Article Type: Research Article

Deep-Learning Convolutional Neural Network: Inner and Outer Bladder Wall Segmentation in CT Urography

Marshall N. Gordon
Lubomir M. Hadjiiski
Kenny H. Cha
Ravi K. Samala
Heang-Ping Chan
Richard H. Cohan
Elaine M. Caoili

Department of Radiology, The University of Michigan, Ann Arbor, Michigan 48109-0904

Running Title: Bladder Wall Segmentation using DL-CNN and Level Sets

Correspondence:

Marshall N. Gordon

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/mp.13326

Department of Radiology
University of Michigan
1500 E. Medical Center Drive
MIB C474
Ann Arbor, MI 48109-5842
Telephone: (734) 936-8282
Fax: (734) 615-5513
E-mail:mngordon@umich.edu

ABSTRACT

Purpose: We are developing a computerized segmentation tool for the inner and outer bladder wall as a part of an image analysis pipeline for CT urography (CTU).

Materials and Methods: A data set of 172 CTU cases was collected retrospectively with IRB approval. The data set was randomly split into two independent sets of training (81 cases) and testing (92 cases) which were manually outlined for both the inner and outer wall. We trained a deep-learning convolutional neural network (DL-CNN) to distinguish the bladder wall from the inside and outside of the bladder using neighborhood information. Approximately 240,000 regions of interest (ROIs) of 16 x 16 pixels in size were extracted from regions in the training cases identified by the manually outlined inner and outer bladder walls to form a training set for the DL-CNN; half of the ROIs were selected to include the bladder wall and the other half were selected to exclude the bladder wall with some of these ROIs being inside the bladder and the rest outside the bladder entirely. The DL-CNN trained on these ROIs was applied to the cases in the test set slice by slice to generate a bladder wall likelihood map where the gray level of a given pixel represents the likelihood that a given pixel would belong to the bladder wall. We then used the DL-CNN likelihood map as an energy term in the energy equation of a cascaded level sets method to segment the inner and outer bladder wall. The DL-CNN segmentation with level sets was compared to the 3D hand-segmented contours as a reference standard.

Results: For the inner wall contour, the training set achieved the average volume intersection, average volume error, average absolute volume error, and average distance of 90.0±8.7%, -4.2± 18.4%, 12.9±13.9%, and 3.0±1.6mm, respectively. The corresponding values for the test set were 86.9±9.6%, -8.3±37.7%, 18.4±33.8%, and 3.4±1.8mm, respectively. For the outer wall contour, the training set achieved the values of 93.7±3.9%, -7.8±11.4%, 10.3±9.3%, and 3.0±1.2mm,

respectively. The corresponding values for the test set were 87.5±9.9%, -1.2±20.8%, 11.9±17.0%, and 3.5±2.3mm, respectively.

Conclusions: Our study demonstrates that DL-CNN assisted level sets can effectively segment bladder walls from the inner bladder and outer structures despite a lack of consistent distinctions along the inner wall. However, even with the addition of level sets, the inner and outer walls may still be over segmented and the DL-CNN assisted level sets may incorrectly segment parts of the prostate that overlap with the outer bladder wall. The outer wall segmentation was improved compared to our previous method and the DL-CNN assisted level sets was also able to segment the inner bladder wall with similar performance. This study shows the DL-CNN assisted level set segmentation tool can effectively segment the inner and outer wall of the bladder.

Key Words: Computer-Aided Diagnosis, Deep-Learning, Segmentation, CT Urography, Bladder, Bladder Wall

1. INTRODUCTION

The American Cancer Society estimates that in 2017, 76,900 people will be diagnosed with bladder cancer (58,950 in men and 18,010 in women) and that bladder cancer will be responsible for 16,390 deaths (11,820 in men and 4,570 in women)[1].

Early detection of bladder cancer is possible using a multidetector row CT (MDCT) urography exam, which can image the bladder, kidneys, and ureters in a single scan[2-6]. Unfortunately, the interpretation of a CT urography (CTU) exam is a time-intensive process. A CTU scan has on average 300 slices in each exam with a range of 200 to 600 slices. Since multiple lesions are possible in different areas of the scan, the entire set of slices must be carefully examined by the radiologist to determine if a lesion is present, which involves frequent adjustment of images in order to best view each slice. Additionally, CTU scans often have many urinary anomalies and the radiologist must determine the likelihood that each anomaly is benign or cancerous. These challenges in CTU interpretation lead to large variance among radiologists in detecting bladder cancer. Studies showed that the sensitivity ranges from 64% to 97% among radiologists[7,8].

Due to the large inter-radiologist variability and the substantial chance that a radiologist may miss a potentially cancerous lesion, it is imperative to develop techniques that make diagnosing bladder cancer easier and more accurate. To accomplish this goal, we are developing a computer-aided detection (CAD) system to assist with diagnosis of urothelial neoplasms including asymmetrical bladder wall thickenings. A critical step in developing a CAD system is the segmentation of the bladder and bladder wall as it determines the search region for the subsequent steps. If the bladder is under-segmented such that parts of the bladder are not included within the segmented region, the search may miss the lesions that appear outside the segmented region. Without segmentation or with an over-segmented bladder, non-bladder structures may be identified as lesions and also distract the detection of true lesions, thereby increasing the chance of both false positives and false negatives. Additionally, lesions typically manifest in the wall of the bladder. For these reasons, accurate segmentation of the bladder wall from the interior of the bladder and the outer structures is a crucial step in the development of a CAD system for urinary cancer.

Segmentation of the bladder wall in CTU is a challenging task due to a number of confounding factors. Bladders may be partially or completely filled with intravenous contrast material or have no contrast material at all. The different conditions create an inconsistent boundary between the wall of the bladder, the interior, and the outside structures. The boundary between the contrast material inside the bladder, the bladder wall, and the surrounding structures is a weak boundary and the CTU images are often very noisy. A common difficulty in segmenting the bladder wall is the prostate in male patients. For some cases, the prostate protruding into the bladder has a similar appearance as the bladder wall. Often the interface between the bladder wall and the prostate is difficult to distinguish. A bladder wall segmentation method based on simple thresholding tends to leak into the portions of the prostate and the surrounding structures. The boundary between the non-contrast and contrast regions of the bladder interior strongly resembles a wall, further confounding the segmentation process. Adding to these problems is the irregular sizes and shapes of the bladder. Even healthy bladders can be shaped oddly or appear to have anomalies, making it extremely difficult to accurately segment the bladders with any conventional methods.

Previous researchers have attempted to segment the bladder. Li et al.[9] and Duan et al.[10] used magnetic resonance (MR) images to segment the bladder wall. Duan et al.[11] used a window setting scheme to segment and detect tumors in MR images. Han et al.[12] used a Markov random field model

and level sets to segment MR images. Chai et al[13] developed a segmentation system for cone beam CT, but relied on population data as prior knowledge which can cause errors in cases that deviate from the training set. These methods have the advantage of being relatively fast on modern hardware, but these studies all employed small data sets of no more than 22 patients. Other researchers have attempted to use level sets to segment the bladder wall. Ma et al.[14] used level sets to segment the inner wall in MR images, but relied on a separate Chan-Vese model to segment the outer wall. Chi et al. used coupled level sets to segment the bladder wall in MR images, but only on a small data set of 11 patients[15]. Hadjiiski et al.[16,17] initially proposed a level set segmentation method for bladders in CTU. They[18] subsequently developed a new system based on the characteristics of bladder CTU images called conjoint level set analysis and segmentation system (CLASS). Cha et al.[19] further improved the segmentation accuracy of the CLASS method and expanded the data set to the one used in the current study. This method showed improved results over previous attempts to segment the bladder, but relied on two manually marked bounding boxes for the contrast-filled and non-contrast filled regions, respectively human inputs as a starting point.

In this study, we explored the use of a Deep Learning Convolutional Neural Network (DL-CNN) to segment the bladder wall. Convolutional neural networks (CNN)[20] was first applied to medical image pattern recognition by Lo et al.[21] for lung nodule detection and has been used for the detection of breast lesions[22-30] since the early 1990's. These early applications used relatively shallow CNN and typically trained with smaller data sets due to the extensive training time and limited availability of medical images. As graphical processing units (GPU) become more commonly available and powerful, in combination with the design of better convergence and regularization techniques, the application of CNN with large number of layers and trained with larger data sets to complex pattern recognition tasks becomes more practical. Krizhevsky et al[31,32] demonstrated the effectiveness of a DL-CNN in image classification on the ImageNet ILSVRC-2010 and ILSVRC-2012 data sets[33], and the CIFAR-10 data set[34]. Since then there have been strong interests in applying DL-CNN to all kinds of medical imaging problems[35]. We explored the use of a DL-CNN in the segmentation of the entire bladder (Cha et al.[36]) and the bladder wall in previous pilot studies (Gordon et al.[37]). In Cha et al., a DL-CNN with level sets was shown to be more effective in segmenting the bladder than previous level set systems or other standard image segmentation techniques. In the current study, we further investigated the segmentation of the bladder wall as a

step towards bladder lesion detection. The DL-CNN in this study was trained to distinguish the bladder wall and the output of the DL-CNN was used to guide level set segmentation of the inner and outer bladder wall.

