Automatic segmentation of the optic nerve in transorbital ultrasound images using a deep learning approach

Automatic segmentation of the optic nerve in transorbital ultrasound images using a deep learning approach

Kristen M. Meiburger
Biolab, Polito^{BIO}Med Lab, Department of Electronics and Telecommunications
Politecnico di Torino
Torino, Italy
kristen.meiburger@polito.it

Andrea Naldi
Department of Neurosciences
Università di Torino
Torino, Italy
andrea.naldi@aslcittaditorino.it

Piergiorgio Lochner
Saarland University Medical Center
Homburg, Germany
piergiorgio.lochner@uks.eu

Francesco Marzola
Biolab, Polito^{BIO}Med Lab, Department of Electronics and Telecommunications
Politecnico di Torino
Torino, Italy
francesco.marzola@polito.it

***Abstract*—Transorbital sonography is able to provide reliable information about (a) intra-cranial pressure estimation through the optic nerve sheath diameter (ONSD) measurement, and (b) optic nerve atrophy in patients with multiple sclerosis through the optic nerve diameter (OND). In this study, we present the first method for the automatic measurement of the OND and ONSD using a deep learning technique (UNet with ResNet50 encoder) for the optic nerve segmentation. The dataset included 201 images from 50 patients. The automated measurements were compared with manual ones obtained by one operator. The mean error was equal to 0.07 ± 0.34 mm and -0.07 ± 0.67 mm, for the OND and ONSD, respectively. The developed system should aid in standardizing OND and ONSD measurements and reduce manual evaluation variability.**

Keywords—optic nerve diameter, optic nerve sheath diameter, segmentation, deep learning, CNN

I. Introduction

Transorbital sonography (TOS) allows the non-invasive evaluation of the optic nerve (ON) structures and has been used mainly for the assessment of the optic nerve sheath diameter (ONSD) for estimating and monitoring increased intra-cranial pressure in many neurological disorders including traumatic brain injury, coma, ischemic and hemorrhagic stroke, idiopathic intracranial hypertension, hydrocephalus, posterior reversible encephalopathy [1], [2]. In addition, TOS has found application in the detection of ON atrophy in patients with multiple sclerosis [3] by measuring the ON diameter (OND), thus opening the way for a wider use of the technique also in neurodegenerative diseases.

Manual evaluation of the OND and ONSD have shown to have good intra- and inter-observer reproducibility when using high frequency (> 7MHz) linear probes [4], but it is also highly affected by the operator's expertise [5]. Fig. 1 shows an example of how the manual computation of the OND and ONSD are done.

Deep learning segmentation methods based on convolutional neural networks (CNNs) have recently become the go-to technique in the medical image analysis field. The main advantage of CNNs is that they automatically learn high-level features on the image and then can provide a semantic segmentation by associating each pixel to a label. One of the main drawbacks to deep learning methods is the need of a large annotated database, which has somewhat, but not totally, been mitigated with the employment of transfer learning [6].

Recently, there have been a number of methods presented in literature to aid the computation of the OND and ONSD, using either semi-automatic or completely automatic techniques [7]–[11]. The majority of these studies, however, utilize a rather limited database size with a maximum of 88 images, presented in the study by Rajajee et al. [11], and all methods are based on traditional image-processing techniques, which are often fine-tuned to the specific database that is used entirely as both a training and a test set.

In this study, we present an automatic segmentation method of the optic nerve in transorbital ultrasound images employing a large dataset and using a deep learning approach for the first time.

