Automatic segmentation of the cerebral ventricle in neonates using deep learning with 3D reconstructed freehand ultrasound imaging

Matthieu Martin
CREATIS Laboratory, Lyon, France*
matthieu.martin@creatis.insa-lyon.fr

Bruno Sciolla
CREATIS Laboratory, Lyon, France*

Michael Sdika
CREATIS Laboratory, Lyon, France*

Xiaoyu Wang
IMT Atlantique, Brest, France

Philippe Quetin
CH Avignon, Avignon, France

Philippe Delachartre
CREATIS Laboratory, Lyon, France*

*Univ Lyon, INSALyon, Universit Claude Bernard Lyon 1, UJM-Saint Etienne, CNRS, Inserm, CREATIS UMR 5220, U1206, F69100, LYON, France.

Abstract—Preterm neonates can be subject to ventriculomegaly, which is an enlargement of the cerebral ventricle system (CVS) that can lead to brain damage. In clinical practice, 2D coronal hand-held ultrasonographic scans are performed to assess CVS dilation. Estimating CVS volumes from 2D images is, however, imprecise and time consuming since 3D information is lacking. To address this issue, we propose a 3D reconstruction method and an automatic deep learning segmentation algorithm. The accuracy of the 3D reconstruction was assessed by calculating Mean Absolute Distance (MAD) between manual segmentation of the corpus callosum (CC) on a ground reference and the 3D reconstructed volume, a mean value of 1.55 mm was obtained. The accuracy of the segmentation was evaluated using DICE, Hausdorff distance (d_H) and MAD, respective average values of 0.816, 13.6 mm and 0.62 mm were obtained. The computation time of a segmentation for one $256 \times 256 \times 256$ volume was 5 s.

Index Terms—preterm neonate, 3D reconstruction, ventriculomegaly, segmentation, deep learning

I. INTRODUCTION

The number of preterm births is increasing and estimated to 15 million every year in the world. Preterm neonates, particularly with a very low birth weight (< 1500 g), are at high risk for CVS dilation. This state corresponds to an abnormal accumulation of cerebrospinal fluid inside the CVS, which can pressure the brain against the skull, and lead to neurological impairments such as cerebral palsy, hearing loss or blindness [1], [2].

In usual clinical practice, the diagnosis and monitoring of ventriculomegaly is performed with 2D freehand ultrasound (US) imaging. When ventriculomegaly is detected, the CVS volume is estimated in order to determine the amount of fluid that must be punctured. In order to get an estimation of the CVS volume, clinicians usually realize a manual segmentation or various measurements on 2D images. However this practice is imprecise (since 3D information is missing), time consuming and operator dependent.

In this article, we propose an algorithm to reconstruct 3D US images from 2D freehand image sequences. In addition, the 3D volumes are parametrized according to the standard MRI coordinate system (axial, coronal and sagittal planes). To allow for an operator-independent assessment of ventriculomegaly and to spare clinical time, automatic and fast segmentation algorithms must be developed. The automatic segmentation of CVS in 3D US data has been achieved by [4] using an algorithm based on level-set method and by [5] using an atlas-based MRI/3D US registration method. Nevertheless the method proposed by [4] requires 54 min to segment one volume, which is too long to be used in clinical routine, and no segmentation time was mentioned by [5]. Recently, deep learning based algorithms have demonstrated fast and accurate segmentation results for medical images [7], [8]. In [6], deep neural networks were used to segment the CVS in 2D US images and a segmented time of 22 ms for 512×512 images was reported.

In this paper, we propose a pipeline CVS for segmentation using a deep learning approach after reconstruction of 3D images from stacks of 2D freehand angular sweep. The segmentation of the CVS obtained is fast and accurate. To the best of our knowledge, this is the first work to achieve the CVS segmentation in 3D US data using deep learning.

