

DEEP LEARNING WITH ORTHOGONAL VOLUMETRIC HED SEGMENTATION AND 3D SURFACE RECONSTRUCTION MODEL OF PROSTATE MRI

Ruida Cheng^a, Nathan Lay^b, Francesca Mertan^c, Baris Turkbey^c, Holger R. Roth^b, Le Lu^b, William Gandler^a, Evan S. McCready^a, Thomas Pohida^d, Peter Choyke^c, Matthew J. McAuliffe^a, Ronald M. Summers^b

^a Imaging Sciences Laboratory, ^d Computational Bioscience and Engineering Laboratory, Center for Information Technology, NIH, Bethesda, MD

^b Imaging Biomarkers and CAD Laboratory, Clinical Center, NIH, Bethesda, MD

^c Molecular Imaging Program, NCI, Bethesda, MD

ABSTRACT

Automatic MR whole prostate segmentation is a challenging task. Recent approaches have attempted to harness the capabilities of deep learning for MR prostate segmentation to tackle pixel-level labeling tasks. Patch-based and hierarchical features-based deep CNN models were used to delineate the prostate boundary. To further investigate this problem, we introduce a Holistically-Nested Edge Detector (HED) MRI prostate deep learning segmentation and 3D surface reconstruction model that facilitate the registration of multi-parametric MRI with histopathology slides from radical prostatectomy specimens and targeted biopsy specimens. Application of this technique combines deep learning and computer aided design to provide a generalized solution to construct a high-resolution 3D prostate surface from MRI images in three orthogonal views. The performance of the segmentation is evaluated with MRI scans of 100 patients in 4-fold cross-validation. We achieve a mean Dice Similarity of 88.6%.

Keywords — Prostate segmentation, HED, Deep learning

1. INTRODUCTION

Multi-parametric magnetic resonance imaging (MRI) improves the clinical diagnosis and staging of prostate cancer. The prostate gland boundary from three-dimensional MRI can be segmented to generate a surface model which informs the clinician of the whole prostate volume. In addition to whole prostate volume, the surface model can be used to correlate prostate MRI to histology specimens. Shah *et al.* used patient-specific molds (PSM) created from prostate surface models to correlate MRI to histopathology and developed a decision support system for generating cancer probability maps from multi-parametric MRI [1]. The PSM is created using prostate surface models from MR images of each patient obtained prior to surgery.

The prostate margin segmentation was performed manually. Since the multi-slice MRI scans were obtained with thick slices relative to the in-plane resolution (due to scan time constraints), a surface model generated from only one view, such as the axial view, would give poor resolution along the slice direction. For the PSM, this limitation was overcome by segmenting the prostate in three orthogonal MRI scans and combining them either as binary masks or as clouds of points before reconstructing the surface. The same prostate surface model could also be used in another method to register MRI with real time trans-rectal ultrasound (TRUS) in order to target lesions during biopsy [2]. An automated segmentation and surface reconstruction approach would reduce the amount of tedious manual labor needed to segment multiple MR images.

In previous works, only a few methods were investigated for MRI prostate segmentation in the deep learning domain. Guo *et al.* [3] proposed a deformable segmentation method to learn latent features by a stacked sparse auto-encoder and then used a sparse patch matching to infer the prostate boundary. Fausto *et al.* [4] presented a 3D volumetric deep CNN to model an axial image and optimize the training with a Dice coefficient objective function. It utilized a V-Net, a 3D based convolutional and de-convolutional architecture to learn the 3D hierarchical features. Liao *et al.* [5] proposed representation learning, a stacked independent subspace analysis network that is adopted to learn the most effective features in a hierarchical and unsupervised manner. All the previously proposed methods focused on the MR axial prostate images alone. We investigate the capability to apply the HED deep learning architecture to low resolution orthogonal volumetric MR prostate images, which leverages the 2D HED deep learning model to all the axial, sagittal, and coronal images, and then generates the final higher resolution 3D surface from the segmented contours.

As compared to works described above, our approach shows that it is applicable to apply the HED pixel level labeling model to axial, sagittal and coronal images so that one can form an end-to-end trainable system for MR

prostate image segmentation which combines the strengths of deep learning and 3D surface reconstruction to extract a high resolution surface from thick-sliced MRI. 3D deep CNNs model [4] requires substantially large amount of memory for 3D activation maps and computation during training and testing phases, which are prohibitively expensive. The orthogonal volumetric HED 2D deep CNNs model only deals with 2D hierarchical features and 2D activation maps. The processing time is much faster than the 3D volume based deep CNNs model.

