Localization of carotid arteries at traversal ultrasound images with Faster RCNN

Martin Kostelansky

Faculty of Electrical Engeneering
Czech Technical University
kostema3@fel.cvut.cz

Abstract

The computer vision has entered many domains in the recent years and together with deep learning are taking place in many different domains. In our work we aim to sort localize carotid artery at traversal ultrasound images. We compare two different approaches, including state-of-the-art Faster R-CNN and train them on Ultrasound image database dataset from Signal processing laboratory at VUT Brno and show that model trained on the data from multiple different scanners can still generalize for taken by a different machine.

I. Introduction

In the last decade, convolutional neural nets have been setting new benchmarks in many domains, namely image recognition [1, 2]. Many neural network architectures exist, such as recurrent CNN [3], very deep CNN [1], NN with depth-wise separable convolutions [4] and deep residual CNN [5], etc. This research ha been widely applied in many fields, including medicine and processing of biological images. Convolutional neural nets have shown their potential in segmentation[6], localization[7] and classification[8].

In our work, we use dataset containing thousands of carotid artery images. We aim reproduce results achieved in [9] and to use such model on our dataset.

II. Dataset

We use Ultrasound image database dataset from Signal processing laboratory at VUT Brno with labels describing the positions of carotid arteries. Model trained on this data aims to be used on our own dataset of traversal images of carotid arteries.

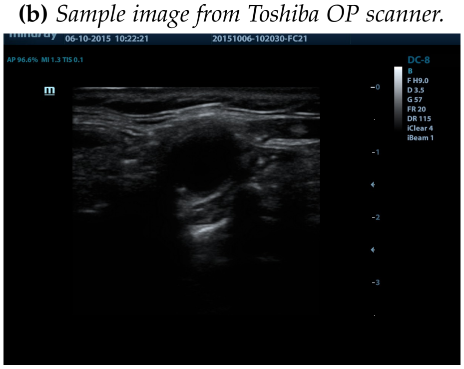
i. Data splitting

The data were split into three groups - train, validation and test. Train set contains 283 images s (388 x 400 pixels) and the validation one 538 images (388 x 400 pixels), both taken by Ultrasonix OP scanner. The test set is composed by 433 test images (283 x 322 pixels) from Toshiba Nemio XG scanner. This distribution follows was created by reproducing [9] (this will be later referenced as *DataSplit1*).

In the experiment, we created additional train-validation-test split combining data from different scanners, since we aim to use the model on different dataset, we expect this to increase the robustness of the model. The all 1254 images were split into train-validation-test with ratio 60-25-15 (this will be later referenced as *DataSplit2*).

(a) *Sample image from Ultrasonix OP scanner.*

(b) *Sample image from Toshiba OP scanner.*



(c) *Sample of our dataset.*

Figure 1: *Comparison of images taken by different scanners.*

III. Models

i. ResNet

We have taken an advantage of the fact, that there is only one artery at each image and created a model predicting only one set of coordinates. We have used Resnet50 [5] pretrained on the ImageNet dataset [10], where we have create new last fully-connected layer with four neurons and trained it to predict left-bottom and right-upper positions of a bounding box. Since the sizes of the training images differs, and for this architecture input needs to be of the shape 224x224, the labels need to be resized as well.

ii. Faster R-CNN

As the second approach we followed [9] in the setting of the Faster R-CNN. We trained it to predict only the location of an artery.

IV. Results and discussion

Both of the approaches we trained on both data splits. In the case of ResNet, the best model was selected by measuring Smooth L1 loss on the validation set and the Faster R-CNN by the percentage of validation images in which the IoU of labels and predicted values was higher tha 0.85. The same evaluation was used on the test dataset with two different thresholds - 0.6 and 0.85.

The test results are summarized in the Table 1. As we can see, the Resnet50 approach has struggled with this task and in the case of DataSplit1 was not able to learn such problem. When trained on the DataSplit2, the overall performance was improved by the fact that the images from both of scanners were present in the training set, but still has not achieved satisfiable results.

