Feasibility Study of Catheter Segmentation in 3D Frustum Ultrasounds by DCNN

Lan Min[a], Hongxu Yang[a], Caifeng Shan[b], Alexander F. Kolen[b], Peter H.N.de With[a]

[a]Eindhoven University of Technology, Eindhoven, The Netherlands
[b]Philips Research, Eindhoven, The Netherlands

ABSTRACT

Nowadays, 3D ultrasound (US) has been employed rapidly in medical intervention therapies, such as cardiac catheterization. To efficiently interpret 3D US images and localize the catheter during the surgery, an experienced sonographer is required. As a consequence, image-based catheter detection can be a benefit to sonographer to localize the instrument in the 3D US images timely. Conventionally, the 3D imaging methods are based on the Cartesian domain, which is limited by bandwidth and information lose when it is converted from the original acquisition space—Frustum domain. The exploration of catheter segmentation in Frustum space helps to reduce the computational cost and improve efficiency. In this paper, we present a catheter segmentation method in 3D Frustum image via a deep convolutional neural network (DCNN). To better describe 3D information and reduce the complexity of DCNN, cross-planes with spatial gaps are extracted for each voxel. Then, the cross-planes of the voxel are processed by the DCNN to distinguish it, whether it is a catheter voxel or not. To accelerate the prediction efficiency on whole US Frustum volume, a filter-based pre-selection is applied to reduce the computational cost of the DCNN. Based on experiments on the ex-vivo dataset, our proposed method can segment the catheter in Frustum images with 0.67 Dice score within 3 seconds, which indicates the possibility of real-time application.

Keywords: Catheter segmentation, 3D Frustum ultrasound, DCNN, *ex-vivo* dataset

1. INTRODUCTION

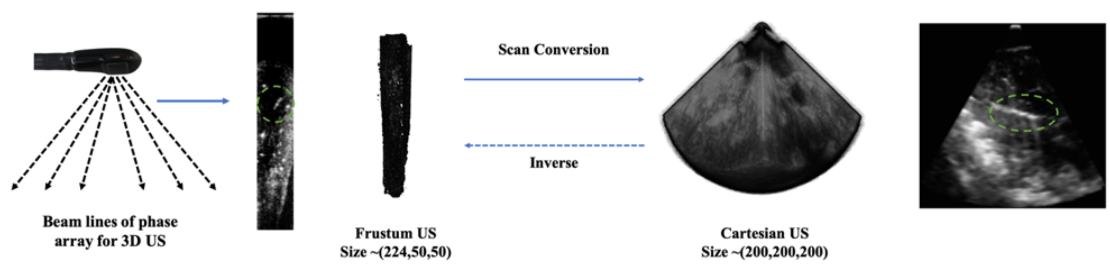

Figure 1: 3D Frustum US and its corresponding 3D Cartesian US. Catheters are indicated in the green circles, which shows the catheter in Frustum domain has deformation and irregular shape than standard Cartesian space.

In recent years, intervention therapy has been widely used because of its lower risk and shorter recovery time for patients, especially for cardiac catheterization for cardiovascular diseases.[1,2] During the intervention procedures, the surgeon or sonographer locates the position of catheter under the assistance of image visualization modalities, such as fluoroscope and ultrasound (US). However, fluoroscope, such as X-ray, has limitations like high radiation, lack of 3D spatial information and harmful contrast agent, which make researchers to find an

Further author information:
Hongxu Yang: E-mail: h.yang@tue.nl

alternative solution for intervention guidance, such as 3D US. More recently, 3D US stands out because it achieves real-time performance with richer spatial information, which makes the intervention therapy to be more efficient.

