2D and 3D-UNet for segmentation of SEEG electrode contacts on post-operative CT scans

Anja Pantović[1], Irène Ollivier[2], Caroline Essert[1]

[1]ICube Laboratory, Université de Strasbourg, CNRS, Strasbourg, France;
[2]Department of Neurosurgery, Strasbourg University Hospital, Strasbourg, France

ABSTRACT

Stereoelectroencephalography (SEEG) is a minimally invasive surgical procedure, used in the treatment of pharmacoresistant epilepsy to precisely locate areas of the brain where seizures originate. An accurate localization of SEEG electrodes is crucial to design a resection plan before surgically removing epileptogenic zone. We propose to train a deep neural network to accurately segment electrode contacts without making any manual adjustments. We trained a 2D and a 3D version of the U-Net[1,2] neural network architecture to handle this task, taking post-operative CT scans as input. We evaluated our models on 18 image datasets of patients using different metrics, and provided a comparison of the two approaches. The presented models are robust to electrode bending and do not need any prior information to make quick and accurate predictions. To the best of our knowledge, deep learning has not been used yet for this task.

1. INTRODUCTION

Epilepsy is one of the most common serious brain disorders, characterized by an enduring predisposition to generate epileptic seizures.[3] Over 30% of epilepsy patients are pharmacoresistant and for them, the disease could potentially be cured by surgically removing the epileptogenic zone.[4] Prior to the intervention, such zones must be identified and precisely localized, as they are patient-specific.

Stereoelectroencephalography (SEEG) provides a way to detect the seizure onset zone and hence guide further surgical decision making. During this procedure, 10 to 18 electrodes are implanted into the patient's brain. Each electrode contains between 5 and 18 metallic contacts that record the electrical activity within suspected areas during a few days.[5,6] Surgeons then analyze the signal recorded by each contact. By linking that information with their spatial localization, they can delineate the epileptogenic zones and choose a resection plan. This is why an accurate localization of all the implanted electrode contacts is crucial.

Electrode contact segmentation is usually performed manually by neurosurgeons. Having to segment between 100 and 250 contacts per patient is an extremely time-consuming task which often results in inaccurate segmentation. One of the challenges is the artefacts that occur when electrodes are placed close to each other. The difficulty further increases when electrodes are inserted vertically into the brain. To help the neurosurgeons, several approaches have been proposed to improve and accelerate the segmentation of SEEG electrodes, saving time and increasing the efficiency of the procedure.[7–10] Arnulfo et al.[9] suggested an approach based on a geometrical-constrained search. Meesters et al.[10] extracted guiding screws with a multi-scale filter and determined the possible tip locations inside a wedge-shaped region. This method did not take into account electrode bending, assuming rigid electrodes. Granados et al.[7] suggested a method aiming for an automatic segmentation of both bolts and contacts. Their contact search strategy was based on the direction of the bolt, given the distance and angle constraints. The proposed method, however, needs manual adjustments to handle electrodes crossing. All of these methods use co-registered post-implantation CT scans with pre-implantation MRI and hence rely on pre-operative plans. Benadi et al.[8] developed two interactive and an automatic method carried out using 3D Slicer platform with post-operative CT scans. The automatic method does not take into account electrode bending and inaccurately segments contacts that are surrounded by strong artefacts.

✉Caroline Essert, essert@unistra.fr

To the best of our knowledge, deep learning has not been used yet for solving the SEEG electrode segmentation task. However, with the rise of medical imaging, deep learning has encountered significant advances in the field of machine learning, becoming one of its most successfully used tools. Deep neural networks are being used with CT and MRI scans and trained for different classification and segmentation tasks, leading to the automation of tasks that clinicians would otherwise perform manually. In order to improve and automate the process of SEEG electrode contact segmentation, we propose an approach based on the U-Net[1] encoder-decoder architecture implementing the 2D and the 3D network. The 3D U-Net architecture was first proposed by Çiçek et al.[2] Their network learns from sparsly annotated medical data and provides a dense 3D segmentation.

2. METHODS

2.1 Data Acquisition and Preprocessing

Our dataset consists of postoperative CT scans from 18 different pharmacoresistant epilepsy patients, obtained at the Department of Neurosurgery of Strasbourg University Hospital. For each patient we have the electrode implantation planning specifying the total number of electrodes implanted as well as the model of each electrode. Each CT slice has a resolution of 512×512 pixels viewed from the axial plane, with the number of slices in each scan varying between 166 and 259. As border pixels of each slice do not capture any meaningful information, corresponding to the black background, all slices were cropped to the uniform dimension of 384×384 to reduce memory usage and decrease the computation time. The complete area of the head was preserved after this operation.