2. METHODS

The proposed model is composed of three major building blocks: 1) fully automatic Holistically-Nested Networks volumetric segmentation model using orthogonal MRI prostate images; 2) Ball-pivoting and Poisson based 3D surface reconstruction model to create a high resolution 3D surface from low MRI spatial resolution along the z-axis; 3) 3D prostate mold creation and 3D printing for pathology biopsy analysis. The schema of the proposed method is shown in Figure 1.

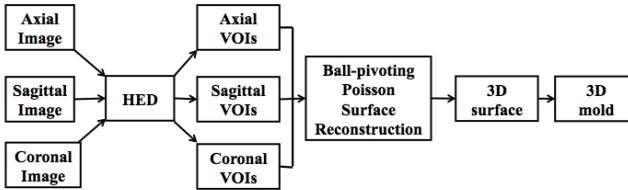


Figure 1. Schema of the proposed pipeline

The input data are MR images. The T₂-weighted MR images of the entire prostate were obtained in three orthogonal planes (sagittal, axial and coronal) each at the scan resolution of 0.2734x0.2734x3.0 mm³; field of view 140 mm; image slice dimension 512x512. The center of the prostate is the focal point for the MRI scan.

2.1. HED volumetric segmentation on orthogonal images

During the training phase, we extracted 2D slices from 3D MR orthogonal images as well as corresponding binary mask slices from the annotations for the training data. In this work, we explicitly learn the prostate interior binary mask labeling models via HED. Each orthogonal image (axial, sagittal, coronal) is trained with its own HED model. The three orthogonal models combined together constitute the volumetric segmentation model of prostate MRI.

The HED was first proposed by [8] as a deep learning method for detecting edge and object boundaries in natural images. It emphasizes an end-to-end edge detection system, a system inspired by fully convolutional neural networks with additional deep supervision on top of VGGNet [9]. Thus, images of different size can be handled by HED via training and detection phases. The HED networks comprise a single stream deep network with multiple side-outputs

(Figure 2), and the side-outputs are inserted after each convolutional layer. The outputs of HED are multi-scale and multi-level with the side-outputs plane size becoming smaller and receptive field size becoming larger. Each side-output produces a corresponding edge map at a different scale level, and one weighted-fusion layer is added to automatically learn how to combine outputs from multi-scale, as shown in Figure 2. The entire network is trained with multiple error propagation paths (dashed lines). Ground truths are used at each of the side-output layers to compensate the weak edges lost and play an important role in supervised learning. The HED architecture is a simple feed forward neural network that produces multi-scale outputs in a single path. With the per-pixel labeling cost function [8], the HED can be effectively trained using only several thousand annotated image pairs. This enables the automatic learning of rich hierarchical feature representations (contexts) that are critical in resolving spatial ambiguity in the prostate segmentation task.

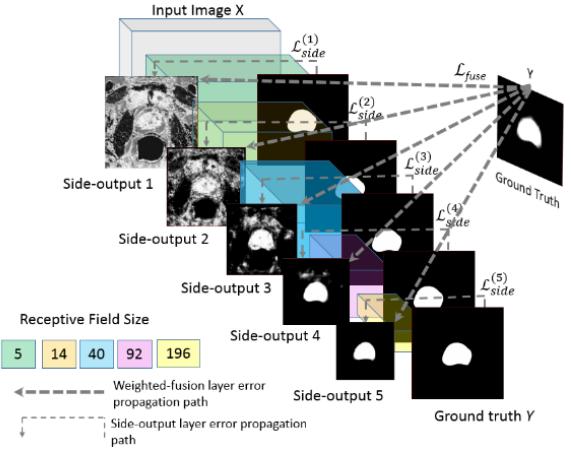


Figure 2. HED architecture for axial images training (adapted from [8] with permission)

A HED network has M side-output layers, where each side-output layer functions as a classifier with corresponding weights $w = (w^1, \dots, w^M)$. All standard network layers parameters are denoted as W. The object function of the M side-output layers is,

$$\mathcal{L}_{side}(W, w) = \sum_{m=1}^M \alpha_m l_{side}^{(m)}(W, w^{(M)}) \quad (1)$$