This paper is organized as follows. First, the data set for this study is described. Second, the details and use of the DL-CNN is described. Third, the level set segmentation method is presented and lastly, the results are reported and discussed.

2. MATERIALS AND METHODS

A DL-CNN was trained to distinguish between ROIs inside the bladder wall and ROIs not within the bladder wall. The DL-CNN outputs a likelihood map where the brightness of a given pixel represents the likelihood that the pixel falls within the bladder wall. The likelihood map is then used as a gradient image in cascaded level sets to segment the inner and outer bladder wall. The process is shown in the flowchart in Figure 1.

2.1 Data set

With IRB approval, a data set of 172 patients who had undergone CTU was collected for this study. The cases were collected retrospectively from the Abdominal Imaging Division of the Department of Radiology at Michigan Medicine. All cases were acquired with established CTU protocol during the patients' clinical care. Within this set of 172 cases, we designated 81 as part of the training set and the remaining 91 as part of the test set. The cases were balanced in difficulty by appearance and shape of the bladder between the training and the test set.

In the training set, 42 bladders contained focal mass-like lesion, 21 had wall thickenings and 18 were normal. 40 of the bladders that contained focal mass-like lesion were malignant. Sixteen of the wall thickened bladders were also malignant. Sixty-one bladders in the training set were partially filled with IV contrast material, 8 were fully filled, and 12 were not filled. In the test set, 42 bladders contained focal mass-like lesion, 36 had wall thickenings and 13 were normal. Forty-two of the bladders that contained focal mass-like lesion and 23 of the bladders with wall thickenings were

malignant. Eighty-four of the test set bladders were partially filled with contrast material, 4 were fully filled, and 3 were not filled.

The inner and outer bladder walls were hand outlined by an abdominal radiologist with more than 20 years of experience for both the training and the test sets using an in-house developed graphical user interface that we named MiViewer. The hand outlines of the training set were used to classify the regions of interest for the DL-CNN training. The outlines were also used as the reference standard for the evaluation of the segmentation performance (see Section 2.4). Both the outer bladder wall and the inner bladder wall were outlined on each 2D slice to form a 3D surface contour of both bladder walls. The 172 bladders contained a total of approximately 16,000 slices, which equates to about 100 slices per bladder.

2.2 Bladder likelihood map generation using deep-learning convolutional neural network (DL-CNN)

We applied the DL-CNN called Cuda-Convnet developed by Krizhevsky[31] to the classification of voxels of 2D CTU slices as being inside the bladder wall or outside the bladder wall. The DL-CNN was trained using regions of interest (ROIs) extracted from the 2D slices of the training cases and labeled as within or not within the bladder wall. After training, the DL-CNN was applied to all pixels of the CTU images in the training and test set to estimate the likelihood that an ROI was within the bladder wall. The output was then assigned to the center pixel of the ROI, resulting in a likelihood map where brighter pixels were more likely to be within the bladder wall and the stack of the resulting 2D likelihood maps on the CT slices formed a 3D likelihood map for the bladder wall.

2.2.1 DL-CNN components

The components and layers of the DL-CNN are described here. More information on these components can be found in the literature[31,36].

Neurons: A neuron is defined by the following function,

$$f(x) = \max(0, x) \tag{1}$$

where x is the sum of the weighted inputs to the neuron. The activation function used in the DL-CNN is a RE-LU (eq. (1)), which converges faster than the typically used sigmoid activation function[38].

Convolution layer: The convolutional layer uses convolutional kernels to convolve with the input ROI. The resulting values are then input to the corresponding neurons in the kernel maps within the convolutional layer. The output of these neurons is determined by the activation function in eq. (1).

Locally-connected convolutional layer: The locally-connected convolutional layer performs the same set of operations as the convolutional layer, but different convolutional kernels are applied at each location of the input image and then collected into the corresponding neurons instead of using a single convolutional kernel.

Fully-connected layer: The fully-connected layer connects all inputs from the previous layer. Every map element multiplied by a trained weight is used as input. The fully-connected layer outputs values, which are passed onto the next fully connected layer or finally to a Softmax layer to obtain the final output of the DL-CNN. There can be one or more fully-connected layers before the Softmax layer in a DL-CNN architecture.

Pooling layer: Pooling layers summarize and reduce the outputs of groups of neighboring neurons. Our pooling layers use overlapping pooling due to its tendency to reduce overtraining.

Local Response Normalization layer: Local normalization layers aid in the generalization of training by normalizing input parameters to subsequent layers of the DL-CNN[39]. The output of neurons was normalized by the following equation:

$$b_{x,y}^{i} = \frac{a_{x,y}^{i}}{\left(1 + \frac{\tau}{N}\sum_{j=\max\left(0,i-\frac{N}{2}\right)}^{\min\left(n,i+\frac{N}{2}\right)}\left(a_{x,y}^{j}\right)^{2}\right)^{\varepsilon}} \quad (2)$$

where $b_{x,y}^{i}$ is the response-normalized neuron activity, $a_{x,y}^{i}$ is the neuron activity computed by applying the kernel i at the coordinates (x, y), n is the number of kernel maps, and N, τ, and ε are constants. Our implementation of the DL-CNN used N = 9, τ = 0.001, and ε = 0.75, which was demonstrated to be effective[31,36]. These parameters were taken from Krizhevsky et al.[31] and were further reinforced as ideal choices from our own experimentation.

2.2.2 DL-CNN architecture

The architecture of the DL-CNN used in this study consists of five main layers: two convolutional layers, two locally-connected convolutional layers, and one fully-connected layer, in that order (Figure 2). The first two convolutional layers are followed by a pooling and by a normalization layer before proceeding to the next layer.

The first two convolutional layers consist of 64 kernels of size 5 x 5. An ROI is input to the first convolutional layer. The output of the first layer is pooled and normalized and then used as input in the second convolutional layer. After another pooling and normalization layers, the output of the second convolutional layer is sent to a locally-connected convolutional layer with 64 kernels of size 3 x 3. The second locally connected convolutional layer contains 32 kernels of size 3 x 3. The fully connected layer outputs two values used as input to the Softmax layer defined by the function:

$$f(x_i) = \frac{e^{x_i}}{\sum_j e^{x_j}} \tag{3}$$

where x_i is each input value. The layer outputs a single value in [0,1], which represents the likelihood of the input ROI being in the bladder wall.

2.2.3 DL-CNN training

The DL-CNN was trained using the 81 cases in the training set. Each 2D axial slice of the CTU scan was divided into overlapping ROIs of M x M pixels. Following the inner and outer wall of the bladder of each slice hand outlined by a radiologist, the ROIs were labeled as within or outside the bladder wall. Three ROI sizes, M = 8, 16, and 32 were studied and the size of 16 x 16 pixels was chosen. For each 16 x 16 ROI, the central 8 x 8 pixel image was used in the labeling criterion. If 70% of the inner 8 x 8 pixel image fell within the bladder wall, i.e., the pixels were located between the hand-outlined contours of the inner wall and the outer wall, the ROI was labeled as inside the bladder wall. 70% was chosen to ensure that a large fraction of the ROI area was within the wall, while not making the classification requirement so strict as to having too few ROIs that could be labeled as within the wall because some of the walls were very thin. If 95% of the inner ROI fell within the interior of the bladder, or if less than 10% of the ROI was located inside the bladder delineated by the outer wall contour, the ROI was labeled as not within the bladder wall. These classification requirements were chosen to properly label ROIs along the wall as not within

the wall while still obtaining a sufficient number of ROIs within the wall. Figure 3 shows a bladder with the ROI images superimposed on top.

Approximately 240,000 ROIs were extracted from the training cases after balancing. Exactly half of the ROIs were labeled as within the bladder wall and the other half as not within the bladder wall. Subsets of the ROIs are presented in Figure 4.

The DL-CNN was trained for 1500 iterations based on our previous experience[36] on the training set to produce the bladder likelihood maps. The training process typically took about seven to eight hours to complete with a Nvidia Tesla K20 GPU.

2.2.4 Bladder likelihood map generation

We applied the trained DL-CNN to the CTU scans to estimate the likelihood of a given voxel being within the bladder wall. For each scan, a rectangular box enclosing the bladder was drawn to denote the volume of interest (VOI). The trained DL-CNN was applied to each voxel within the VOI. For each voxel, a 16 x 16 pixel ROI of the axial slice centered at that voxel was input to the DL-CNN, which output a likelihood score between 0 and 1 that represented the likelihood that the center pixel of the ROI was within the wall. The scores for each ROI were assigned to the center voxel in order to create a map of the output scores from the DL-CNN. The collection of likelihood scores for all voxels in the VOI provided the 3D bladder wall likelihood map. Examples of bladder wall likelihood maps on the axial slices are shown in Figure 5.

2.3 Segmentation of bladder walls using level sets

After the generation of the bladder likelihood maps, we developed a system to segment the inner and outer bladder wall using level sets. The system takes in the 3D bladder wall likelihood map and incorporates it within the level set energy functions to guide the segmentation and creates a contour of the inner or outer wall depending on which set of parameters is used. The system consists of four stages: (1) preprocessing, (2) initial segmentation, (3) 3D level set segmentation, and (4) 2D level set segmentation.

The first stage consists of preprocessing techniques to create a set of gradient vector images. Smoothing, anisotropic diffusion, gradient filters, and the rank transform of the gradient magnitude are applied to slices within the VOI. The vector images created at this stage are used during the propagation of the 3D level sets.