Annotation of the electrode contacts was done manually using the 3D Slicer[11] platform and its segmentation tools. The number of annotated electrode contacts as well as the total number of electrodes annotated was compared to the planning document to assure a high accuracy of the manual segmentation.

2.2 Electrode Contact Segmentation

To segment SEEG contacts from the post-implantation CT scans, we have implemented a 2D and a 3D variant of the encoder-decoder symmetric U-Net[1,2] architecture. The 3D network is illustrated on Figure 1. The left side of the symmetric "U" shape is called contracting path and represents a general convolutional process, consisting of the repeated application of two 3×3 convolutions (or $3 \times 3 \times 3$ for the 3D model). Each convolutional layer is followed by a ReLU activation and batch normalization. Then, spatial dimensions are reduced by applying a 2×2 (or $2 \times 2 \times 2$) max pooling operation. The right side of the "U" shape is the expansive path and is built by upsampling of the feature map followed by a 2×2 (or $2 \times 2 \times 2$) transpose convolution, reducing the number of feature channels by half. At the final layer, a 1×1 (or $1 \times 1 \times 1$) convolution is used after which each pixel is classified either as an electrode contact or a background pixel.

Both networks were trained and tested on the same patient dataset, with slight dimension adjustments for the 3D model due to its heavy memory usage. For the 2D network, the data was organized as 3788 independent CT slices of dimension 384×384. For the 3D U-Net, slices were grouped by patient. To fit the data into GPU RAM memory, the 3D network was trained on smaller patches. Each volume was subdivided into cubes of dimension $96 \times 96 \times 96$. Depending on the total number of CT slices per patient, these volumes were either of dimension $383 \times 384 \times 192$ or $384 \times 384 \times 288$. The initial number of slices for each patient was then either reduced by removing border slices or, where reduction was not possible due to the loss of useful information, extended by adding black slices to obtain a dimension divisible by 96.

Training of both networks was done in nearly the same initial hyper parameter set up. To ensure generalization of the results, both networks were cross validated using the 'leave-one-out' approach. In each training phase, training set consisted of 17 patient data while volume belonging to a single patient was left out to be used for testing. We used the normalization process for all the CT scans with 0 mean and 1 standard deviation. The training/validation split for the 2D model was such that 90% of training data were used for training the network, while the other 10% were used for validation. For the 3D model volume of 1 patient from the training data was always kept out for validation, while 16 remaining patient volumes were used for training. Validation set is used to provide an unbiased evaluation of a model on the training dataset while tuning the hyperparameters. For the entire training process we used Adam optimizer with the initial learning rate set to $1 * 10^{-5}$. The loss function

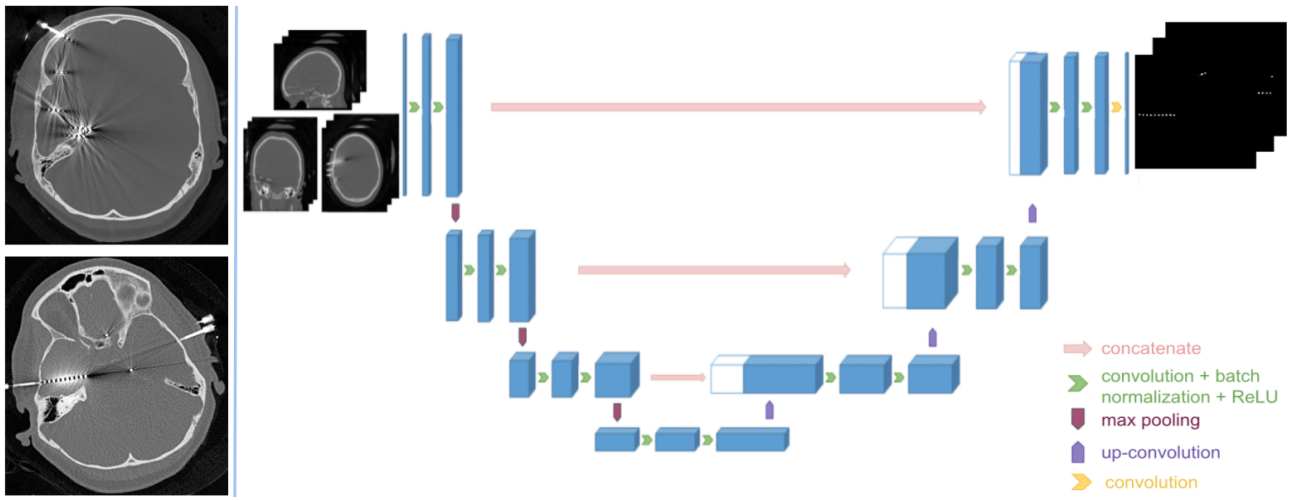

Figure 1. Pre-operative CT slices and the 3D U-Net netwrok architecture.[2] The network takes CT slices as input and returns predicted masks of electrode contacts.

was binary crossentropy. Testing was done on the set of slices belonging to a single patient, unseen during the training phase. Networks were trained during 150 epochs for each cross-validation fold and evaluated based on the computation time and accuracy.