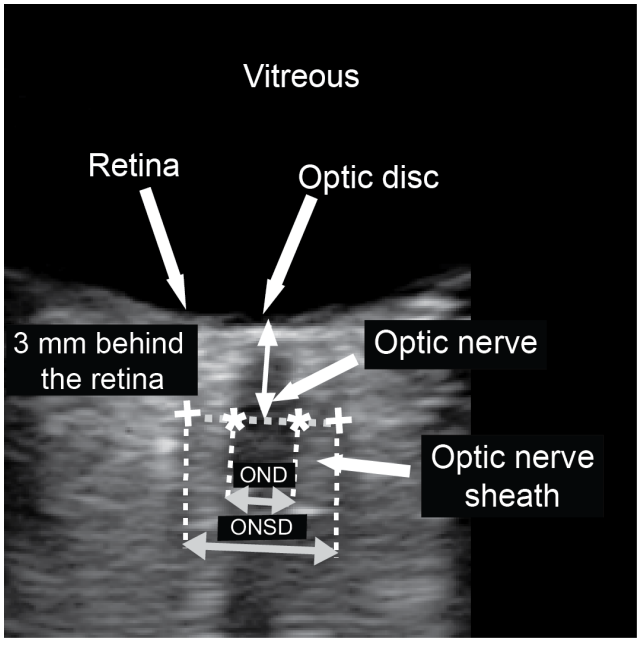

Fig. 1. Example of manual optic nerve diameter (OND) and optic nerve sheath diameter (ONSD) measurement.

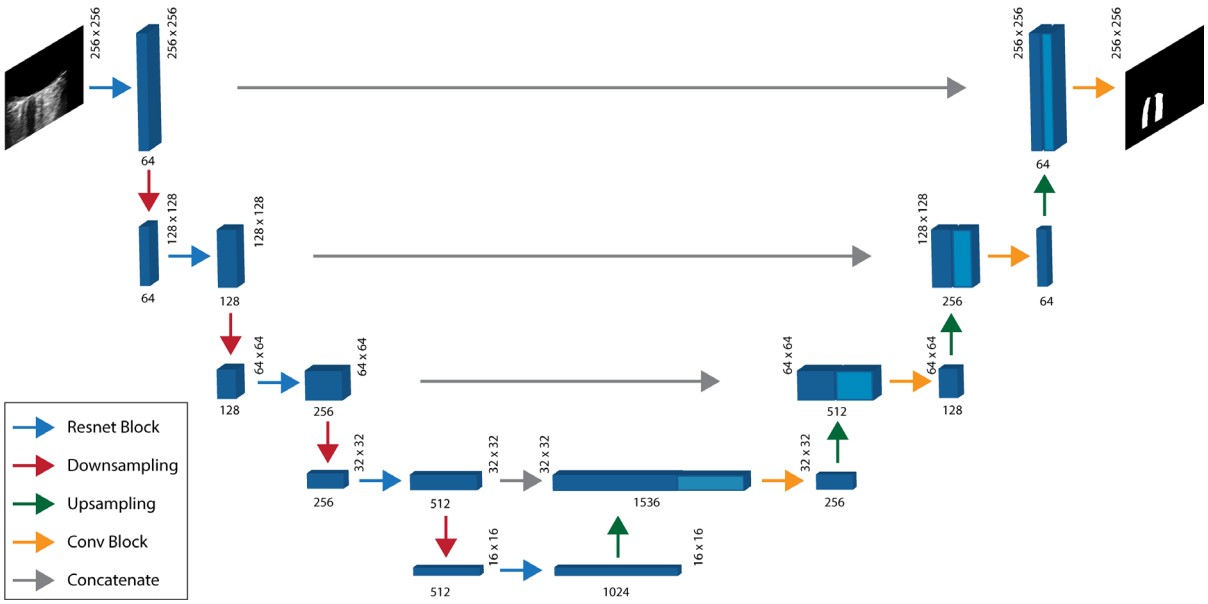

Fig. 2. Representation of the Unet with ResNet50 encoder deep learning model.

II. Materials and Methods

A. Dataset and CNN architecture

The ultrasonographic studies were performed with 5 different ultrasound machines (MyLab Gold 30 and MyLab Seven, Esaote, Genova, Italy; Toshiba Medical System Aplio 300 and AplioXG, Nasu, Japan; Vivid 7 sonography system, GE Healthcare, Milwaukee, USA), equipped with linear transducers in the 3-11 MHz or 7.2-14 MHz frequency range and a lateral resolution of <0.4mm. The dataset employed in this work contains 201 images obtained from 50 subjects. Written consent was obtained from each participant. Since the size of the dataset is too limited to have an unbiased estimation of the model performance, we divided the dataset into 5 folds, and then performed cross-validation. For each run, we used 3 folds for training, one for validation, and one for testing. The number of images in each fold ranges from 31 to 46 images because not all the subjects have the same number of images and we enforced that all the images of a subject must belong to the same fold. The images were automatically cropped into 256x256 pixel squares containing the optic nerve and were input to the network. The automatic cropping was done searching for anechoic circular structures in the image by the circular Hough transform.