\mathcal{L}_{side} denotes an image-level loss function for side-outputs, computed over all pixels in a training image pair X and Y (MR image and binary mask). The prostate interior map prediction $\hat{Y}_{side} = \sigma(\hat{A}(m)_{side})$ can be obtained at each side-output layer, where $\hat{A}(m)_{side}$ are the activation of the side-output layer m. The loss function at the fusion layer \mathcal{L}_{fuse} is defined as,

$$\mathcal{L}_{fuse}(W, w, h) = Dist(Y, \hat{Y}_{fuse}) \quad (2)$$

where $\hat{Y}_{fuse} \equiv \sigma(\sum_{m=1}^M h_m^{A_{side}})$ with $h = (h_1, \dots, h_M)$ being the fusion weight. $Dist(\cdot, \cdot)$ is the distance between the fused predictions and the ground truth label map, which is set to be cross-entropy loss. The overall objective loss function of HED is minimized via back propagation with stochastic gradient descent.

$$\mathcal{L}_{HED}(W, w, h) = argmin(\mathcal{L}_{side}(W, w) + \mathcal{L}_{fuse}(W, w, h)) \quad (3)$$

During testing phase, with new image X , the prostate binary mask prediction maps \hat{Y}_{fuse} and \hat{Y}_{side} are generated from both side-output layers and the weighted fusion layer. HNN denotes the edge maps produced by HED networks.

$$(\hat{Y}_{fuse}, \hat{Y}_{side}^{(1)} \dots \hat{Y}_{side}^{(5)}) = HNN(X, (W, w, h)). \quad (4)$$

After HED prediction, the \hat{Y}_{fuse} layer generated probability map is used to search the final contour. The HED-generated fusion layer probability map sometimes contains isolated components that present a coarse level prostate prediction. A morphology filter runs against the probability map to remove noise and searches for the largest region to represent the prostate. Then, it generates the prostate shape binary mask from the probability map and converts the mask to the final VOI contour. The tested image fused with probability map is shown in Figure 3.

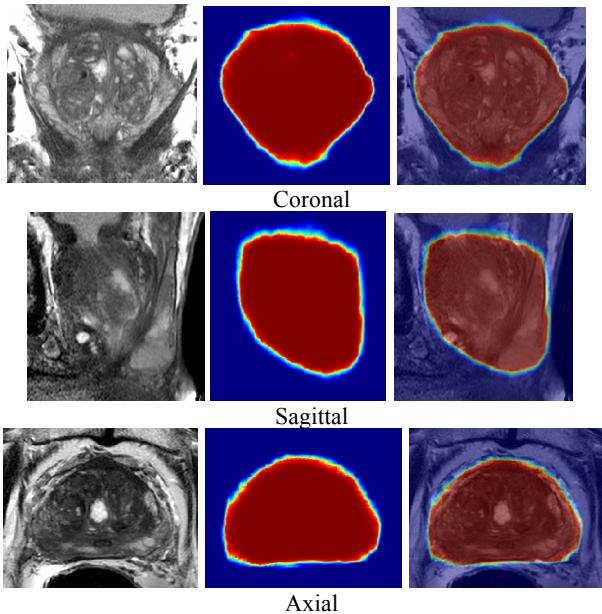


Figure 3. Probability map. Column 1: Original MR images. Column 2: HED predicted probability map. The dark red region represents the higher likelihood of the segmented prostate shape, which is used to delineate the final prostate contour. Column 3:

Fused image from original MR image and predicted probability map.

2.2. Ball-pivoting and Poisson 3D surface reconstruction

The HED automatic prostate segmentation generates VOIs for each axial, sagittal and coronal image. Each VOI contour is generated from the high resolution xy plane in each view. The distance between VOI contours is large due to the low resolution along the z axis on each orthogonal image. We select 100 points to represent each VOI contour and merge the resulting three VOIs in the DICOM space, forming a rough point cloud. The point cloud reflects the high resolution density from each image, and compensates for the low z-axis resolution. The Ball-Pivoting and Poisson surface reconstruction algorithms are run against the point cloud to construct the high resolution prostate 3D surface.

The Ball-Pivoting Algorithm (BPA) [10] takes the point cloud as input without normal information to build a rough surface. The point cloud merged from the three VOIs has low density sample points due to the low z-axis resolution. As a result, the BPA generated mesh might leave holes or non-smooth patches on the surface. Thus, the Poisson surface reconstruction algorithm is used to remove bumpy patches and to create the final smooth surface. The Poisson algorithm (PA) utilizes an implicit function to approximate the surface as a solution to a Poisson equation [11]. Figure 4 demonstrates the Ball-Pivoting and Poisson surface reconstruction results respectively.

