Automatic Vessel Recognition and Segmentation: a Novel Deep Learning Architecture with Transfer Learning Approach

Giovanni Faoro, Stefano Pane, Veronica Iacovacci, Sara Moccia and Arianna Menciassi

May 28, 2022

Automatic Vessel Recognition and Segmentation: a novel Deep Learning Architecture with Transfer Learning Approach

Giovanni Faoro, Stefano Pane, Veronica Iacovacci, Sara Moccia, Arianna Menciassi

The BioRobotics Institute, Scuola Superiore Sant'Anna, Pisa, Italy

Department of Excellence in Robotics and AI, Scuola Superiore Sant'Anna, Pisa, Italy

giovanni.faoro@santannapisa.it

INTRODUCTION

Ultrasound (US) imaging stands as a valid alternative to X-rays based methodologies for navigation and intraoperative tracking of vascular probes, thanks to its non-ionizing nature. However, US images quality is highly operator dependent, being subject to probe's orientation and contact force. In recent years, researchers have worked to develop Robotic US Systems (RUSS), granting the acquisition of good quality real-time US images, without the need of an expert operator [1]. Besides, to facilitate US images analysis, deep learning strategies have been developed. Applications in the field include the automatic segmentation of vessels, which is fundamental during endovascular procedures.

Intraoperatively, an automatic method to classify images based on the presence of vessels and selectively segment only vascular images would be valuable. For example, during hand-held probe procedures it would increase the quality of information feedback. In RUSS, it would enable automatic adjustment of probe positioning serving as alternative to manual positioning by highly trained sonographers. Additionally, a method for precisely discriminating the presence of vessels in the image plane could increase safety in visual-servoing platforms, by preventing possible control instabilities generated by imaging artifacts. However, segmentation architectures typically assume that the processed image contains vessels to be segmented [2], but this is not granted in real intraoperative settings especially at the beginning of the procedure when the imaging probe is not yet optimally positioned.

To address these unmet needs, in this paper we propose a multi-task convolutional neural network (CNN) architecture able to distinguish between vessel and no vessel images, in addition to segmenting them. The goal of such architecture is to enable robust and automatic US images analysis in real intraoperative settings.

MATERIALS AND METHODS

One of the most common deep learning architectures for medical image segmentation is U-Net, which is characterized by a contracting (encoder) and an expansion (decoder) path [3]. To accomplish our goal, we built a multi-task architecture, as a modified version of U-Net (Fig. 1), by adding a classification branch after the contracting path that is able to detect the presence of vessels in the image. The classification branch is made of flatten layer, dense layer (activated with the rectified linear unit), dropout (with 0.5 probability), dense layer (activated with the hyperbolic tangent function), dropout (with 0.5 probability) and dense layer (activated with the softmax function). In our multi-task architecture, the contracting path is shared among classifier final layers and segmentation decoder, minimizing the computational cost.

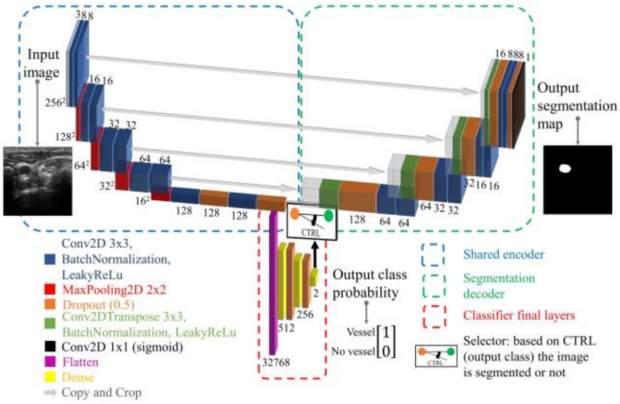

Fig.1 Our multi-task CNN architecture for vessel detection and segmentation from US images. The encoder path is shared to minimize computation; based on the output class probability the input image is further processed for vessel segmentation or not.

Our multi-task CNN is fed with images with size 256x256 pixels and provides as output a 2x1 probability vector, representing the class probability (vessel or no vessel image), and a 256x256 probability map, representing the segmentation output. Classification and segmentation branches were trained separately, considering categorical cross-entropy (for the classification task) and Dice loss (for the segmentation task). The Dice loss is defined as 1- Dice Similarity Coefficient (DSC), which is computed as two times the cardinality of output and ground truth intersection over the sum of cardinalities. To tackle overfitting issues relevant to the relatively small size of our dataset, we decided to exploit transfer learning [4].

We collected two small datasets of common carotid artery (CCA) US images available online[1]: 240 B-mode

[1] https://splab.cz/en/research/zpracovani-medicinskychsignalu/databaze/artery, 17/01/2022, 16:50
https://www.researchgate.net/publication/261703132_100-IMT-Images_of_the_CCA, 17/01/2022, 16:50

images (120 with no vessel, 60 with vessels in long axis, 60 with vessels in short axis) were used to train the classifier, 240 B-mode images (120 with vessels in long axis, 120 with vessels in short axis) to train the segmentation network. Gold-standard (GS) annotation was obtained semi-automatically by using GIMP environment (GNU Image Manipulation Program 2.10.28). The images from the datasets had different size and were resized to 256x256 pixels.
Initial CNN weights were retrieved from a state-of-the-art U-Net architecture trained on US images [5], since transfer learning is expected to achieve better performances when original and target tasks are similar [4]. The encoder layers were frozen and the two datasets were used to fine-tune the classification branch and the U-Net decoder path.