The CNN network that was employed was a Unet with a ResNet50 Encoder pretrained on the ImageNet dataset [12]; a representation of the model is shown in Fig. 2.

We applied on the fly augmentation performing random affine transformations, sharpening, and blurring of the images. The network was trained using Adam's optimization algorithm with an initial learning rate of 0.0002, batch size of 8, and Dice Loss as the objective function. The learning rate decreases by a factor of 10 every 10 epochs without improvements in the IoU, and the training stops if this condition reaches 25 epochs, the maximum number of epochs is 50. For the network implementation and training, we adopted the same code we used in our previous work [1]. For each fold, the network reached its optimal state between epoch 33 and epoch 49, with a training time of under 10 minutes using an Nvidia RTX 3070 with 8 GB of VRAM. The loss path on the validation set for each fold is visualized in Fig. 3.

The output of the network is then post-processed to ensure the presence of only the left and right ON sheaths, maintaining only the two larger connected areas.

B. Validation

The deep learning approach produces a binary mask of the optic nerve structure as an output. To validate the model, two approaches were employed: a first one compared the manual masks with the network output masks using the Precision, Recall, and Dice parameters, limiting the manual and automatic segmentation comparison only in the image rows where both the manual and the automatic segmentation are present. Secondly, since the parameters that are of actual clinical use are the OND and ONSD, the manual and automatic masks were analyzed to automatically compute the diameters. This was done by a custom-made code in Matlab 2020b, and is based on (1) locating the optic nerve centerline thanks to the binary masks and finding the line perpendicular

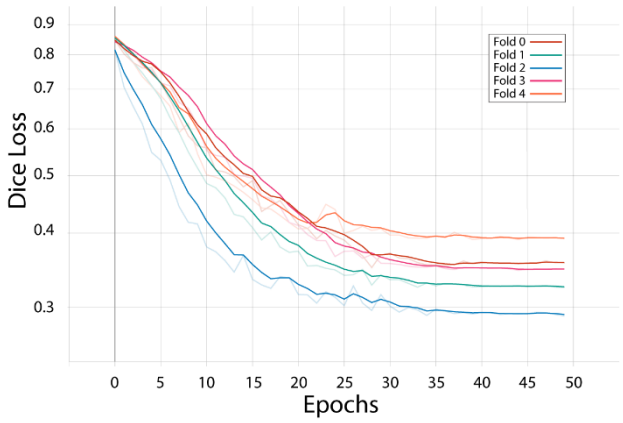

Fig. 3. Loss path on the validation set for each fold

to it, (2) locating the optic bulb as a circular anechoic region within the image, and finally (3) computing the Euclidean distance between the intersection points between the line perpendicular to the centerline and the segmentation mask at a 3mm distance below the optic bulb.

III. Results

Merging the test sets in each fold we achieved a Dice score of 0.79 ± 0.11, a Precision of 0.77 ± 0.13, and a Recall equal to 0.82 ± 0.12. The mean OND error was equal to 0.07 ± 0.34 mm. The mean absolute error for the ONSD measurements was equal to -0.07 ± 0.67 mm.

We performed a Wilcoxon paired test to investigate the differences in the OND and ONDS measurements between a human operator and the automatic algorithm, resulting in a p-value of 0.004 and 0.057 respectively.

Fig. 4 portrays an example of the obtained segmentation and diameter computation results.

IV. Discussion and Conclusions

This study is the first in literature that employs a deep learning technique for the segmentation of the optic nerve in transorbital ultrasound images. The proposed method is fully automated, starting from the optimal cropping of the image that is input to the deep network, to the post processing computation of the OND and ONSD. The total implementation time of the CNN, including the training and testing for each of the 5-folds was equal to 20 minutes.

The obtained mean errors and mean absolute errors are comparable to those obtained in literature in previous studies. In particular, in our previous study [10], we presented a method based on dual snakes that obtained a mean error for the OND and ONSD equal to 0.06 ± 0.35 mm and 0.06 ± 0.522 mm, respectively. More recently, Rajajee et al. [11] presented a method based on clustering and the mean error was equal to 0.012 ± 0.046. Table 1 shows a comparison between our obtained results and other recent clinical studies focusing on optic nerve segmentation in ultrasound images. We can see how here we show results for the largest dataset (i.e., 201 images) acquired with different ultrasound devices and in different centers. Moreover, the deep learning model was trained and tested on completely different images, showing unbiased and more robust results. The other studies in literature do not mention the division of their datasets into training/testing, so it is assumed that the training and test set coincide, which can polarize results and make the developed system less robust to new images.