II. METHOD

A. Data description

In this study, 15 2D freehand angular sweep (Fig. 1a) image sequences (Fig. 1b) were acquired by a pediatrician in coronal orientation. The acquisition were realized through anterior fontanel with a 2.5D Siemens Acuson 9L4 transducer over 14 patients at the average age of 35.8 ± 0.8 SA. Sequences were composed of T images with T ranging from 136 to 306 depending on the acquisition. In addition, a sagittal image (Fig. 4a) was acquired (during the same examination) for every patient in order to be used as a reference for the reconstruction algorithm. To evaluate the accuracy of the
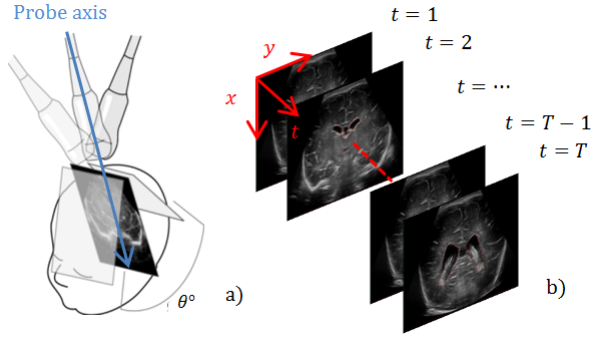
Fig. 1. Representation of a clinician manual sweep of angle θ (a) and an acquired image sequence composed of T images in coronal view (b).

segmentation and to train the convolutional neural network, manual CVS segmentation of the 15 volumes were realized by the clinician. The segmentation were performed on the 2D angular sweep sequence and reconstructed using the algorithm that will be presented in the next subsection.

B. 3D reconstruction

First, to cope lateral probe displacements and neonate head movements, each 2D image of the sequence was realigned to the following one using a 2D rigid registration routine. The probe motion was then modeled using Eq. 1.

$$
\begin{aligned}
f: \quad I \quad &\rightarrow \quad I' \\
(x,t) \quad &\mapsto \quad \begin{pmatrix} C_{x'} + v_{x'}t + (x + \Delta)\cos(\omega t + \theta_0) \\ C_{z'} + v_{z'}t + (x + \Delta)\sin(\omega t + \theta_0) \end{pmatrix}
\end{aligned}
\tag{1}
$$

The model was designed to represent a rotation with a rotation center translating at constant speed. The interpretation of the parameters is given in Table I, axes x' and z' are defined on the reference sagittal image (Fig. 4a).

TABLE I
INTERPRETATION OF THE MODEL PARAMETERS

Parameter	Interpretation
$C_{x'}$	Initial x' position of the probe rotation center
$C_{z'}$	Initial z' position of the probe rotation center
$v_{x'}$	Velocity in x' of the probe rotation center
$v_{z'}$	Velocity in z' of the probe rotation center
Δ	Distance between the probe rotation center and an image
ω	Angular velocity
θ_0	Initial angular position of the probe

To determine the model parameters, we manually placed anatomical landmarks (red dots in Fig. 4a and blue dots in Fig. 4c) on the reference image (I') and on one or several 2D sagittal views (I) of the image sequence. The model parameters were then chosen such that the set of landmarks $S = \{(x_i, t_i) \in I, i \in [[1, N]]\}$ match as best as possible the set of targeted landmarks $S' = \{(x'_i, z'_i) \in I', i \in [[1, N]]\}$ after reconstruction. The parameters optimization was performed

by minimizing the sum of the Euclidean distance d (Eq. 2) between S' and $f(S)$. All parameters were initialized to 0 and minimization was realized with gradient descent.

$$
\underset{C_x, C_t, v_x, v_t, \Delta, \omega, \theta_0}{\arg\min} \sum_{i=1}^{N} d(f(x_i, t_i), (x'_i, z'_i))
\tag{2}
$$

Once the parameters optimization was performed, the images of the angular sequence were mapped onto a 3D Cartesian grid using the optimal model parameters. The volume interpolation was performed using a voxel based method [9] with a spherical weighted interpolation of radial $\sqrt{5}$. The 3D reconstructed volume were finally manually centered on the corpus callosum splenium and rotated to the MRI coordinate system.