3D US image-based medical tool detection or segmentation has been studied over last years.[3–6] Traditionally, image processing algorithms were directly applied to the 3D Cartesian space, which is converted from the Frustum space, as shown in Figure 1. The Frustum space is an image coordinate space following a spherical coordinate and is the original 3D US image acquisition space. US images are captured by the phase array, which is formed into a cubic volumetric image, as shown in the left part of Figure 1. Traditionally, Frustum images are scan-converted into Cartesian coordinate, which projecting the real physical word and make it much easier to be interpreted. However, when compared to the Frustum space, catheter detection in the Cartesian space introduces more challenges in terms of application. First, during the image domain conversion procedure, the information is degraded due to interpolation, filtering and distortion. Second, the difficulties of converting Cartesian image back into Frustum image makes the system hard to project a mask in Cartesian domain back Frustum images, which introduces challenges of 3D US visualization rendering from Frustum domain. Third, the data communications between the different machines are rely on Frustum space, which is because of limited bandwidth and computation resources. As a result, exploiting the possibility of catheter segmentation is necessary in Frustum space. In this paper, we propose a catheter segmentation method in 3D Frustum images through the deep learning approach. The proposed method consists of two parts. First, the voxel-of-interests voxels are pre-selected by a Frangi vesselness filter, which can reduce the number of voxels to be processed by deep convolutional network (DCNN). Second, the remaining voxels are classified by our proposed multi-plane DCNN to distinguish the category of the voxels. The proposed pipeline provides a semi-real time segmentation performances in 3D Frustum image with a high Dice score of 0.67. The results indicate the feasibility of catheter segmentation in a Frustum image domain. To the best of our knowledge, this paper is the first paper for study the segmentation task in Frustum domain.

2. METHODS

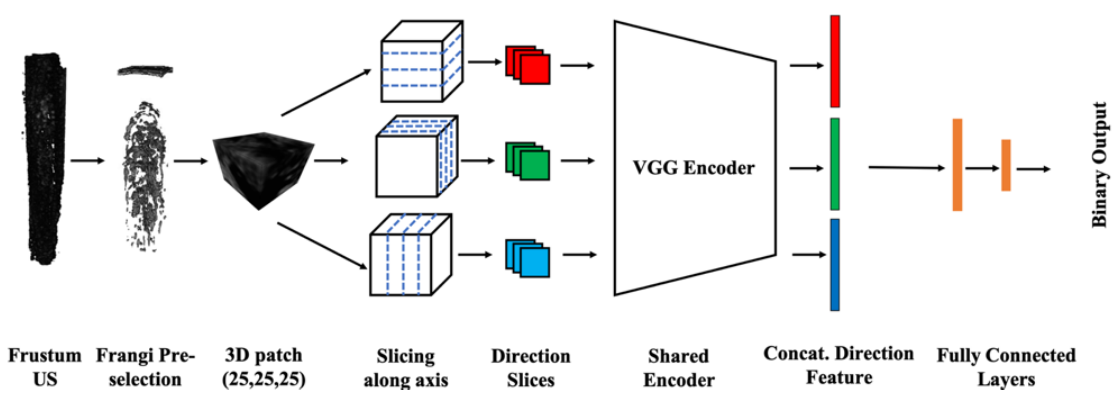

Figure 2: The proposed catheter voxel classification method in 3D Frustum US image. The input volume is first processed by a Frangi filter to select the most possible voxels belonging to the catheter. Then the voxel-based DCNN is applied to distinguish their categories.

The method of catheter segmentation in 3D Frustum US is shown in Figure 2, which consists of (1) Frangi-based Voxel-of-Interest pre-selection,[6] which selects the most possible voxels belonging to catheter, (2) local patch extraction for pre-selected voxels, (3) direction-slices extraction for local patch, and (4) binary classification by DCNN. The input Frustum volume is firstly processed by a Frangi vesselness filter, which is designed to extract the tubular shape local information. An adaptive intensity threshold is applied on the filtered image to select 25k possible voxels belonging to the catheter.[6] For each candidate voxel selected from filtered 3D Frustum image, its corresponding local patch is extracted with size (25,25,25) considering it as the center.[5,6] Then, three orthogonal slices are extracted through the center voxel. For each slice passing the center, its adjacent two slices with spatial gap d are extracted to form a three-channel image for a better spatial description,[7] which finally has an image size of (25,25,3). The three-channel images of three different axes are encoded by a VGG-based network[8] separately, which consists of five convolutional layers and two Maxpooling layers, to extract object-specific high-level feature vector. Three vectors are extracted by sharing the same VGG encoder to save the memory usage.[5,6] Then, these feature vectors are concatenated and processed by two fully connected layers with size 64 and 32. Finally, the binary prediction is generated by a softmax layer to distinguish the category of the voxel. With iterative prediction of all selected voxels in the 3D Frustum volume, the catheter is segmented from the image.