TABLE I. Segmentation comparison

	Segmentation comparison results			
	Data-set	***Method***	***OND error (mm)***	***ONSD error (mm)***
[7]	23 images	Threshold + distance transform	-	Pearson Correlation range = 0.35 – 0.95
[8]	50 videos	Super pixel analysis	-	MSE = 0.0018
[9]	42 images	Assymetry features + active contour	-	ME = -0.08 ± 0.45 ME = -0.05 ± 0.41
[10]	75 images	Dual snakes	ME = 0.06 ± 0.35 MAE = 0.28 ± 0.22 MSE = 0.12 ± 0.17	ME = 0.06 ± 0.52 MAE = 0.41 ± 0.32 MSE = 0.27 ± 0.37
[11]	88 images	Clustering	-	ME = 0.012 ± 0.046
Unet	201 images	UNet with ResNet 50 encoder	ME = 0.07 ± 0.34 MAE = 0.26 ± 0.23 MSE = 0.12 ± 0.20	ME = -0.07 ± 0.67 MAE = 0.48 ± 0.48 MSE = 0.45 ± 0.93

[a.] ME: Mean error; MAE: Mean absolute error; MSE: Mean square error.

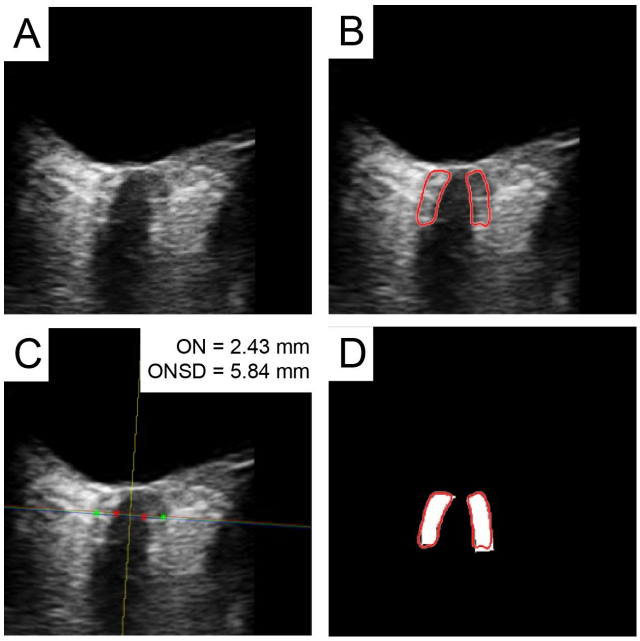

Fig. 4. Example segmentation results. A) Original 256x256 pixels image. B) Deep learning network output overlaid on original image. C) Automatically determined centerline (yellow line), points used for optic nerve diameter measurement (red asterisks), points used for optic nerve sheath diameter measurement (green asterisks). D) Automatic segmentation (red borders) overlaid on manual segmentation mask.

This study presents some limitations. First of all, while the employed dataset is larger than those in previously published studies, it is still limited, especially for the application of deep learning methods. Secondly, we did not present any inter- or intra-operator variability study. In fact, in literature there is still a variability of manual OND and ONSD measurements, which mainly reflect operator experience, an un-standardized image acquisition and measurement method [13], and the use of different ultrasound devices used for image acquisition. Due to this un-standardization, we recognize that the actual OND and ONSD measurements could be considered untrustworthy, due to the blooming effect in ultrasound images that is evident especially when measuring small structures in the absence of a standard gain setting [14]. The smallness of the measured structure also directly influences the OND and ONSD error measurements. In fact, the calibration factor (mm/pixel) in the images in our dataset ranged from 0.0316 to 0.1210, which demonstrates how one pixel more or less in the computation of the automatic vs. the manual diameter has to potential to have a large impact on the obtained

performance. The measured diameters by the deep learning method compared to the manual diameters showed error values that are still not ideal. This is noticeable especially with the ONSD measurement, where there is a bias by the automated method that produces a mean absolute error of 0.48 mm. Still, the error results obtained are smaller than the 1mm difference that is usually expected between subjects with normal or elevated ICP. Hence, an approach like the one proposed should be able to distinguish these patient groups, although not specifically demonstrated here.