The multi-task CNN was implemented with TensorFlow and Keras libraries and ran on NVIDIA Tesla K80 GPU provided by Google Collaboratory. The classifier was trained with SGD optimizer (learning rate: 0.001, momentum: 0.8). The segmentation branch was trained with Adam optimizer with a learning rate of 0.03, adapted during learning with an InverseTimeDecay routine. Batch size was set to 20 and the maximum number of epochs to 100, but early stopping with a patience of 10 was used to avoid overfitting. 6-fold cross-validation was used for robust testing. For each fold, training images were doubled in number by applying horizontal flip. Training images were further split, after shuffle, into training set (80%) and validation set (20%). Supplementary data augmentation techniques, as random rotation, zoom and shift, were applied online during training.

Architecture performances were evaluated defining True Positive (TP), True Negative (TN), False Negative (FN) and False Positive (FP) outputs in accord to the GS annotation. These were used for the computation of standard metrics as: accuracy (defined as TP+TN over the sum of all) and F1 score (defined as 2TP over 2TP+FP+FN) for the classification task; DSC for the segmentation task.

RESULTS

The average accuracy and F1 score for the classification task were 93.54% and 92.06%, respectively, with an average computation time on Google Colab of 5.01±0.61 ms per image (Fig. 2a). The average DSC for the segmentation task was higher than 90%, i.e., 92.22%, with an average computation time of 8.95±1.04 ms per image, enabling real-time applications (Fig. 2b).

DISCUSSION

Deep learning architectures able to perform vessel detection and segmentation together, as the one proposed here, would be valuable in real life dynamic environments, both in hand-held and robot-held probe procedures. By exploiting transfer learning, we were able to achieve promising segmentation performances (DSC equal to 92.22%) with a relatively small dataset (i.e., 240 images), compared to a previous work that shows CCA segmentation training a U-Net with over 2000 images [2]. Additionally, the proposed architecture allowed real-time deployment.

These preliminary results indicate that such multi-task CNN could be efficiently integrated in a robotic platform, potentially enabling robust visual-servoing procedures for, e.g., catheter navigation or aortic screening. Additionally, applications can be enlarged to different districts by further fine-tuning the network.

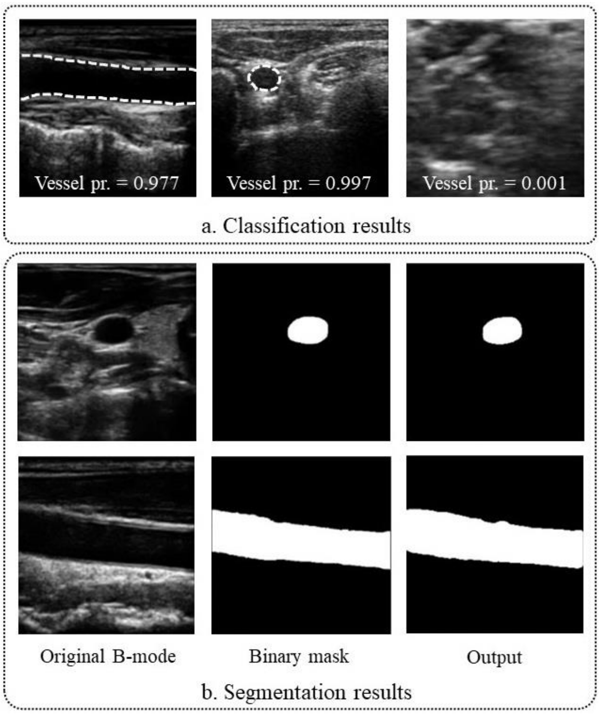

Fig. 2 Results in classification and segmentation tasks. a. Different B-mode images with the output vessel probability. From left to right, a TP long axis view, TP short axis view and a TN background images are shown. b. Segmentation results on a vessel short axis (first row) and long axis (second row) view.

REFERENCES

[1] Z. Jiang, Z. Li, M. Grimm, M. Zhou, M. Esposito, W. Wein, W. Stechele, T. Wendler and N. Navab, Autonomous Robotic Screening of Tubular Structures based only on Real-Time Ultrasound Imaging Feedback, *IEEE Transactions on Industrial Electronics*, 2021

[2] M. Xie, Y. Li, Y. Xue, L. Huntress, W. Beckerman, S. Rahimi, J. Ady, U. Roshan, Vessel lumen segmentation in carotid artery ultrasounds with the U-Net convolutional neural network, *IEEE International Conference on Bioinformatics and Biomedicine*, 2020

[3] N. Siddique, S. Paheding, C. P. Elkin, and V. Devabhaktuni, U-Net and its variants for medical image segmentation: A review of theory and applications, *IEEE Access*, vol. 9, pp. 82031-82057, 2021

[4] M. H. Hesamian, W. Jia, X. He, P. Kennedy, Deep Learning Techniques for Medical Image Segmentation: Achievements and Challenges, *J Digit Imaging*, 32:582 596, 2019

[5] Guo Y, Duan X, Wang C, Guo H, Segmentation and recognition of breast ultrasound images based on and expanded U-Net, *PLoS ONE* 16(6): e0253202, 2021