During the training, VGG-based encoder is initialized based on the pre-trained parameters from ImageNet challenge.[9] The remaining parameters of fully connected layers are initialized randomly. To generalized the network of a different direction and noise level, data augmentations are randomly formed on-the-fly, which includes randomly rotation, mirror, contrast adjust, etc. To deal with imbalanced catheter/non-catheter distribution in Frustum images, we perform a two-stage hard sample mining strategy,[5,6] which promises the information of the image can be used for DCNN training. More specifically, non-catheter samples are down-sampled to the same size of catheter samples, which is used to train the DCNN. Then, whole training images are validated by DCNN to extract misclassified voxels as the hard-classified samples. These samples are used to update the DCNN to learn the most meaningful information in the training dataset. The parameters are learned by minimizing weighted binary cross-entropy using Adam optimizer with learning rate 1e-5. Moreover, dropout layers are included between fully connected layers during the training with dropout rate as 0.5. The training procedure is terminated after converge with mini-batch size equals to 128.

3. RESULTS

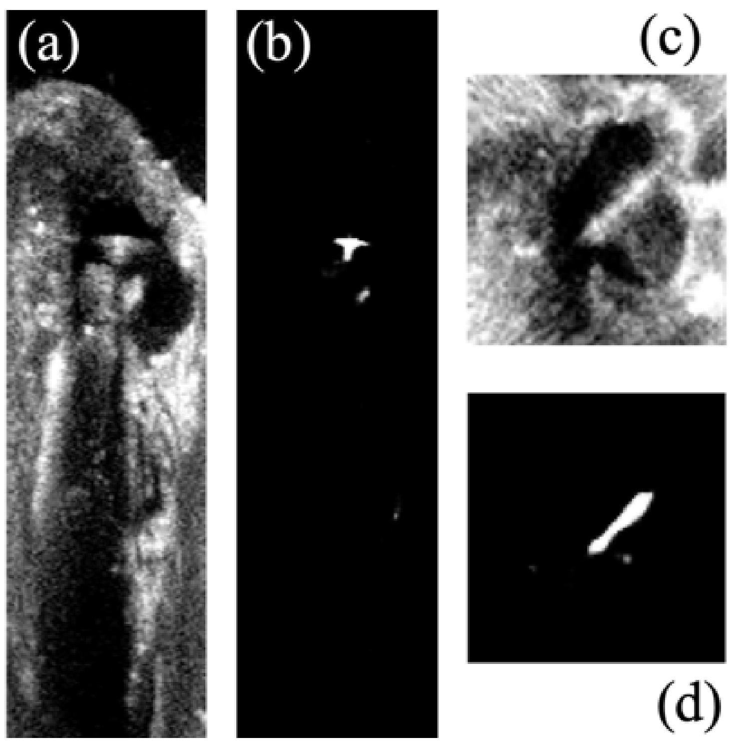

Figure 3: Example slices of catheter and its corresponding segmentation in Frustum image. (a)(c) Catheter in B-mode Frustum image along elevation and axial direction, respectively. (b)(d) Segmented catheter in Frustum image, corresponding to (a) and (c), respectively.