Our future studies will focus on continuing to increase the dataset size and including an inter- and intra-operator variability study. Moreover, other deep learning architectures and the potential combination of methods to obtain an Ensemble network will be investigated, and we aim to reduce the network's error measurements when computing the final OND and ONSD. Moreover, we would like to look into the influence of different gain and device settings on the network's segmentation performance.

Overall, the results are encouraging and the use of an automated artificial intelligence system such as the one proposed here will hopefully promote the standardization of OND and ONSD measurements and in the future provide a more comparable interpretation of results among studies.

REFERENCES

[1] C. Robba, S. Bacigaluppi, D. Cardim, J. Donnelly, A. Bertuccio, and M. Czosnyka, "Non-invasive assessment of intracranial pressure," *Acta Neurol. Scand.*, vol. 134, no. 1, pp. 4–21, Oct. 2015.

[2] I. Soliman *et al.*, "New Optic Nerve Sonography Quality Criteria in the Diagnostic Evaluation of Traumatic Brain Injury," *Crit. Care Res. Pract.*, vol. 2018, pp. 1–7, Apr. 2018.

[3] N. Carraro, G. Servillo, V. Maria Sarra, A. Bignamini, G. Pizzolato, and M. Zorzon, "Optic Nerve and its Arterial-Venous Vascularization: An Ultrasonologic Study in Multiple Sclerosis Patients and Healthy Controls," *J. Neuroimaging*, vol. 24, no. 3, pp. 273–277, May 2014.

[4] P. Lochner *et al.*, "Optic nerve sheath diameter: present and future perspectives for neurologists and critical care physicians," *Neurol. Sci.*, pp. 1–11, Jul. 2019.

[5] R. Copetti and L. Cattarossi, "Optic nerve ultrasound: artifacts and real images," *Intensive Care Med.*, vol. 35, no. 8, pp. 1488–1489, 2009.

[6] S. ialin Pan and Q. Yang, "A Survey on Transfer Learning," *IEEE Trans. Knowl. Data Eng.*, vol. 22, no. 10, pp. 1345–1359, 2010.

[7] S. Gerber *et al.*, "Automatic Estimation of the Optic Nerve Sheath Diameter from Ultrasound Images BT - Imaging for Patient-Customized Simulations and Systems for Point-of-Care Ultrasound," 2017, pp. 113–120.

[8] R. Soroushmehr *et al.*, "Automated Optic Nerve Sheath Diameter Measurement Using Super-pixel Analysis," in *41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2019, pp. 2793–2796.

[9] R. R. F. Stevens *et al.*, "An Automated Algorithm for Optic Nerve Sheath Diameter Assessment from B-mode Ultrasound Images," *J. Neuroimaging*, vol. 31, no. 4, pp. 724–732, 2021.

[10] K. M. Meiburger *et al.*, "Automatic Optic Nerve Measurement: A New Tool to Standardize Optic Nerve Assessment in Ultrasound B-Mode Images," *Ultrasound Med. Biol.*, vol. 46, no. 6, pp. 1533–1544, 2020.

[11] V. Rajajee *et al.*, "Novel Algorithm for Automated Optic Nerve Sheath Diameter Measurement Using a Clustering Approach," *Mil. Med.*, vol. 186, no. Supplement_1, pp. 496–501, 2021.

[12] Jia Deng, Wei Dong, R. Socher, Li-Jia Li, Kai Li, and Li Fei-Fei, "ImageNet: A large-scale hierarchical image database," pp. 248–255, 2009.

[13] S. Bloria, P. Bloria, and A. Luthra, "Is it the time to standardize the procedure of ultrasound guided optic nerve sheath diameter measurement?," *Saudi J. Anaesth.*, vol. 13, no. 3, p. 255, 2019.

[14] M. De Bernardo and N. Rosa, "Transorbital sonography to evaluate optic nerve in hypertensive encephalopathy," *J Stroke Cerebrovasc Dis*, vol. 27, p. 1124, 2018.