The second stage generates the initial segmentation surface and consists of three parts. First, we apply a threshold to the bladder wall likelihood maps to create a binary mask. The mask DL_{Mask} is generated according to the following criterion:

$$DL_{Mask}(x,y) = \begin{cases} 1, & DL_{Score}(x,y) \geq \theta \\ 0, & DL_{Score}(x,y) < \theta \end{cases} \quad (4)$$

where x and y are coordinates in the image, DL_{Mask} is the value of the resulting binary mask at the given coordinate, DL_{Score} is the value of the likelihood map score at the given coordinate, and θ is a chosen value for the threshold. The value of θ was determined experimentally[36] by histogram analysis. By using the training cases a histogram of the DL-CNN likelihood scores for the pixels inside and outside of the bladder within the VOIs was generated. The threshold of 0.85 provided a good separation of the two classes and the best contours that did not leak to the outside of the bladder wall while closely approaching the hand segmentation for cases in the training set.

The second part of the initial segmentation process is the creation of an ellipsoid as the object region for segmentation. An ellipsoid with major and minor axis 1.5 times the width and height of the VOI, respectively, is placed at the centroid of the bladder mask. The intersection between this ellipsoid and bladder mask is the object region. This ellipsoid prevents the segmentation from leaking into the surrounding structures outside the bladder by limiting the area to which the contours can expand. This limited object region prevents interference from non-bladder structures and solves the problem of high likelihood scores on other regions of the CT scan such as the pelvic bone, which often receives high likelihood scores.

Finally, neighboring components in the object region are connected using a morphological dilation filter with a 2-voxel-radius spherical structuring element, a 3D flood fill algorithm, and a morphological erosion filter with a 2-voxel-radius spherical structuring element. This process allows for the extraction of an initial segmentation surface $\phi_0(\vec{x})$ from the bladder wall likelihood map for use in the level set propagation.

The third stage of the system propagates the initial segmentation towards the inner or outer bladder wall using cascading level sets. The level set contours are generated according to the chosen level set equation:

$$\frac{\partial}{\partial t}\Psi(\vec{x}) = -\alpha A(\vec{x})\nabla\Psi(\vec{x}) - \beta P(\vec{x})|\nabla\Psi(\vec{x})| + \gamma\kappa(\vec{x})|\nabla\Psi(\vec{x})|, \quad (5)$$

$$\Psi(\boldsymbol{x}, n = 0) = \phi_0(\vec{x})$$

α, β, and γ are the coefficients for the advection, propagation, and curvature terms, respectively. $A(\vec{x})$ is a vector field image where each voxel in the image is a vector. $P(\vec{x})$ is a scalar speed term between 0 and 1[23]. $\kappa(\vec{x})$ is the mean curvature of the level set at the point $\vec{x}$ and can be defined as

$$\kappa(\vec{x}) = \mathrm{div}(\nabla\Psi(\vec{x})/|\nabla\Psi(\vec{x})|). \quad (6)$$

The symbol ∇ denotes the gradient operator, div is the divergence operator, $\phi_0(\vec{x})$ is the initial segmentation, and n is the number of iterations. The bladder wall likelihood map is used as the $\Psi(\vec{x})$ term in the fourth level set as an integral part of the advection, propagation and curvature terms, while the original CTU volume is used for the first three level sets.

The advection term drives the contour towards regions of high gradient according to the $A(\vec{x})$ vector. The propagation term controls the expansion of the contours according to local pixel information. The curvature term causes the contour to maintain relative shape and curvature while expanding.

Four different 3D level sets are applied sequentially. The first level set expands the initial contours slightly and smooths edges. The second level set brings the contours towards sharp edges and draws it slightly in regions of low gradient. The third level set finishes drawing the contour toward edges. The fourth level set draws the contours towards the inner and outer bladder walls, using the bladder wall likelihood map for the $\nabla\Psi(\vec{x})$ term in equation 5. The level set equations are run twice, once for the outer contour and once for the inner contour, using different parameters.

As the final step in the level set system, 2D level sets are applied to each slice of the 3D segmented object, using the 3D level set generated contours on each slice as an initial contour, to

further refine the contours. Further details on the methods of cascading level sets and the necessity of 2D level sets to refine segmentation can be found in our previous publications[16,36,40].

Table 1 shows the coefficient information for each round of level sets. The differences in the runs for outer and inner wall segmentation are the parameters α, β, and γ and the number of iterations (n) in the fourth level sets. Notably, the propagation coefficient β is positive when running the level set equations for the outer wall so the contour expands to the outer wall and negative for the inner wall so the contour propagates inwards. The coefficients used for the first 3 rounds of level sets are the same as those used in our previous bladder segmentation[36]. The parameter q in Table 1 is defined by the linear function $\sigma M + \phi$, where M is the 2D diagonal distance of the VOI in mm, $\sigma = 0.06, and\ \phi = -0.11$ as shown in previous work[36].

2.4 Evaluation Methods

Segmentation performance was evaluated by comparing the generated contours to 3-D hand segmented contours. The inner and outer walls were each independently compared to the hand outlined inner and outer wall contours. We calculated the volume intersection ratio, the volume error, the absolute volume error, and the minimum distance as performance metrics, defined below. Additionally, we compared the area between the inner and outer wall contours to that of the hand outlined-contours using the volume intersection, the volume error, and the absolute volume error.

The volume intersection is the ratio of the intersection between the given volume enclosed by the contour generated by the level sets and the reference volume enclosed by the reference contour, to the reference volume:

$$R^{3D} = \frac{V_R \cap V_U}{V_R} \tag{7}$$

where V_U is the given volume and V_R is the reference volume.

The volume error is defined as the ratio of the difference between the reference volume and the given volume to the reference volume:

$$E^{3D} = \frac{V_R - V_U}{V_R} \tag{8}$$

Since the given volume is segmented by the level sets, a positive error represents under-segmentation of the bladder wall, while a negative error indicates over-segmentation. We also calculate the absolute volume error $|E^{3D}|$ to show the average deviation from the reference contours. Other performance indicators can be derived from the volume intersection ratio and the volume error[41].

AVDIST is the average distance between the closest points on the reference and the given contours:

$$AVDIST = \frac{1}{2}\left(\frac{\sum_{x\in R} min\{d(\boldsymbol{x},\boldsymbol{y}): \boldsymbol{y}\in U\}}{N_R} + \frac{\sum_{y\in U} min\{d(\boldsymbol{x},\boldsymbol{y}): \boldsymbol{x}\in R\}}{N_U}\right) \tag{9}$$

R is the reference contour, U is the given contour and N_R and N_U are the number of voxels along each respective contour. The function d is the minimum Euclidian distance between a given voxel $\boldsymbol{x}$ on the contour R to a voxel $\vec{\boldsymbol{y}}$ on the contour U. The minimum distances of all points along R are calculated and then averaged. The minimum distances of all points along U are also calculated and averaged by repeating the process with the roles of R and U switched. The overall average distance AVDIST between the two contours is then calculated from the two average minimum distances.

The volume intersection ratio, the volume error, and the absolute volume error are also calculated for the region between the inner and outer wall contours. These metrics are calculated in the same way as for the outer and inner wall contours, except that V_R and V_U are the volume of the shell enclosed between the outer wall contour and the inner wall contour.

The volume intersection ratio is also calculated for the lesions and the outer wall contour to determine what percent of the lesions is successfully enclosed by the segmentation.

3. RESULTS

3.1 Segmentation performance of inner and outer wall contours

Examples of segmentation results from the test cases are shown in Figure 6. The DL-CNN bladder wall likelihood maps that were used to guide the segmentation are also shown.

Histograms for the training and test sets for the volume intersection ratio, the volume error, and the average distance are shown in Figure 7.

The average segmentation performances for the training and test sets using the different performance metrics are shown in Table 2.

For the training set, the inner wall contour achieved average volume intersection ratio, average volume error, average absolute volume error, and average distance of 90.2±8.7%, -4.3± 18.2%, 12.6±13.7%, and 3.0±1.6mm, respectively. For the test set, the inner wall achieved values of 87.2±10.5%, -5.3±28.2%, 15.6±24.0%, and 3.2±1.7mm, respectively. For the training set, the outer wall contour achieved the values of 93.2±5.8%, -7.2±12.3%, 10.4±9.6%, and 3.0±1.2mm, respectively. For the test set, the outer wall contour achieved values of 89.5±9.8%, -6.2±20.5%, 14.6±15.6%, and 3.5±2.0mm, respectively.

3.2 Performance of combined wall contours

For the segmented bladder wall, the training set achieved volume intersection ratio, volume error, and absolute volume error of 61.0±11.3%, -13.7±49.1%, and 34.5±37.3%, respectively. The test set achieved values of 54.6±10.4%, 10.7±28.0%, and 25.1±15.8%, respectively.

Due to the small volume of the wall compared to the entire bladder volume enclosed by either the inner or the outer walls, small differences between the contours lead to large fluctuations in the measured values, which is the main reason that the performance metrics are worse for the segmented wall than the individual inner and outer wall contours.