To validate our proposed method, we have collected eight 3D Frustum US images in an isolated porcine heart (ex-vivo dataset), which have an average volume size of (376,92,88) in Frustum space for axial, lateral and elevation direction. Their corresponding volume size in Cartesian space are around (208,336,320), which has more than seven times point size than the Frustum image. The catheters in Frustum images are manually annotated as the ground truth. From statically calculation in Frustum images, around 600 voxels are belonging to the catheter while the rest are belonging to empty space or tissue. As a consequence, Recall, Precision and Dice score are considered as the evaluation metrics to evaluate segmentation performance in the imbalanced classes distribution. Moreover, we also exploit the influence of spatial gap d as the ablation study. We consider our previously proposed handcrafted features with Random Forest classifier as the baseline.[4] A two-fold cross validation is applied on the dataset.

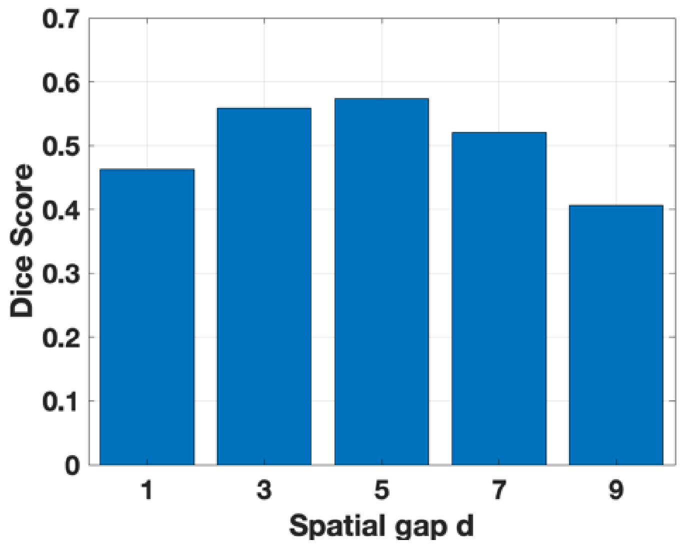

Figure 4: The influence of spatial gap d between adjacent slices, which is shown in Figure 2. The results are measure by Dice score, which indicates a proper spatial selection can improve the discriminating information extraction from VGG Encoder.

Table 1: Performance comparison between our proposed method and conventional handcrafted feature with supervised classifier. Values are representing mean (std.)

Method	Recall	Precision	Dice Score	Time (sec.)
MSMD Features by RF w/o Pre-selection	0.476 (0.328)	0.399 (0.253)	0.428 (0.282)	∼135
MSMD Features by RF w Pre-selection	0.471 (0.325)	0.501 (0.303)	0.477 (0.308)	∼1.4
Proposed DCNN w/o Pre-selection	0.795 (0.205)	0.512 (0.201)	0.574 (0.149)	∼226
Proposed DCNN w Pre-selection	0.785 (0.201)	0.655 (0.228)	0.673 (0.149)	∼2.2

Example images of the segmented catheter in Frustum image without pre-selection are shown in Figure 3. The catheter and its corresponding segmentation are shown in different view by 2D slicing. The influences of spatial gap d of three-channel images are shown in Figure 4, which indicates a proper spatial gap selection could improve the discriminating information extraction of local patch. As a consequence, $d = 5$ is considered as the best setting for the proposed DCNN method, as it extract the best spatial information description from the experimental results.