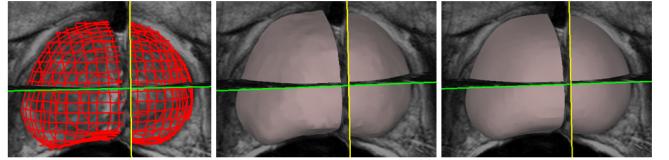


Figure 4. Ball-Pivoting and Poisson surface reconstruction. (1) After HED segmentation, merging the VOIs contours from axial, sagittal, and coronal images to constitute a point cloud. (2) Ball pivoting surface reconstruction generated bumpy surface. (3) Poisson surface reconstruction generated final smooth surface.

2.3. 3D prostate mold creation and 3D printing

The Java-based open source MIPAV [7] visualization software uploads the final prostate surface reconstruction. This surface is decimated and converted to a stereolithography (STL) surface file format for subsequent semi-automated design and fabrication of the physical 3D prostate mold. The prostate mold design process utilizes the STL file and other MR imaging setup parameters to generate a surrounding box with both the prostate void and integrated knife slots spaced 6mm apart. The position and orientation of the prostate void relative to the knife slots establishes spatial congruence between the resulting tissue blocks and MRI slices. The prostate mold (Figure 5) is fabricated via 3D printing. As a result of sectioning the

prostate in the custom mold, the tissue blocks (and subsequent histological sections) are more directly applicable in the validation of mpMRI as a means of prostate tumor localization for clinical diagnosis and treatment planning.

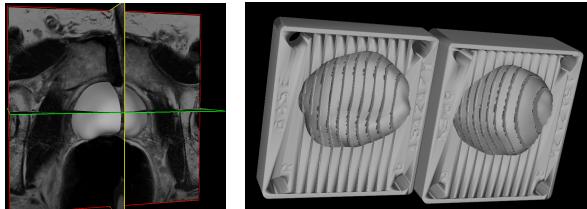


Figure 5. 3D mold creation and 3D printing

3. RESULTS

The HED segmentation method performance was evaluated with 100 prostate MRI scans per view. Four-fold cross validation comparing segmentation performance against the ground truth binary masks was conducted. The segmentation performance (Table 1) was evaluated with 1) Dice similarity coefficient (DSC), 2) Hausdorff distance (mm), and 3) Average Symmetric Surface Distance (SSD, mm). The HED deep learning segmentation model elevates the mean DSC to 88.6%, and mean Hausdorff distance to 17.43 mm. All the metrics are calculated without trimming any ending contours or cropping data to the [5%, 95%] probability interval to avoid outliers for distance based measure. The mask based performance measure uses the EvaluateSegmentation tool [13] to compare the ground truth and segmented masks. Manual reference segmentations drawn by the experts approximate the ground truth. For volumetric measure, the 3D surface volume is calculated from the binary surface volumetric mask by multiplying the total surface volumetric voxels with a single voxel volume. MIPAV [7] 3D visualization and surface volume tool is used to conduct the volumetric measure. The absolute 3D surface volume difference is computed by $|Vol_{seg} - Vol_{gt}| / Vol_{gt}$. The resulting mean 3D surface volume difference is 11.4%. Figure 6 illustrates the 3D surfaces overlapping between proposed method generated surface and the ground truth surface of one patient. Noticeable volume differences are shown on top and bottom of the coronal and sagittal views. The HED orthogonal volumetric segmentation model generates a comparable result to other deep learning based segmentation methods [3, 4, 5] in the literature. The higher Hausdorff distance primarily results from not trimming ending contours. The proposed method facilitates the HED deep learning segmentation model to axial, coronal and sagittal images of MRI prostate. Even in case that HED generates erroneous points on the 2D contour, the 3D BPA ball rolling mechanism can effectively correct the errors in 2D segmentation, and ensure approximate 3D surface creation. By merging the axial, sagittal and coronal VOIs contours into a point cloud in DICOM space, the low

resolution issue from each orthogonal view's slice direction can be essentially eliminated. The Ball-pivoting and Poisson algorithms finally build a smoothed high resolution 3D prostate surface for 3D mold printing. The processing time of one test image including HED orthogonal segmentation and 3D surface reconstruction is under one minute, which is significantly better than other literature approaches.