3.3 Lesion Intersection

To determine if the DL-CNN with level sets enclosed bladder lesions within the contours, we evaluated the volume intersection between the lesion as the reference contour and the outer wall as the given contour. Histograms of the result are shown in Figure 8. For the DL-CNN outlined walls, the average volume intersection ratio is 80.3 ± 23.8% for the training set and 81.6 ± 16.6% for the test set. 70.2% of the lesions have a volume intersection ratio over 75% with the DL-CNN assisted level sets contours compared to 89.1% for the hand outlined contours.

3.4 Level Set Method Comparison

Table 3 shows a comparison between performing segmentation by using the likelihood map as the $\Psi(\vec{x})$ term in only the fourth level set and by using the likelihood map as the $\Psi(\vec{x})$ term in all four stages of the cascading level sets. When the likelihood map was used only in the fourth level set the volume intersection index was significantly better (p-value < 0.01) for the outer wall in both the training and test set. For the inner wall, the difference was not significantly different in either the training or the test set.

To further demonstrate the effectiveness of the level set method with the DL-CNN, we compared the DL-CNN with level sets method to contours created using the DL-CNN without level sets (Table 4). The DL-CNN without level sets tended to over segment the inner wall but under-segment the outer wall. With the refinement by the level sets, both the over-segmentation of the inner wall and under-segmentation of the outer wall were reduced significantly ($p < 0.01$), as indicated by the volume error, absolute volume error, and average distance.

4. Discussion

We developed a segmentation method to extract the bladder wall from the interior of the bladder and the surrounding structures. The method uses a deep learning convolutional neural network combined with a series of cascading level sets to detect the contours for the inner and outer walls of the bladder in CTU scans. Segmentation of the wall presents many challenges; some are associated with segmentation of the bladder in general, and others are unique to the segmentation of the wall specifically. The training and test sets of CTU scans contain bladders completely filled with contrast material, partially filled, and not filled. The boundary between the wall and the interior of the bladder is drastically different depending on whether contrast material is present. The combination of the DL-CNN and level sets is able to account for the difference in the regions, demonstrating that the DL-CNN can overcome the strong barrier between regions, even for narrow regions such as the bladder wall.

The small thickness of the bladder wall presents other problems for segmentation as well. Bladder walls can vary in thickness drastically compared to the full bladder. Thin walls can be only

few pixels wide compared to bladder walls with urothelial thickenings. The DL-CNN often has difficulties with this variation while the level sets are able to account for this difference. The DL-CNN and the level sets play complementary roles in segmenting the bladder wall and generating the likelihood maps.

The presence of lesions also presents another confounding factor for segmentation of the wall. Our DL-CNN assisted level sets method often generates inner wall contours through the lesions rather than around them (Figure 9), although the bladder wall likelihood maps typically include the lesions accurately during the estimation of the bladder wall.

The 16 x 16 ROI size was chosen after experimenting with ROIs of 8 x 8, 16 x 16, and 32 x 32 pixels in size. The 8 x 8 pixel ROIs tended to generate too much noise in the bladder wall likelihood maps, specifically along the barrier between the contrast and non-contrast region. 8 x 8 pixel ROIs also had the tendency to create gaps in parts of the wall that were extremely thin (Figure 10b). 32 x 32 pixel ROIs resulted in walls that were far too thick in the bladder likelihood maps. Due to the small thickness of the wall, ROIs of 32 x 32 pixels were too large to accurately represent the wall (Figure 10d). The 16 x 16 pixel ROIs were chosen as a middle ground because they provided bladder wall likelihood maps that did not have as many gaps or as much noise as those obtained with the 8 x 8 pixel ROIs, but provided thinner walls that resulted in more accurate contours than the 32 x 32 pixel ROIs (Figure 10c).

To further prevent holes in the bladder wall likelihood maps, we developed a method for the generation of the ROIs that can capture the thin regions of the bladder wall, while still using an ROI size large enough to accurately capture the wall features in thicker walls. In this method, a 16 x 16-pixel ROI is used to train the deep learning algorithm, but only the central 8 x 8-pixel area of the ROI is used for the labeling of the ROI, referred to as 8 x 8 jittering. If 70% of this inner 8 x 8 pixel area falls between the hand-outlined outer and inner wall contours, the ROI is labeled as within the bladder wall; otherwise, the ROI is classified as not within the bladder wall. Using the central smaller area rather than the whole ROI allows more ROIs to be correctly labeled as within a thin wall, while preventing more ROIs outside the wall from being mislabeled as within the wall.

This jittering technique combined with the added level set equation were important methodological changes from the method used in Cha et al.[36] The addition of another round of level

sets that incorporate the DL-CNN likelihood map in the equation allow the level sets to segment both the inner and outer bladder wall as opposed to just the entire bladder as a whole.

The only small drawback in the use of 16 x 16 ROIs is a slight increase in the training time; an increase from 5.5 to 6.5 hours due to an increase in the number of training ROIs compared to the use of 32 x 32 ROIs. However, both 16 x 16 and 32 x 32 ROI sizes take about 4 min to generate a bladder likelihood map during deployment so this drawback is a minor one. The ROI size does not affect the time required for running the level sets either, which takes about 2 to 5 minutes per contour.

The likelihood map was chosen to be used only in this fourth level set based on our experimental results (Table 3). When the likelihood map was used only in in the fourth level set the segmentation results were more accurate. The fourth level set is crucial to differentiating between the outer wall and the inner wall for the level set segmentation.

In the cascaded level sets pipeline, each level set shows progressively improved results over the previous level set, justifying the need for each subsequent level set. Table 5 shows the volume intersection ratios from the third to the last level sets.

The inclusion of lesions within the bladder wall contours is important for the bladder wall segmentation as the segmented wall defines the search region for bladder lesions in the subsequent steps of a CAD system. A segmentation method that can reliably include all lesions is therefore critical for the CAD system to identify cancerous lesions or find wall thickenings. The results of this study show that our proposed method is promising for this application.

Compared to the study by Cha et al.[36] on the same data set, the current study showed better performance. The purpose of the study by Cha et al.[36] was to segment the entire bladder so that they only needed to segment the outer bladder wall and also used a DL-CNN with level sets. The volume intersection ratios for the training set and test set achieved in the current study were significantly better (p value < 0.01) than those in the previous study. The volume intersection ratios were 93.7% and 89.5% for the training and test sets, respectively, in this study compared to 84.2% and 78.0%, respectively, by Cha et al[36]. The differences in the absolute percent volume errors for the training set and test set are not significant, however.

In order to demonstrate the effectiveness of our proposed deep learning approach to this problem, we compared the performance for the outer wall segmentation from this study with the previous attempt at segmenting the same data set using an adaptive level set system (CLASS)[16]. We also applied a commonly used image segmentation tool, ITK-SNAP, to a subset of our CTU data as it was used by Ma et al. [13,14] to segment the bladder walls in T2-wieghted MR images. The performance of the DL-CNN with level sets is significantly better (p-value < 0.01) than that of ITK-SNAP and CLASS for the metrics of volume intersection and absolute volume error. Table 6 shows the comparison of the performances by CLASS and DL-CNN. CLASS is a level set system that does not use a DL-CNN to assist with segmentation. The improved results of the DL-CNN with level sets show the value of a deep learning approach to bladder segmentation. Table 7 shows the average values for a smaller data set of 30 bladders used for ITK-SNAP alongside values from the same data set for CLASS and DL-CNN. Even with the limited scope of previous experiments, the performance of the DL-CNN with level sets is significantly better than that of ITK-SNAP and CLASS with a p value less than 0.01 for the metrics of volume intersection and absolute volume error on the smaller data set.

One limitation of our current segmentation method was observed for CT scans with thick slices and an unusually small bladder. For small bladders, the level sets cannot change fast enough to account for the rapid differences between slices when slices are 5mm thick. This problem could potentially be overcome by selecting a different set of parameters that is optimized for thick slices and the small bladders. An automated pre-processing triage stage that can recognize small bladders with thick slices has to be developed and call for the specific set of parameters for the segmentation of these outlier cases. We will continue to optimize the parameters of the level sets in future studies.

The small thickness of the wall also presents problems in the evaluation of the DL-CNN and level set contours. While the outer and inner contours individually achieved good results compared to the reference standard, the overlap of the wall between the inner and outer contours achieved much lower results in terms of the volume intersection and the volume error. Since the bladder wall can be very thin, small deviations in the wall contours can lead to massive errors that are not observed when comparing the bladder as a whole. Cases in both the training and the test sets with

good performance for the inner and outer wall contours had noticeably poorer performance in the wall between the two contours due to these small deviations. Examples of bladder wall segmentations are presented in Figure 11.

Another limitation associated with the method is the reliance on hand outlines from a single radiologist. Generating outlines from multiple radiologists and taking the average will reduce the bias for the reference standard, which will allow better estimation of the DL-CNN performance. However, due to the excessive time required to generate manual hand outlines for such a large data set (a total of approximately 16,000 slices of bladder outlines), only one set of hand outlines was obtained and used for evaluation.

5. Conclusions

Our results in this study show that the DL-CNN assisted level set method is useful for segmentation of the bladder wall in CTU scans. This method can specifically segment the bladder wall by detecting both the inner and outer wall contours of the bladder. The DL-CNN can accurately differentiate the bladder wall from the interior of the bladder and the surrounding structured background. The segmentation of the wall is an important first step for automated lesion detection and bladder analysis. Further studies are underway to improve the performance of the wall segmentation process, especially for the inclusion of lesions within the wall contours. This study lays the important groundwork for further efforts to develop computerized decision support systems for diagnosis and treatment of bladder cancer.