C. Segmentation

To achieve the segmentation of the 3D reconstructed volumes, we trained a convolutional neural network with the U-Net architecture described in Fig. 2. From the 15 reconstructed volumes, we extracted $256 \times 256 \times 256$ sub-volumes centered on the corpus callosum splenium. Respectively 8, 3 and 4 volumes from 7, 3 and 4 different patients were used for training, validation and test. A dropout layer with $p = 0.5$ was used just before the last layer. The network was trained with batches composed of 10 randomly drawn 128×128 coronal sub-images and the network parameters optimization was performed by minimizing soft-DICE loss, as it was reported to give better segmentation results for one class classification problem [8]. The stochastic batch gradient descend was performed with Adam optimizer and the gradient step was set to 1^{-4} during the first 10000 steps, 2^{-5} during the following 10000 steps and 5^{-6} until the end of training. The loss was calculated over the entire training and validation set every 1000 steps. We stopped training when the loss had not decreased for 10000 steps (early stopping) and we kept the network that gave the last best performance over the validation set. Training and validation curves can be seen on Fig. 3.
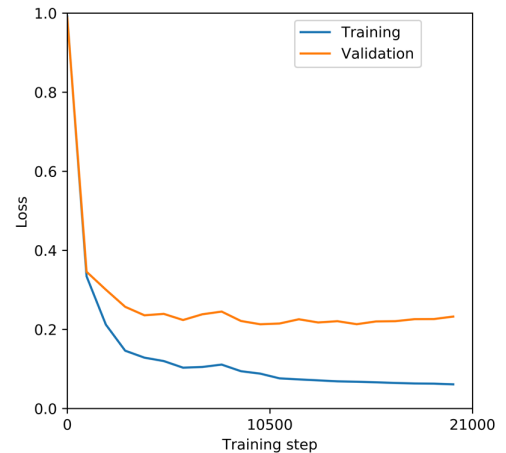

Fig. 3. Training and validation loss as a function of the number of training steps
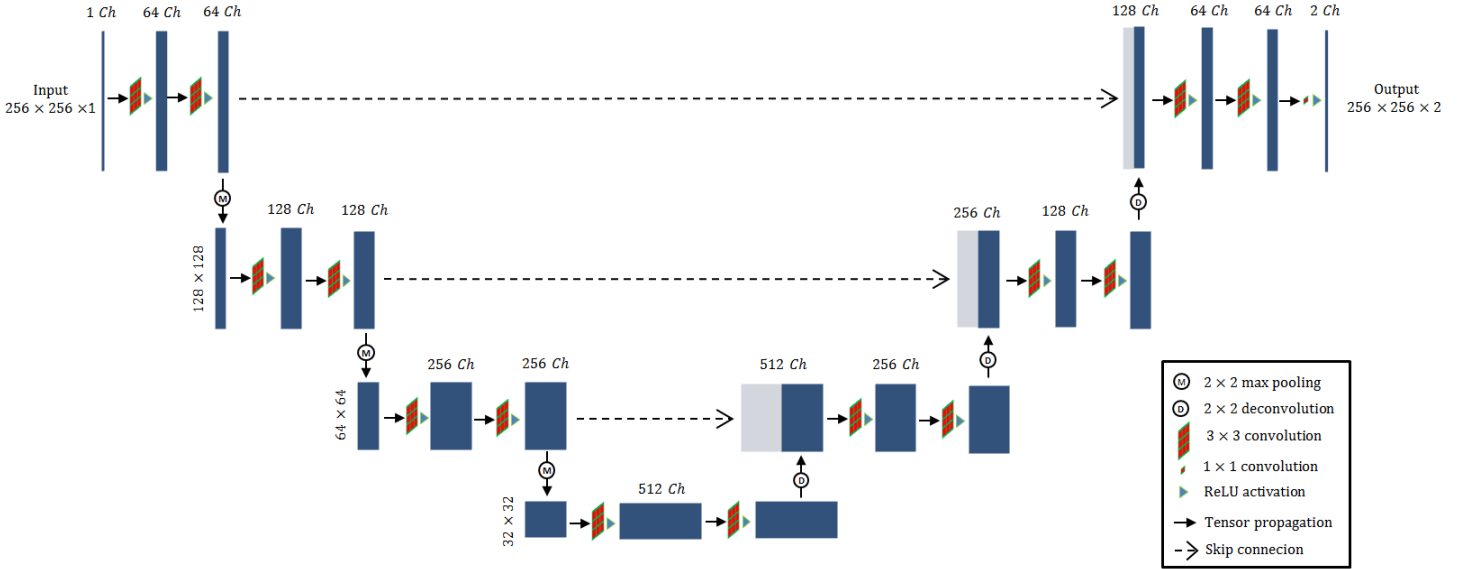
Fig. 2. Architecture of the trained U-Net, the tensors (represented in blue) spatial dimensions are given on the left of each floor and the tensor channel dimension is written on the top of each tensor.

The network was implemented on python using tensorflow [10]. As a post-processing, we extracted the 5 biggest connected binary region with a volume greater than 1% of the database mean CVS volume and applied a Gaussian filter of size $9 \times 9 \times 9$.