3. RESULTS

To validate our methods, we have trained and tested both networks on CT scans from 18 different pharmacoresistant epilepsy patients, who underwent the SEEG procedure, as described in subsection 2.1. Training was performed on 4 GPUs, NVIDIA GeForceRTX 2080 Ti, with 4×11GB of RAM. Network performance was evaluated and compared based on 3 different metrics - pixels accuracy, Jaccard index and Dice coefficient. Pixel accuracy computes the percentage of pixels in an image that are classified correctly. The Jaccard index, also known as Intersection over Union (IoU), is the area of overlap between the predicted segmentation and the ground truth divided by the area of union between the two. Dice Coefficient is computed as 2 times the area of overlap divided by the total number of pixels in the predicted segmentation and the ground truth.

Network performance and above mentioned evaluation metrics were computed and averaged over all folds after the cross validation. Results are represented in Table 1.

Table 1. 2D U-Net and 3D U-Net network performance after 150-epoch-training performed on 4 NVIDIA GeForceRTX 2080 Ti GPUs. Average results of the 2D and 3D model after leave-one-out cross validation, evaluated and compared on 3 different metrics: pixel accuracy, intersection over union (IoU) and the Dice coefficient.

Model	**Training time**	**Prediction time**	**Test loss**	**Accuracy**	**IoU**	**Dice coefficient**
2D U-Net	45 minutes	1.4 seconds	0.0002	0.999	0.704	0.807
3D U-Net	73 minutes	5.6 seconds	0.0002	0.999	0.701	0.806

It is worth noting that SEEG electrode contacts occupy a very small percentage of the entire CT volume. This, from early on, makes the pixel accuracy very high, not providing too much information on how precise the prediction really is. Furthermore, this together with the metal artifacts that appear around electrode contacts, makes manual segmentation highly prone to human error. Here each pixel has a strong impact on the results and this sensitivity makes it difficult to reach a high Dice coefficient and Jaccard index. Apart from the evaluated metrics, we analysed where the misclassified pixels come from and discussed how to overcome such these faults.

Figure 2 (a-e) and Figure 2 (f-j) show the predictions obtained with the 2D and the 3D model respectively. True postivive pixels are represented in green, false positives in red and false negative pixels are displayed in

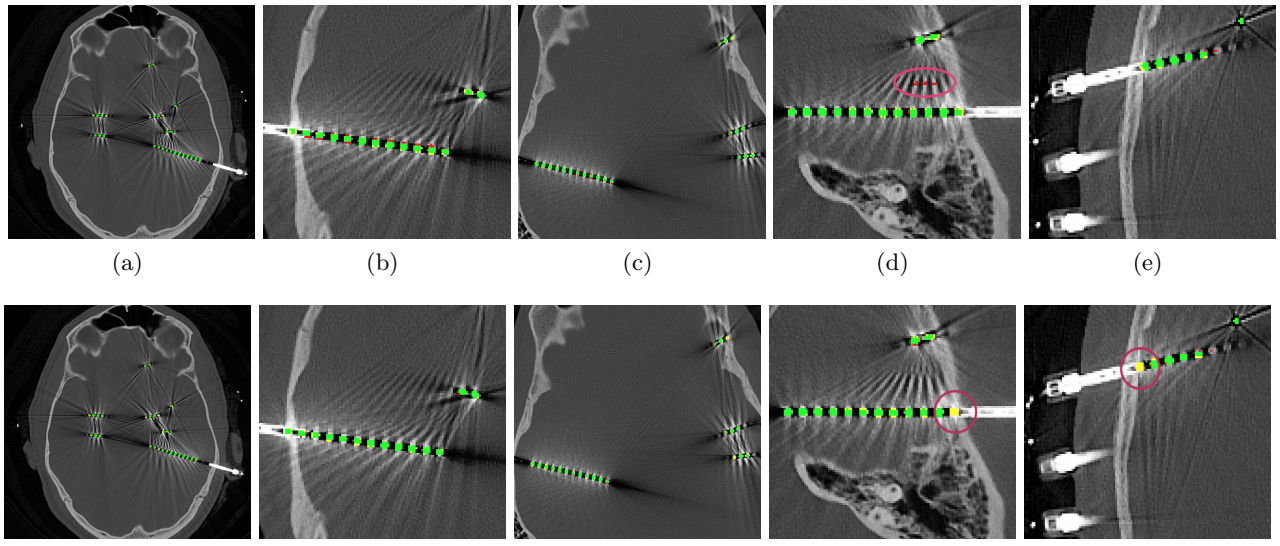

Figure 2. Predictions obtained with the 2D (top,a-e) and the 3D model (bottom,f-j). True positive pixels are represented in green, false positives in red and false negatives in yellow. Most of the contacts are properly segmented by both models (a-c, e, f-h). In some cases, the 2D model misclassifies artefacts as electrode contacts (d). The 3D model often fails to classify contacts that coincide with the skull or the screw in the CT scan (i,j).