ACKNOWLEDGEMENT

This work is supported by National Institutes of Health grant number U01CA179106.

REFERENCES

[1]*American Cancer Society. Cancer Facts & Figures 2017.* , (American Cancer Society, Inc., Atlanta, 2017).

[2]S. A. Akbar, K. J. Mortele, K. Baeyens, M. Kekelidze, and S. G. Silverman, "Multidetector CT urography: Techniques, clinical applications, and pitfalls," Seminars in Ultrasound CT and MRI 25, 41-54 (2004).

[3]E. M. Caoili, R. H. Cohan, M. Korobkin, J. F. Platt, I. R. Francis, G. J. Faerber, J. E. Montie, and J. H. Ellis, "Urinary tract abnormalities: Initial experience with multi-detector row CT urography " Radiology 222, 353-360 (2002).

[4]W. C. Liu, K. J. Mortele, and S. G. Silverman, "Incidental extraurinary findings at MDCT urography in patients with hematuria: Prevalence and impact on Imaging costs," American Journal of Roentgenology 185, 1051-1056 (2005).

[5]C. L. McCarthy and N. C. Cowan, "Multidetector CT urography (MD-CTU) for urothelial imaging," Radiology (P) 225, 237 (2002).

[6]M. Noroozian, R. H. Cohan, E. M. Caoili, N. C. Cowan, and J. H. Ellis, "Multislice CT urography: State of the art " British Journal of Radiology 77, S74-S86 (2004).

[7]S. B. Park, J. K. Kim, H. J. Lee, H. J. Choi, and K.-S. Cho, "Hematuria: portal venous phase multi detector row CT of the bladder--a prospective study," Radiology 245, 798-805 (2007).

[8]G. S. Sudakoff, D. P. Dunn, M. L. Guralnick, R. S. Hellman, D. Eastwood, and W. A. See, "Multidetector computerized tomography urography as the primary imaging modality for detecting urinary tract neoplasms in patients with asymptomatic hematuria," Journal of Urology 179, 862-867 (2008).

[9]L. Li, Z. Wang, X. Li, X. Wei, A. H. L., W. Huang, S. Rizvi, M. H., D. P. Harrington, and Z. Liang, "A new partial volume segmentation approach to extract bladder wall for computer aided detection in virtual cystoscopy," Proc. SPIE 5369, 199-206 (2004).

[10]C. Duan, Z. Liang, S. Bao, H. Zhu, S. Wang, G. Zhang, J. J. Chen, and H. Lu, "A coupled level set framework for bladder wall segmentation with application to MR cystography," IEEE Trans Med Imaging 29, 903-15 (2010).

[11]C. J. Duan, K. H. Yuan, F. H. Liu, P. Xiao, G. Q. Lv, and Z. R. Liang, "An Adaptive Window-Setting Scheme for Segmentation of Bladder Tumor Surface via MR Cystography," IEEE Transactions on Information Technology in Biomedicine 16, 720-729 (2012).

[12]H. Han, L. Li, C. Duan, H. Zhang, Y. Zhao, and Z. Liang, "A unified EM approach to bladder wall segmentation with coupled level-set constraints," Medical Image Analysis 17, 1192-1205 (2013).

[13]X. F. Chai, M. van Herk, A. Betgen, M. Hulshof, and A. Bel, "Automatic bladder segmentation on CBCT for multiple plan ART of bladder cancer using a patient-specific bladder model," Physics in Medicine and Biology 57, 3945-3962 (2012).

[14] Z. Ma, R. N. Jorge, T. Mascarenhas, J. M. Tavares, "Novel Approach to Segment the Inner and Outer Boundaries of the Bladder Wall in T2-Weighted Magnetic Resonance Images," Annals of Biomedical Engineering 39, 2287-2297 (2011).

[15]S. M. B. J. Wenjun Chi , Niall R. Moore and Julia A. Schnabel, "Segmentation of the Bladder Wall Using Coupled Level Set Methods," Chicago, IL, USA, 30 March-2011.

[16]L. Hadjiiski, H.-P. Chan, Y. Law, R. H. Cohan, E. M. Caoili, H. C. Cho, C. Zhou, and J. Wei, "Segmentation of Urinary Bladder in CT Urography (CTU) using CLASS," Proc. SPIE 8315, 83150J1-83150J7 (2012).

[17]L. M. Hadjiiski, B. Sahiner, H.-P. Chan, E. M. Caoili, R. H. Cohan, and C. Zhou, "Automated segmentation of urinary bladder and detection of bladder lesions in multi-detector row CT urography," Proc. SPIE 7260, 72603R1- 72603R7 (2009).

[18]L. Hadjiiski, H. P. Chan, R. H. Cohan, E. M. Caoili, Y. Law, K. Cha, C. Zhou, and J. Wei, "Urinary bladder segmentation in CT urography (CTU) using CLASS," Medical Physics 40, 111906 (2013).

[19]K. H. Cha, L. M. Hadjiiski, H.-P. Chan, E. M. Caoili, R. H. Cohan, and C. Zhou, "CT urography: segmentation of urinary bladder using CLASS with local contour refinement," Phys Med Biol 59, 2767-85 (2014).

[20]K. Fukushima, S. Miyake, and T. Ito, "Neocognitron: A neural network model for a mechanism of visual pattern recognition," IEEE Trans Systems Man Cybernetics SMC-13, 826-834 (1983).

[21]S. C. B. Lo, J. S. Lin, M. T. Freedman, and S. K. Mun, "Computer-assisted diagnosis of lung nodule detection using artificial convolution neural network," Proc. SPIE 1898, 859-869 (1993).

[22]H.-P. Chan, S. C. B. Lo, M. A. Helvie, M. M. Goodsitt, S. N. C. Cheng, and D. D. Adler, "Recognition of mammographic microcalcifications with artificial neural network," Radiology 189(P), 318 (1993).

[23]H.-P. Chan, B. Sahiner, S. C. B. Lo, M. A. Helvie, N. Petrick, D. D. Adler, and M. M. Goodsitt, "Computer-aided diagnosis in mammography: Detection of masses by artificial neural network," Medical Physics 21, 875-876 (1994).

[24]W. Zhang, K. Doi, M. L. Giger, Y. Wu, R. M. Nishikawa, and R. A. Schmidt, "Computerized detection of clustered microcalcifications in digital mammograms using a shift-invariant artificial neural network," Med Phys 21, 517-524 (1994).

[25]H.-P. Chan, S. C. B. Lo, B. Sahiner, K. L. Lam, and M. A. Helvie, "Computer-aided detection of mammographic microcalcifications: Pattern recognition with an artificial neural network," Medical Physics 22, 1555-1567 (1995).

[26]S. C. B. Lo, H.-P. Chan, J. S. Lin, H. Li, M. Freedman, and S. K. Mun, "Artificial Convolution neural network for medical image pattern recognition," Neural Networks 8, 1201-1214. (1995).

[27]B. Sahiner, H.-P. Chan, N. Petrick, D. Wei, M. A. Helvie, D. D. Adler, and M. M. Goodsitt, "Classification of mass and normal breast tissue: A convolution neural network classifier with spatial domain and texture images," IEEE Transactions on Medical Imaging 15, 598-610 (1996).

[28]M. N. Gurcan, B. Sahiner, H.-P. Chan, L. M. Hadjiiski, and N. Petrick, "Selection of an optimal neural network architecture for computer-aided detection of microcalcifications - comparison of automated optimization techniques," Medical Physics 28, 1937-1948 (2001).

[29]M. N. Gurcan, H.-P. Chan, B. Sahiner, L. Hadjiiski, N. Petrick, and M. A. Helvie, "Optimal neural network architecture selection: Improvement in computerized detection of microcalcifications," Academic Radiology 9, 420-429 (2002).

[30]R. K. Samala, H.-P. Chan, Y. Lu, L. M. Hadjiiski, J. Wei, and M. A. Helvie, "Digital breast tomosynthesis: computer-aided detection of clustered microcalcifications on planar projection images," Physics in Medicine and Biology 59, 7457-7477 (2014).

[31]A. Krizhevsky, "cuda-convnet,", http://code.google.com/p/cuda-convnet/

[32]A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," Advances in neural information processing systems, 1097-1105, 2012

[33]O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, and M. Bernstein, "Imagenet large scale visual recognition challenge," International Journal of Computer Vision 1-42 (2014).

[34]A. Krizhevsky, "Learning Multiple Layers of Features from Tiny Images," Master's thesis, Department of Computer Science, University of Toronto, 2009.

[35]G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. van der Laak, B. van Ginneken, and C. I. Sánchez, "A Survey on Deep Learning in Medical Image Analysis," Medical Image Analysis, 42, 60-88, (2017).

[36]K. H. Cha, L. Hadjiiski, R. K. Samala, H. P. Chan, E. M. Caoili, and R. H. Cohan, "Urinary bladder segmentation in CT urography using deep-learning convolutional neural network and level sets," Medical Physics 43, 1882-1896 (2016).