Based on the best performance of our proposed method under the gap setting $d = 5$, the Frangi-based pre-selection is applied prior to DCNN to further reduce the computation time in whole pipeline. The overall performance comparisons between Multi-scale and Multi-definition (MSMD) Features under Random Forest (RF) and our methods are shown in Table 1, and their corresponding examples are shown in Figure 5. In addition, the detailed comparison between our proposed DCNN and MSMD under pre-selection are demonstrated in Figure 6. From the table, our proposed DCNN achieves better performance than the conventional handcrafted feature. It has two reasons: (1) the previous designed feature is used for Cartesian domain that they might not extract sufficient discriminating information for catheter in Frustum domain with limited information capacity; (2) the three-channel & three-direction images can extract proper 3D spatial information and DCNN can to capture the most relevant feature to describe the catheter that it achieves better performance. More crucially, with pre-selection in Frustum image, the number of voxels to be classified by DCNN is drastically reduced from over 3 million to 25 k, which improve the efficiency around 100 times. Meanwhile, with efficiency improvement, the segmentation performance is also improved by pre-selection, which mainly improves the performance of Precision, i.e. fewer outliers. (Similarly, the MSMD method also has a performance improvement when considering the pre-selection procedure while it is still worse than DCNN.) As a consequence, our proposed catheter segmentation method in 3D Frustum images achieves an average Dice score of 0.673. Meanwhile, as shown in Figure 5, DCNN approach provides a more accurate edges than handcrafted approaches, which is because of a richer information description from DCNN. Moreover, the pre-selection can reduce the number of outliers, which is also indicated by the table. When compared to catheter segmentation in Cartesian images, which obtained similar Dice score[6] but cost more than 1400 seconds for full volume processing, catheter segmentation in Frustum images provides a higher efficient and accurate segmentation. As a result, it is much easier to be implemented into real-time applications.

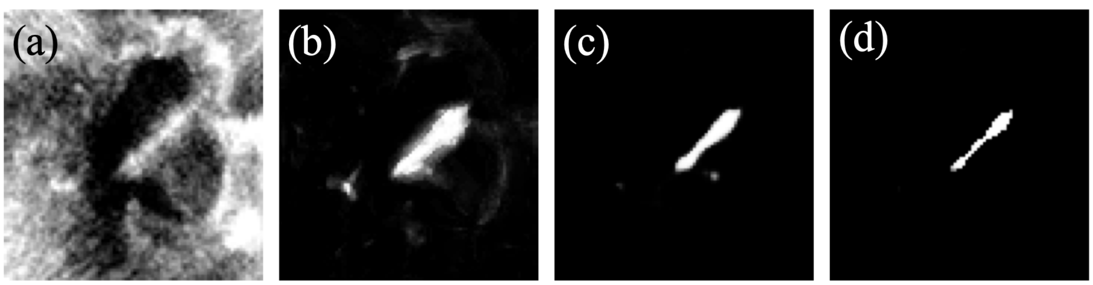

Figure 5: Example slices of the segmented catheter by different methods: (a) B-mode image, (b) Handcrafted feature by RF, (3) Proposed DCNN without pre-selection and (d) Proposed DCNN with pre-selection.

4. CONCLUSIONS

We have proposed an efficient and accurate catheter segmentation method in 3D Frustum US using DCNN approach, which is the first study to perform the segmentation task in Frustum domain. Based on proposed architecture, the catheter in ex-vivo dataset can be segmented with a Dice score of 0.67 within 3 seconds (includes pre-selection and DCNN processing). The results show that our proposed method is much more efficient and can keep a same segmentation accuracy when compared to the studies in Cartesian space. From the experiments, our DCNN shows a better performance than conventional handcrafted-feature-based method.

Image-based catheter segmentation can facilitate the cardiac intervention and therefore improve the operation efficiency and outcome. To improve the catheter segmentation efficiency and keep the accuracy, we propose a catheter segmentation method in 3D Frustum US images, which is the first study to exploit the feasibility of instrument segmentation in 3D Frustum US modality. The experiment results show our proposed method achieves 0.67 Dice score in challenging ex-vivo Frustum images. Based on the result, we will investigate the possibility of Volume-of-Interest-based approach to further reduce the calculation time, which pave the way to the real-time application. Moreover, based on the results from DCNN, proper post-processing is necessary to omit outliers and further improve the performance.

