stability, structure and scale: improvements in multi-modal vessel extraction for SEEG trajectory planning. *International Journal of Computer Assisted Radiology and Surgery*, 10(8):1227–1237, 2015.
- [6] Davide Scorza, Sara Moccia, Giuseppe De Luca, Lisa Plaino, Francesco Cardinale, Leonardo S Mattos, Luis Kabongo, and Elena De Momi. Safe electrode trajectory planning in SEEG via MIP-based vessel segmentation. In *Medical Imaging 2017: Image-Guided Procedures, Robotic Interventions, and Modeling*, volume 10135, page 101352C. International Society for Optics and Photonics, 2017.
- [7] Vejay N Vakharia, Rachel Sparks, Roman Rodionov, Sjoerd B Vos, Christian Dorfer, Jonathan Miller, Daniel Nilsson, Martin Tisdall, Stefan Wolfsberger, Andrew W McEvoy, et al. Computer-assisted planning for the insertion of stereoelectroencephalography electrodes for the investigation of drug-resistant focal epilepsy: an external validation study. *Journal of Neurosurgery*, pages 1–10, 2018.
- [8] Francesco Cardinale, Guglielmo Pero, Luca Quilici, Mariangela Piano, Paola Colombo, Alessio Moscato, Laura Castana, Giuseppe Casaceli, Dalila Fuschillo, Luciana Gennari, Marco Cenozato, Giorgio Russo, and Massimo Cossu. Cerebral angiography for multimodal surgical planning in epilepsy surgery: description of a new three-dimensional technique and literature review. *World Neurosurgery*, 84(2):358–367, 2015.
- [9] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 234–241. Springer, 2015.
- [10] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 770–778, 2016.
- [11] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Identity mappings in deep residual networks. In *European Conference on Computer Vision*, pages 630–645. Springer, 2016.
- [12] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. Going deeper with convolutions. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–9, 2015.
- [13] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In *IEEE International Conference on Computer Vision*, pages 1026–1034, 2015.
- [14] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.
- [15] Alejandro F Frangi, Wiro J Niessen, Koen L Vincken, and Max A Viergever. Multiscale vessel enhancement filtering. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 130–137. Springer, 1998.
- [16] Fausto Milletari, Nassir Navab, and Seyed-Ahmad Ahmadi. V-net: Fully convolutional neural networks for volumetric medical image segmentation. In *International Conference on 3D Vision*, pages 565–571. IEEE, 2016.