[37]M. Gordon, L. Hadjiiski, K. Cha, H.-P. Chan, R. Samala, R. H. Cohan, and E. M. Caoili, "Segmentation of inner and outer bladder wall using deep-learning convolutional neural networks in CT urography," Proc SPIE 10134, 1013402-1013402-7 (2017).

[38]V. Nair and G. E. Hinton, "Rectified linear units improve restricted boltzmann machines,"

[39]S. Ioffe and C. Szegedy, "Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift," (2015).

[40]E. Street, L. Hadjiiski, B. Sahiner, S. Gujar, M. Ibrahim, S. K. Mukherji, and H.-P. Chan, "Automated Volume Analysis of Head and Neck Lesions on CT Scans Using 3D Level Set Segmentation," Medical Physics 34, 4399-4408 (2007).

[41]T. W. Way, L. M. Hadjiiski, B. Sahiner, H.-P. Chan, P. N. Cascade, E. A. Kazerooni, N. Bogot, and C. Zhou, "Computer-aided diagnosis of pulmonary nodules on CT scans: segmentation and classification using 3D active contours," Medical Physics 33, 2323-2337 (2006).

The authors have no conflicts to disclose.

Figure 1: Flowchart of DL-CNN segmentation that shows generation of a bladder wall likelihood map and use of level sets to obtain inner and outer wall contours. The DL-CNN likelihood map was used as a gradient image in the energy equation of a cascaded level sets method.

Figure 2. Block diagram showing our DL-CNN architecture. The input layer is the classified ROIs and the output of the trained DL-CNN creates a likelihood map.

Figure 3: ROIs superimposed on a CTU slice. The darker boxes are ROIs determined to be within the wall and the lighter boxes are ones determined not to be within the wall. The number of ROIs are balanced so that the training ROIs contain exactly half within each category.

Figure 4: 16 x 16 pixel ROIs. (a) ROIs labeled as being within the wall; (b) ROIs labeled as not within the wall.

Figure 5: Bladder wall likelihood maps (b), (d) shown with the CTU slices (a), (c) from which they were generated. The brighter pixels represent a higher likelihood of that pixel occurring in the wall.

Figure 6: (a) and (c) are bladder wall likelihood maps generated from CTU scans from the test set (b) and (d), respectively. The level set contours are overlaid on the CTU slices.

Figure. 7: Histograms of the volume intersection % (a, b), volume % error (c, d), and average distance (e, f) for the training and test sets, respectively.

Figure 8: Histogram of the lesion volume intersection. The graph (a,b) shows results both for the radiologist hand-outline contours and the DL-CNN with level set contours for the training and test set, respectively.

Figure 9: Segmented contours on a CTU slice. The inner contour goes through the lesion rather than around it. The small contours in the lower right of the image are ignored.

Figure 10: Likelihood maps of the CTU slice shown in (a). (b) was generated using 8 x 8 ROIs, (c) by 16 x 16 ROIs, and (d) by 32 x 32 ROIs. All likelihood maps were shown with the same brightness and contrast window settings.

Figure 11: Comparisons between the hand-outlined contours and the computer segmented contours. (a), (c), and (e) show the outer wall contour, and (b), (d), and (f) the inner contour. The dark and light contours in each image represent the hand-outlined contour and the computer-segmented contour, respectively. The 2D area intersection ratio of the thin wall between the inner and outer wall contours (a) 50.0%, (c) 53.0%, and (e) 55.9%.

Table 1: Level set parameters. The parameters listed in the "Fourth" row are listed as inner wall parameter/ outer wall parameter.

Level Sets	**α**	**β**	**γ**	**n**
First	1	2	1	10
Second	1	0.6	q	150
Third	0	1	0	10
Fourth	4.5/3	-2.5/2	2/2	150/100
2D Slices	4	0.2	0.5	100

Table 2: Summary of the performance metrics for the inner and outer bladder walls

		Volume Intersection (%)	**Volume Error (%)**	**Absolute Volume Error (%)**	**Average Distance (mm)**
Training	Inner	90.2±8.7	-4.3± 18.2	12.6±13.7	3.0±1.6
	Outer	93.2±5.8	-7.2±12.3	10.4±9.6	3.0±1.2
Test	Inner	87.2±10.5	-5.3±28.2	15.6±24.0	3.2±1.7
	Outer	89.5±9.8	-6.2±20.5	14.6±15.6	3.5±2.0

Table 3: Comparison between using the likelihood map for all level sets and using the likelihood map for only the fourth level set.

			Volume Intersection (%)	**Volume Error (%)**	**Absolute Volume Error (%)**	**Average Distance (mm)**
Training	Inner	All level sets	93.5±4.5	-14.0±25.0	16.5±23.3	3.0±1.4
		4^{th} level set only	90.2±8.7	-4.3± 18.2	12.6±13.7	3.0±1.6

	Outer	All level sets	81.8±8.0	13.3±9.8	14.1±8.5	3.8±1.5
		4^{th} level set only	93.2±5.8	-7.2±12.3	10.4±9.6	3.0±1.2
Test	Inner	All level sets	88.5±10.3	-8.0±27.7	16.8±23.4	3.3±1.7
		4^{th} level set only	87.2±10.5	-5.3±28.2	15.6±24.0	3.2±1.7
	Outer	All level sets	76.1±11.9	18.0±15.5	19.8±13.0	4.7±2.4
		4^{th} level set only	89.5±9.8	-6.2±20.5	14.6±15.6	3.5±2.0

Table 4: Comparison between bladder contours generated by using the DL-CNN with level sets and contours generated directly from the DL-CNN likelihood map without applying level sets.

			Volume Intersection (%)	**Volume Error (%)**	**Absolute Volume Error (%)**	**Average Distance (mm)**
Training	Inner	DL-CNN without Level Sets	97.6±2.1	-33.8±36.1	33.8±36.1	4.0±1.9
		DL-CNN with Level Sets	90.2±8.7	-4.3± 18.2	12.6±13.7	3.0±1.6
	Outer	DL-CNN without Level Sets	72.8±8.7	24.3±8.8	24.3±8.8	5.2±1.6
		DL-CNN with Level Sets	93.2±5.8	-7.2±12.3	10.4±9.6	3.0±1.2
Test	Inner	DL-CNN without Level Sets	94.2±8.0	-29.3±32.6	30.9±31.1	4.1±2.2
		DL-CNN with Level Sets	87.2±10.5	-5.3±28.2	15.6±24.0	3.2±1.7
	Outer	DL-CNN without Level Sets	76.5±12.0	17.7±15.2	19.5±13.0	4.6±2.5
		DL-CNN with Level Sets	89.5±9.8	-6.2±20.5	14.6±15.6	3.5±2.0

Table 5: The volume intersection ratios after the 3^{rd} level set, the 4^{th} level set, and the full level set cascade. The segmentation by the full cascaded level sets achieved the best performance.

	After 3^{rd} level set (%)	**After 4^{th} level set (%)**	**Full level set cascade (%)**

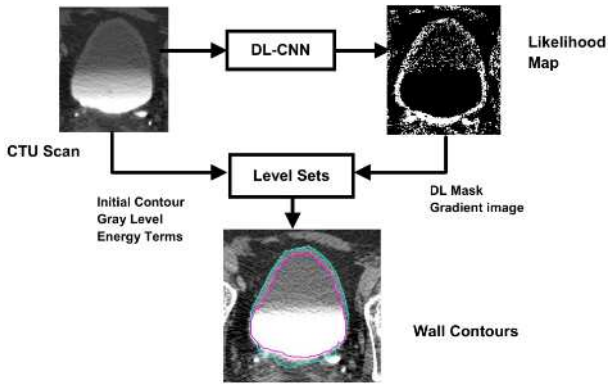
Outer Training Set	77.7±7.9	89.7±6.8	93.2±5.8
Inner Training Set	77.7±7.9	79.3±7.0	90.2±8.7
Outer Test Set	72.4±11.8	84.3±11.4	89.5±9.8
Inner Test Set	72.3±11.5	74.9±10.6	87.2±10.5

Table 6. Comparison of the outer wall segmentation using CLASS versus the proposed DL-CNN assisted level set method.