The Faster R-CNN have crunched the Resnet50 in the performance. When trained on the DataSplit1, although promising IoU on the validation set during training, it had problems to generalize, what was improved by training on the DataSplit2. By this setting model was

Model	DataSplit	IoU>0.6	IoU>0.85
hline	DataSplit1	0.2748	0.053
hline	DataSplit2	1.0	0.80
F R-CNN	DataSplit1	0.	0.
F R-CNN	DataSplit2	0.771	0.12

Table 1: *Evaluation of models on the test data.*

able to predict the coordinates of bounding boxes with IoU> 0.85 in eighty percent of the test samples.

Since our goal was to crate a model able to predict a position af an artery on completely different data, we used the best model on multiple target images. The results can be seen in the Figure 2. We clearly see that the some cases the model almost perfectly predicts the bounding boxes, but in on od the examples struggles.

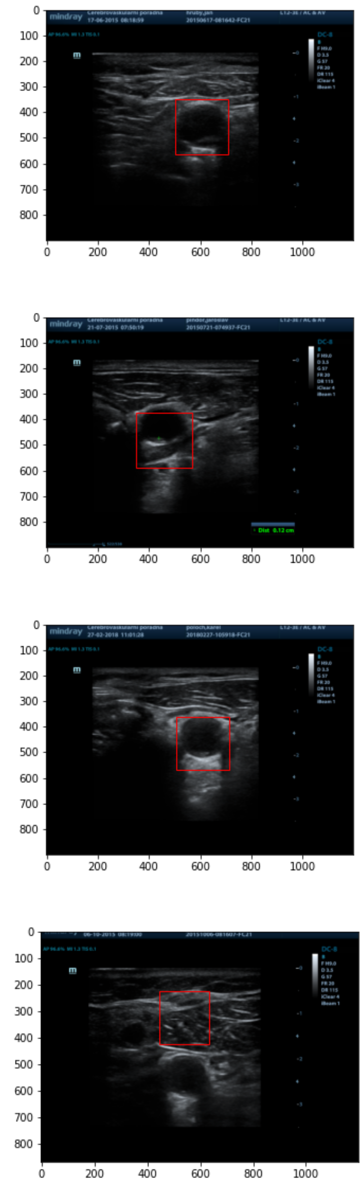


Figure 2: *Bounding boxes predicted by the best model on the target dataset.*

V. Future work

We can truly see the power of Faster R-CNN on localization tasks, but some of our results have not met the benchmarks achieved in [9]. We could fine-tune our model at the current setting or add a generator of background objects to create a negative samples in the training.

Also, to achieve maximal performance on our target dataset, we could create a labeled validation group from this dataset, and select the model during the training which has the best validation set.

VI. Implementation notes

This work has been implemented in Python 3.7 and the models were created using PyTorch 1.6. This project can be found on GitHub.

References

[1] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv 1409.1556*, 09 2014.

[2] S. Ghosh, I. Das, N. Das, and U. Maulik, "Understanding deep learning techniques for image segmentation," *ACM Computing Surveys*, vol. 52, p. 40, 05 2019.

[3] M. Liang and X. Hu, "Recurrent convolutional neural network for object recognition," pp. 3367–3375, 06 2015.

[4] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "Mobilenets: Efficient convolutional neural networks for mobile vision applications," *CoRR*, vol. abs/1704.04861, 2017.

[5] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *CoRR*, vol. abs/1512.03385, 2015.

[6] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," 2015.

[7] B. D. de Vos, J. M. Wolterink, P. A. de Jong, T. Leiner, M. A. Viergever, and I. IÅągum, "Convnet-based localization of anatomical structures in 3-d medical images," *IEEE Transactions on Medical Imaging*, vol. 36, no. 7, pp. 1470–1481, 2017.

[8] J. Chang, J. Yu, T. Han, H. Chang, and E. Park, "A method for classifying medical images using transfer learning: A pilot study on histopathology of breast cancer," in *2017 IEEE 19th International Conference on e-Health Networking, Applications and Services (Healthcom)*, pp. 1–4, 2017.

[9] P. K. Jain, S. Gupta, A. Bhavsar, A. Nigam, and N. Sharma, "Localization of common carotid artery transverse section in b-mode ultrasound images using faster rcnn: a deep learning approach," *Medical & Biological Engineering & Computing*, vol. 58, pp. 471–482, Mar 2020.

[10] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "ImageNet: A Large-Scale Hierarchical Image Database," in *CVPR09*, 2009.