III. RESULTS

To evaluate the reconstruction accuracy we calculated MAD (as defined by Eq. 5) between a manual segmentation of CC on the reference sagittal image (Fig. 4d) and a manual segmentation of CC on its corresponding saggital view after 3D reconstruction (Fig. 4e). We obtained a good mean value of **1.55 $\pm$ 1.59 mm**. Only one sagittal image was used to evaluate the accuracy of the reconstruction because the model reconstruction f does not depend on the sagittal coordinate y. We can then assume that the same performance is approximately found on every sagittal view. In addition, CC is located between the left and right cerebral ventricles so it gives an idea of the error made in the CVS reconstruction.

The segmentation accuracy was calculated over the entire test set. We used DICE, d_H and MAD as respectively defined by Eq. 3, 4 and 5 to quantify the performance of the segmentation.

$$DICE = 2 \times \frac{|P \cap Q|}{|P| + |Q|} \quad (3)$$

$$d_H = \max \left(\min_{x \in \partial P} d(x, \partial Q), \min_{x \in \partial Q} d(x, \partial P) \right) \quad (4)$$

$$MAD = \frac{1}{2} \left(\frac{\sum_{x \in \partial P} \min d(x, \partial Q)}{|\partial P|} + \frac{\sum_{x \in \partial Q} \min d(x, \partial P)}{|\partial Q|} \right) \quad (5)$$

P and Q respectively stand for the binary reference volumes and segmented volumes, ∂P and ∂Q corresponds to their respective borders and $|X|$ to the number of elements in a set X. All quantitative segmentation results (after post-processing) are given in Table II. Qualitative results are shown on Fig. 5.

TABLE II
SEGMENTATION ALGORITHM PERFORMANCE

	DICE	Hausdorff (mm)	MAD (mm)
Mean results	0.816 ± 0.04	13.6 ± 4.7	0.62 ± 0.2

We obtained respective mean value of 0.816 ± 0.04, 13.6 ± 4.7 and 0.62 ± 0.2 for DICE, d_H and MAD. As it is shown by Fig. 3, the training loss is low, which means that the architecture has a good capacity to learn how to segment the CSV on the training cases. However the generalization gap (difference between training and validation loss) remains high and could be explained by the low number of training volumes. The segmentation of one $256 \times 256 \times 256$ volumes (16 millions voxels) was achieved in 5 s using a NVIDIA GeForce GTX 1080 graphic card with 8 Go RAM.

IV. CONCLUSION

This work proposes a 3D reconstruction algorithm to reconstruct 2D freehand angular manual sweep which allows the visualization of the data in axial, coronal and sagittal plan. The segmentation of the 3D CVS automatic segmentation was achieved in a remarkable time with a good accuracy according to all the metrics used. We plan to study the performance of 3D CNN to segment CVS and other cerebral structures in 3D US Data. This preliminary work shows that CNN can be
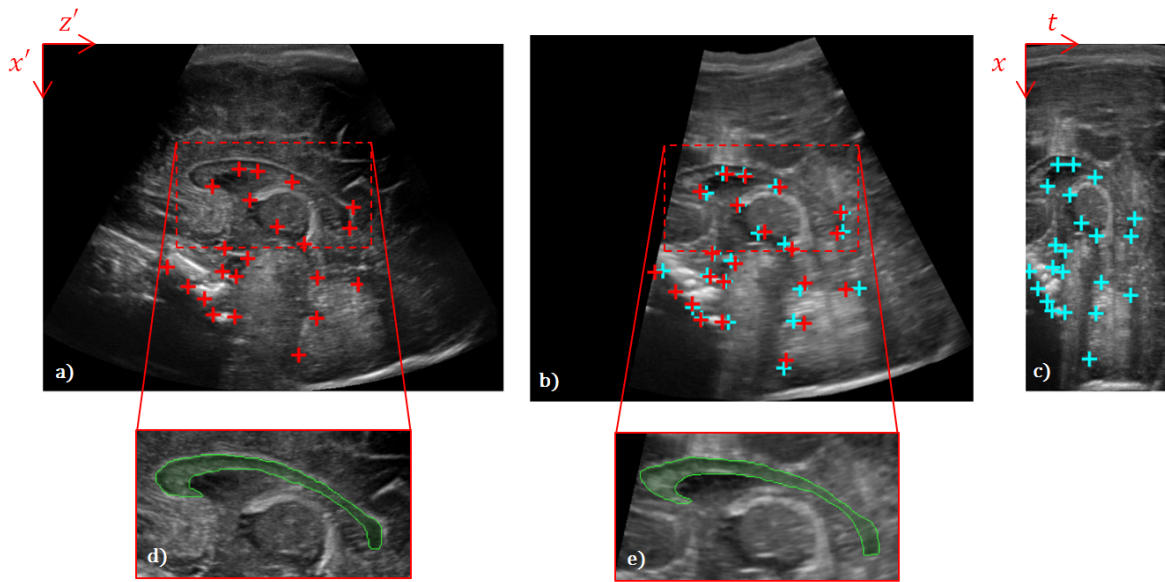
Fig. 4. a) : reference sagittal image, b) : reconstructed volume in sagittal view, c) : image sequence in sagittal view. The red dots on image a) and b) correspond to the landmarks set S', the blue dots on image c) to the landmarks set S and the blue dots on image b) to the landmarks set $f(S)$. d), e) : manual segmentation (green area) of the CC used to evaluate the accuracy of the 3D reconstruction algorithm.
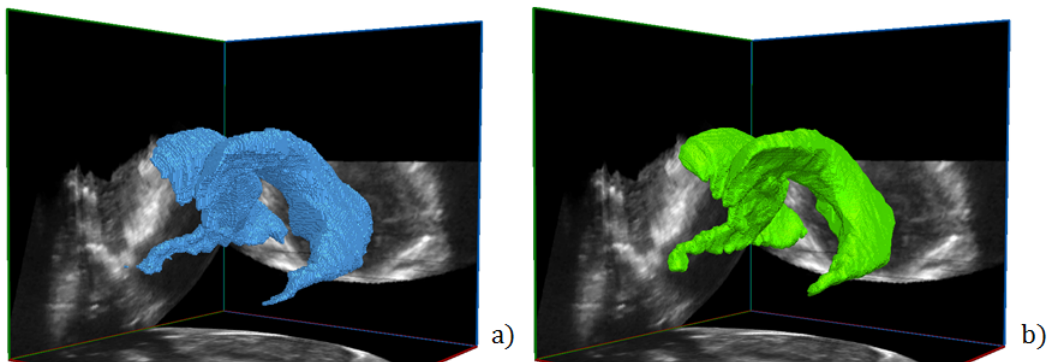