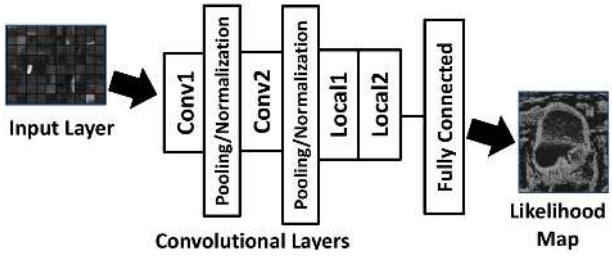
	Volume Intersection (%)	**Volume Error (%)**	**Absolute Volume Error (%)**	**Average Distance (mm)**
DL-CNN	89.5±9.8	-6.2±20.5	14.6±15.6	3.5±2.0
CLASS	84.0±11.4	8.2±17.4	13.0±14.1	3.5±1.9

Table 7. Comparison of the outer wall segmentation by the proposed DL-CNN assisted level set method to those using the ITK-SNAP and CLASS on 30 cases

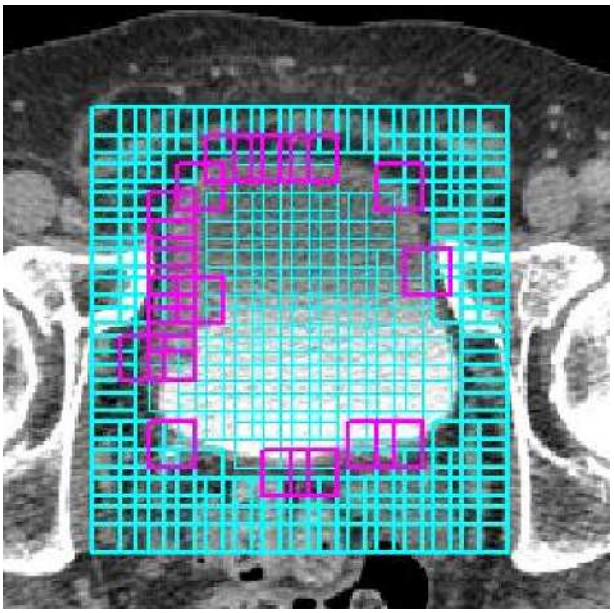
	Volume Intersection (%)	**Volume Error (%)**	**Absolute Volume Error (%)**	**Average Distance (mm)**
DL-CNN	94.4±3.2	-8.5±9.8	10.0±8.3	3.0±1.2
ITK-SNAP	78.8±16.3	8.3±33.1	24.2±23.7	5.2±2.6
CLASS	79.0±8.2	16.1±16.3	19.9±11.1	3.5±1.3



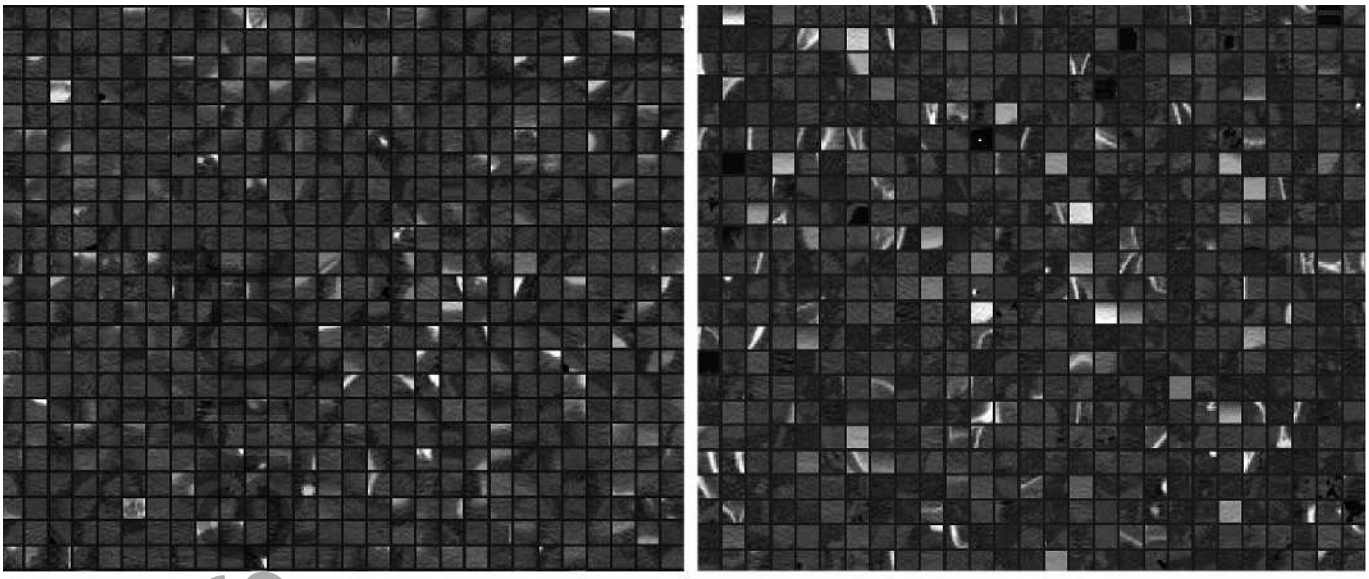
mp_13326_f1.tif



mp_13326_f2.tif

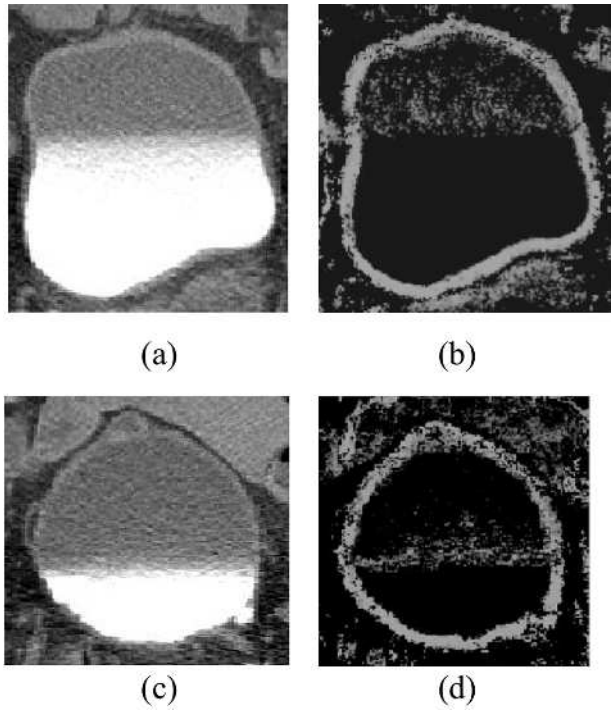


mp_13326_f3.tif

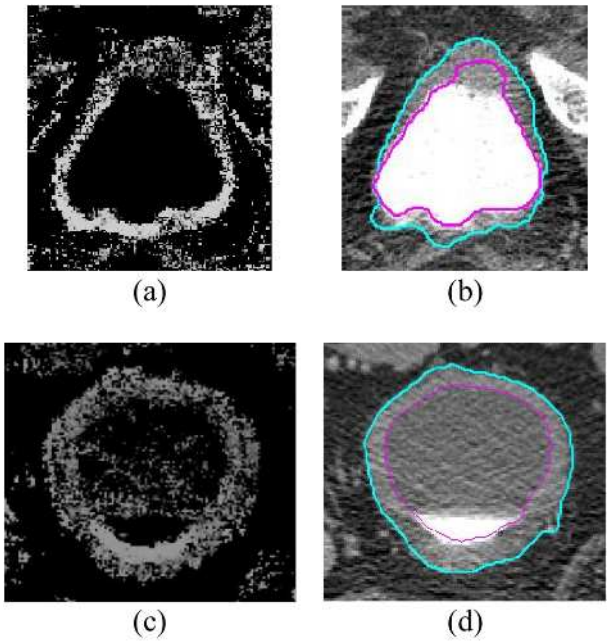


(a) (b)