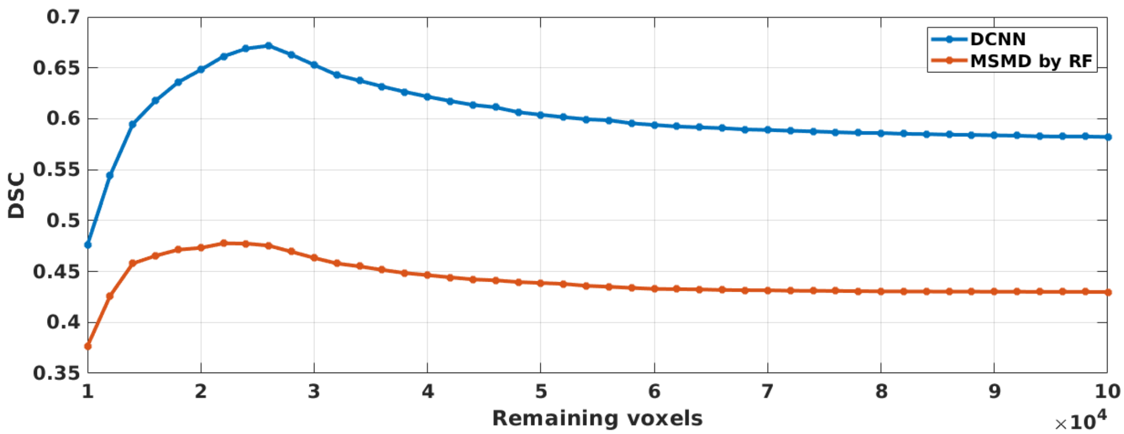

Figure 6: DSC performance curves of two different methods with respect to remaining voxels after the pre-selection. Blue curve represents the DCNN while red curve represents MSMD by RF.

5. STATEMENT

This work is original and in its present form has not been submitted elsewhere in any form. This research was conducted in the framework of "Impulse-2 project for the healthcare flagship–topic ultrasound" at Eindhoven University of Technology in collaboration with Catharina Hospital Eindhoven and Royal Philips.

REFERENCES

[1] Graham, M. M., Knudtson, M. L., O'Neill, B. J., Ross, D. B., to Care Working Group, C. C. S. A., et al., "Treating the right patient at the right time: Access to cardiac catheterization, percutaneous coronary intervention and cardiac surgery," *Canadian Journal of Cardiology* **22**(8), 679–683 (2006).

[2] Camuglia, A. C., Randhawa, V. K., Lavi, S., and Walters, D. L., "Cardiac catheterization is associated with superior outcomes for survivors of out of hospital cardiac arrest: review and meta-analysis," *Resuscitation* **85**(11), 1533–1540 (2014).

[3] Pourtaherian, A., Scholten, H. J., Kusters, L., Zinger, S., Mihajlovic, N., Kolen, A. F., Zuo, F., Ng, G. C., Korsten, H. H., and de With, P. H., "Medical instrument detection in 3-dimensional ultrasound data volumes," *IEEE transactions on medical imaging* **36**(8), 1664–1675 (2017).

[4] Yang, H., Shan, C., Pourtaherian, A., Kolen, A. F., et al., "Catheter segmentation in three-dimensional ultrasound images by feature fusion and model fitting," *Journal of Medical Imaging* **6**(1), 015001 (2019).

[5] Yang, H., Shan, C., Kolen, A. F., and de With, P. H., "Catheter detection in 3d ultrasound using triplanar-based convolutional neural networks," in [*2018 25th IEEE International Conference on Image Processing (ICIP)*], 371–375, IEEE (2018).

[6] Yang, H., Shan, C., Kolen, A. F., et al., "Catheter localization in 3d ultrasound using voxel-of-interest-based convnets for cardiac intervention," *International journal of computer assisted radiology and surgery* **14**(6), 1069–1077 (2019).

[7] Pourtaherian, A., Zanjani, F. G., Zinger, S., Mihajlovic, N., Ng, G. C., Korsten, H. H., et al., "Robust and semantic needle detection in 3d ultrasound using orthogonal-plane convolutional neural networks," *International journal of computer assisted radiology and surgery* **13**(9), 1321–1333 (2018).

[8] Simonyan, K. and Zisserman, A., "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556* (2014).

[9] Yang, H., Shan, C., Kolen, A. F., et al., "Improving catheter segmentation & location in 3d cardiac ultrasound using direction-fused fcn," in [*International Symposium on Biomedical Imaging 2019*], (2019).