Fig. 5. 3D representation of the CVS : a) manual reference drawn by the clinician, b) automatic segmentation

used to segment CVS in 3D US volume and open the way to quantitative 3D US imaging of the preterm neonate brain.

REFERENCES

[1] Maunu, J., et al.: Ventricular dilatation in relation to outcome at 2 years of age in very preterm infants: a prospective Finnish cohort study. Developmental Medicine & Child Neurology, 53(1), 48-54 (2011).

[2] Lind, A., et al.: Associations between regional brain volumes at term-equivalent age and development at 2 years of age in preterm children. Pediatric radiology, 41(8), 953-961 (2011).

[3] Kurian, J., Sotardi, S., Liszewski, M. C., Gomes, W. A., Hoffman, T., & Taragin, B. H. Three-dimensional ultrasound of the neonatal brain: technical approach and spectrum of disease. Pediatric radiology, 47(5), 613-627 (2017).

[4] Qiu, Wu, et al.: Automatic segmentation approach to extracting neonatal cerebral ventricles from 3D ultrasound images. Medical image analysis 35 : 181-191 (2017).

[5] Boucher, M. A., Lippe, S., Damphousse, A., El-Jalbout, R., & Kadoury, S. Dilatation of Lateral Ventricles with Brain Volumes in Infants with 3D Transfontanelle US. arXiv preprint arXiv:1806.02305 (2018).

[6] Puyang Wang, N., et al.: Automatic real-time CNN-based neonatal brain ventricles segmentation. In : IEEE 15th Internation Symposium on Biomedical Imaging (2018).

[7] Ronneberger, O., et al.: U-net: Convolutional networks for biomedical image segmentation. In International Conference on Medical image computing and computer-assisted intervention, pp. 234-241 (2015).

[8] Milletari, F., et al.: V-net: Fully convolutional neural networks for volumetric medical image segmentation. In Fourth International Conference on 3D Vision, pp. 565-571 (2016).

[9] Solberg, O. V., Lindseth, F., Torp, H., Blake, R. E., & Hernes, T. A. N. Freehand 3D ultrasound reconstruction algorithmsa review. Ultrasound in medicine & biology, 33(7), 991-1009 (2007).

[10] Martn Abadi & al. TensorFlow: Large-scale machine learning on heterogeneous systems (2015). Software available from tensorflow.org.