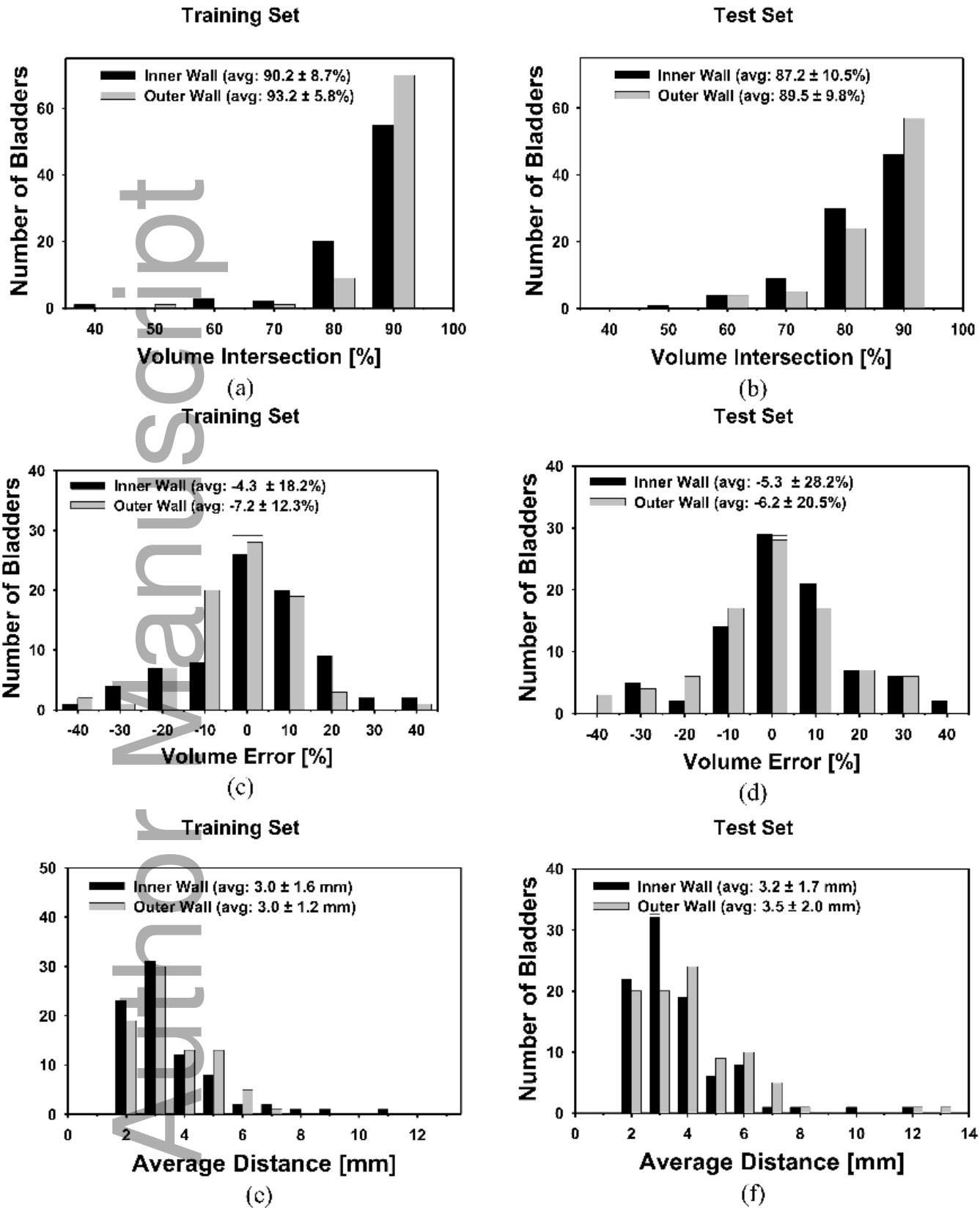
mp_13326_f4.tif



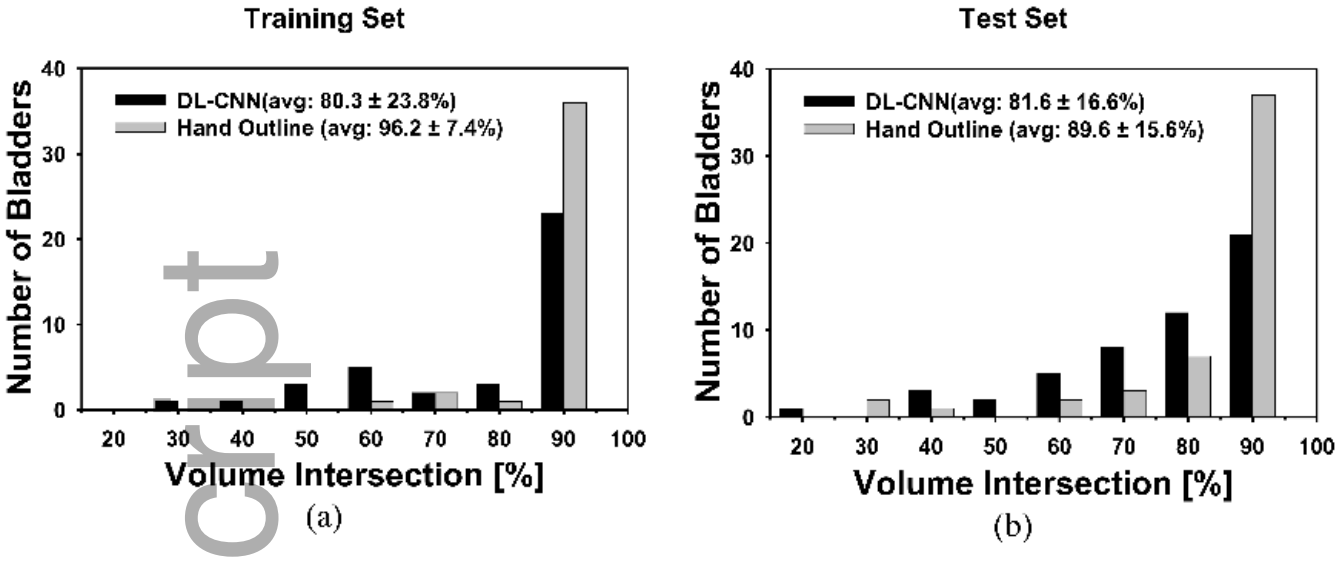
mp_13326_f5.tif



mp_13326_f6.tif

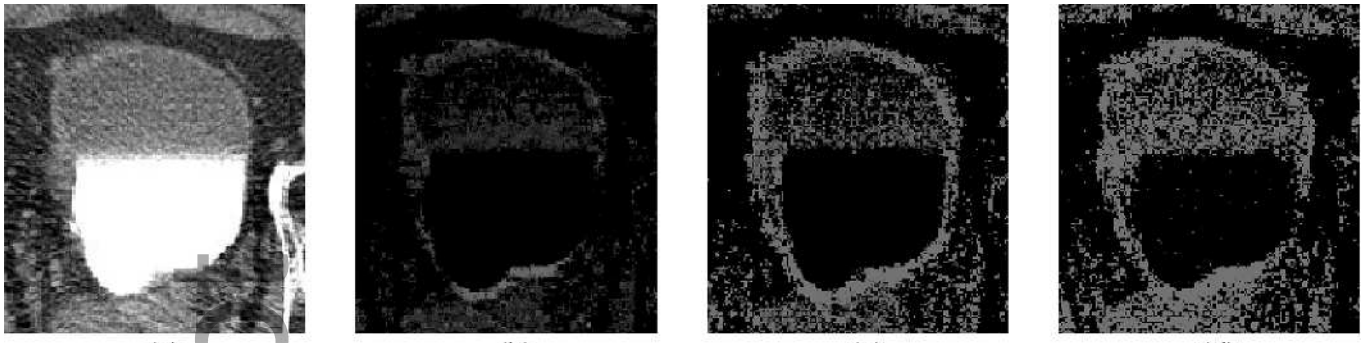


mp_13326_f7.tif

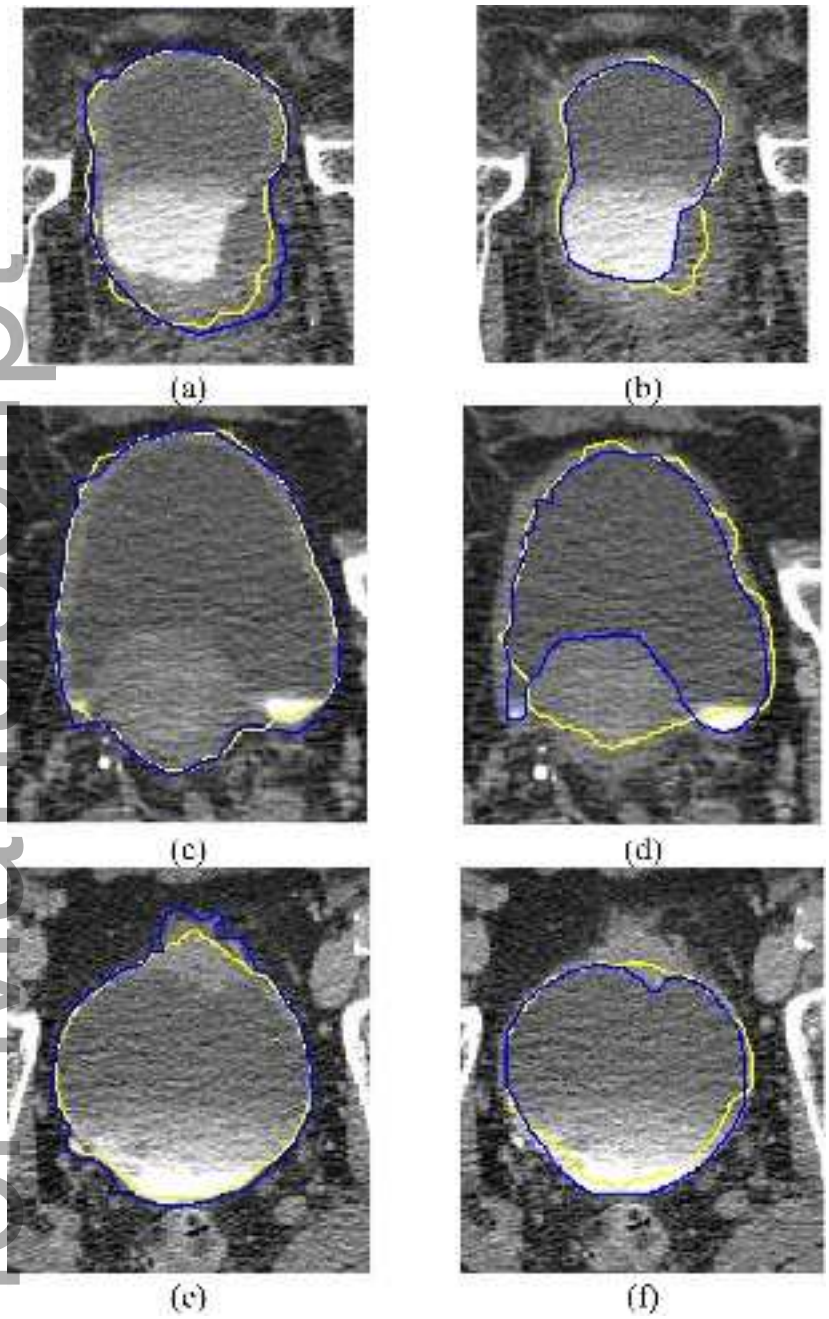


mp_13326_f8.tif

mp_13326_f9.tif



mp_13326_f10.tif



mp_13326_f11.tif