	DSC	Hausdorff[mm]	SSD[mm]	3D Vol diff
axial	88.98%	14.53	0.187	
sagittal	88.57%	20.24	0.278	
coronal	88.26%	17.54	0.254	
mean	88.6%	17.43	0.24	11.4%

Table 1. HED segmentation results

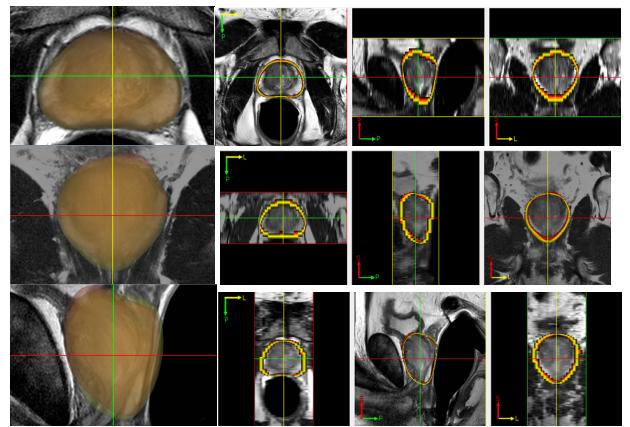


Figure 6. 3D surface comparison (red: ground truth, yellow: HED segmentation and surface reconstruction)

4. CONCLUSION

A novel method for prostate MRI segmentation and surface reconstruction was presented in this paper. Unlike the previous deep learning based methods, the proposed model takes advantage of HED deep learning to provide a robust segmentation mechanism, which is simple, efficient and fast. The BPA and PA surface reconstruction algorithms ensure the rapid prototyping of high resolution 3D prostate surface from thick-slice orthogonal MRI. The primary contributions of the proposed method are the automatic segmentation pipeline, enabling shorter analysis times of axial, coronal, and sagittal patient images, and more importantly, increased stability from the HED deep learning hierarchical feature based segmentation. The experimental results demonstrate the promising of applying HED for 3D orthogonal volumetric segmentation of prostate MRI. The distributed segmentation and reconstruction pipeline provides a unified tool under MIPAV for generating prostate surface models that can be used to improve the correlation of prostate MRI to histopathology [2] and fusion of MRI with TRUS for targeted biopsy [12].

5. ACKNOWLEDGMENT

We would like to thank Nvidia for the K40, Titan X GPU donations to our groups.

6. REFERENCES

- [1] S. Borofsky, S. Haji-Momenian, S. Shah, M. Taffel, "Multi-parametric MRI of the prostate gland: technical aspects," *Future oncology*, 12(21):2445-62, 2016.
- [2] V. Shah, T. Pohida, B. Turkbey, H. Mani, M. Merino, P. Pinto, P. Choyke, M. Bernardo, "A method for correlating in vivo prostate magnetic resonance imaging and histopathology using individualized magnetic resonance-based molds," *Rev Sci Instrument*, vol. 80, pp. 104301, 2009.
- [3] Y. Guo, Y. Gao, D. Shen, "Deformable MR Prostate Segmentation via Deep Feature Learning and Sparse Patch Matching", *IEEE Transactions on Medical Imaging*, 2015.
- [4] F. Milletari, N. Navab, S. Ahmadi, "V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation", <https://arxiv.org/abs/1606.04797>.
- [5] S. Liao, Y. Gao, A. Oto, D. Shen, "Representation Learning: A Unified Deep Learning Framework for Automatic Prostate MR Segmentation", *Med Images Comput Assist Interv*, 2013.
- [7] MIPAV, [Medical Image Processing, Analysis, and Visualization (MIPAV)], <http://mipav.cit.nih.gov>
- [8] S. Xie, Z. Tu, "Holistically-Nested Edge Detection," *IEEE ICCV*, 1395-1403, 2015.
- [9] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *ICLR*, 2015.
- [10] F. Bernardini, J. Mittleman, H. Rushmeier, C. Silva, G. Taubin, "The Ball-Pivoting Algorithm for Surface Reconstruction," *IEEE Transactions on Visualization and Computer Graphics*, vol. 5, no. 4, pp. 349-359, Oct.-Dec. 1999.
- [11] M. Kazhdan, M. Bolitho, and H. Hoppe, "Poisson surface reconstruction," In *Eurographics Symposium on Geometry Processing*, pp. 61–70, 2006.
- [12] S. Xu, J. Kruecker, B. Turkbey, N. Glossop, A. Singh, P. Choyke, P. Pinto, B. Wood, "Real-time MRI-TRUS fusion for guidance of targeted prostate biopsies," *Computer Aided Surg*, vol. 13, no. 5, pp. 255-64, September 2008.
- [13] <http://github.com/Visceral-Project/EvaluateSegmentation>