yellow. Both models correctly segment most of the contacts. As it can be seen in Figure 2d, in some cases the 2D model classifies artefacts as contacts. Using additional information along the 3rd axis, the 3D model (Figure 2i) successfully learns to distinguish between the artefacts and metallic contacts. Both models in some cases fail to classify contacts close to the screw and skull bone on the CT scan. However, the 2D model classifies such contacts correctly (Figure 2 d,e) in more cases than the 3D model (Figure 2 i,j).

4. CONCLUSION

In this paper we proposed a method for segmentation of SEEG electrode contacts using the 2D and 3D U-Net encoder-decoder architecture on post-operative CT scans. Our models have been evaluated on 18 different patients who underwent the SEEG procedure. We reported high accuracy results for both the 2D and the 3D model. These networks could segment post-operative CT slices in less than 6 seconds using 4 Nvidia GeForceRTX 2080 Ti GPUs and in less than 85 seconds using a standard PC with Nvidia GeForce GTX 1060. The proposed solution does not require any prior information about the type nor number of electrodes.

Both models have respective limitations that are overcome by the other: the 2D model is sensitive to artefacts and finds false positives that the 3D model properly avoids, while the 3D model is less efficient with the most external contacts that the 2D model properly detects. The proposed models seem to complement each other. A hybrid network making the most of both models may potentially overcome the above mentioned issues and should be investigated in future work. Furthermore, we plan to implement a method to group and link together the contacts belonging to the same electrode.

5. ACKNOWLEDGMENTS

This work was funded by ArtIC project “Artificial Intelligence for Care” (grant ANR-20-THIA-0006-01) and co-funded by Région Grand Est, Inria Nancy - Grand Est, IHU of Strasbourg, University of Strasbourg and University of Haute-Alsace, France.

REFERENCES

[1] Ronneberger, O., Fischer, P., Brox, T., "U-Net: Convolutional Networks for Biomedical Image Segmentation," *MICCAI* (2015).

[2] Çiçek, Ö., Abdulkadir, A., Lienkamp, S.S, Brox, T., Ronneberger, O., "3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation," *MICCAI* (2016).

[3] Duncan, J., Sander, J., Sisodiya, S., Walker, M., "Adult epilepsy," *Lancet 367(9516):1087-1100* (2006).

[4] Kwan, P., Brodie, M.J., "Early identification of refractory epilepsy," *The New England Journal of Medicine* (2014).

[5] Vaugier, L., Lagarde, S., Mcgonigal, A., Trébuchon, A., Milh, M., et al., "The role of stereoelectroencephalography (SEEG) in reevaluation of epilepsy surgery failures," *Epilepsy & Behavior* (2018).

[6] Boyd, S.G., Cross, J.H., Eltze, C., "SEEG-guided radiofrequency thermocoagulation of epileptic foci in the paediatric population: Feasibility, safety and efficacy," *Seizure* (2019).

[7] Granados, A., Vakharia, V., Rodionov, R., Schweiger, M., Vos, S.B., O'Keeffe, A.G., Li, K., Wu, C., Miserocchi, A., McEvoy, A.W., Clarkson, M.J., Duncan, J.S., Sparks, R., Ourselin, S., "Automatic segmentation of stereoelectroencephalography (SEEG) electrodes post-implantation considering bending," *International Journal of Computer Assisted Radiology and Surgery* (2018).

[8] Benadi, S., Ollivier, I., Essert, C., "Comparison of interactive and automatic segmentation of stereoelectroencephalography electrodes on computed tomography post-operative images: preliminary results," *Healthcare Technology Letters* (2018).

[9] Narizzano, M., Arnulfo, G., Ricci, S. et al., "SEEG assistant: a 3DSlicer extension to support epilepsy surgery," *BMC Bioinformatics* (2017).

[10] Meesters, S., Ossenblok, P., Colon, A., Schijns, O., Florack, L., Boon, P., Wagner, L., Fuster, A., "Automated identification of intracranial depth electrodes in computed tomography data," 976–979 (04 2015).

[11] Kikinis, R., Pieper, S.D., Vosburgh, K., "3D Slicer: a platform for subject-specific image analysis, visualization, and clinical support.," in [*Intraoperative Imaging Image-Guided Therapy*], 277–289 (2014).