Volumetric ConvNets with Mixed Residual Connections for Automated Prostate Segmentation from 3D MR Images

Lequan Yu*, Xin Yang*, Hao Chen*, Jing Qin†, Pheng-Ann Heng*‡
*Department of Computer Science and Engineering, The Chinese University of Hong Kong
†Centre for Smart Health, School of Nursing, The Hong Kong Polytechnic University
‡Guangdong Provincial Key Laboratory of Computer Vision and Virtual Reality Technology, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, China
{lqyu,xinyang,hchen,pheng}@cse.cuhk.edu.hk {harry.qin}@polyu.edu.hk

Abstract

Automated prostate segmentation from 3D MR images is very challenging due to large variations of prostate shape and indistinct prostate boundaries. We propose a novel volumetric convolutional neural network (ConvNet) with mixed residual connections to cope with this challenging problem. Compared with previous methods, our volumetric ConvNet has two compelling advantages. First, it is implemented in a 3D manner and can fully exploit the 3D spatial contextual information of input data to perform efficient, precise and volume-to-volume prediction. Second and more important, the novel combination of residual connections (i.e., long and short) can greatly improve the training efficiency and discriminative capability of our network by enhancing the information propagation within the ConvNet both locally and globally. While the forward propagation of location information can improve the segmentation accuracy, the smooth backward propagation of gradient flow can accelerate the convergence speed and enhance the discrimination capability. Extensive experiments on the open MICCAI PROMISE12 challenge dataset corroborated the effectiveness of the proposed volumetric ConvNet with mixed residual connections. Our method ranked the first in the challenge, outperforming other competitors by a large margin with respect to most of evaluation metrics. The proposed volumetric ConvNet is general enough and can be easily extended to other medical image analysis tasks, especially ones with limited training data.

Introduction

Prostate diseases (e.g., prostate cancer, prostatitis and enlarged prostate) are very common in men. In particular, prostate cancer is the second leading cause of cancer death in American men. It is estimated to have caused 26,120 deaths in 2016 according to American Cancer Society (Siegel, Miller, and Jemal 2016). Due to a huge increase in screening, prostate cancer is now the most commonly diagnosed cancer in men in American. Accurate segmentation of prostate from 3D Magnetic Resonance (MR) images is very useful for treatment planning and many other diagnostic and therapeutic procedures for prostate cancer as well as other prostate diseases. However, manual segmentation from 3D MR images is time-consuming and subjective with limited reproducibility. It heavily depends on experience and has large inter- and intra-observer variations. In this regard, automated segmentation methods are highly demanded in clinical practice.

Automated segmentation of prostate from MR images, however, is very challenging for several reasons (Mahapatra and Buhmann 2014). First, different MR images have global inter-scan variability and intra-scan intensity variation due to different MR scanning protocols, such as with/without endorectal coil (a thin wire placed inside the body to generate detailed MR images). Second, the lack of clear prostate boundaries due to similar appearance of prostate and surrounding tissues (e.g., blood vessels, bladder, rectum and seminal vessels) makes the automated segmentation even harder. Third, prostate has a wide variation in size and shape among different subjects due to pathological changes or different resolutions of images. Figure 1 shows examples of segmented prostate in different MR images and we can see the large variation of prostates.

Over the past few years, lots of automated prostate segmentation methods have been proposed to meet these challenges, such as atlas (registration) based methods, deformable methods and machine learning based methods. Although great progress is achieved, there still exists an ob-

Figure 1: Example of prostate MR images displaying large variations (only show the center slice of prostates in 3D MR images and the yellow contours indicate prostates).
vious gap between the automated segmentation results and manual annotations.

Recently, deep convolutional neural networks (ConvNets) with hierarchical feature learning capability have become the dominant machine learning approach in computer vision field and have achieved promising results in different vision tasks (He et al. 2015; Long, Shelhamer, and Darrell 2015; Xie and Tu 2015). Some researchers have employed ConvNets in automated prostate segmentation (Cheng et al. 2016). Nevertheless, most of these methods employed 2D ConvNets on 2D MR image slices and hence were incapable of taking full use of the 3D spatial information of the volumetric data for more accurate segmentation. Recently, many 3D ConvNets were proposed for both natural video and medical image analysis tasks (Ji et al. 2013; Tran et al. 2015; Kamnitsas et al. 2016). Nevertheless, how to train an efficient volumetric ConvNet under limited training data for medical image analysis applications is still a challenging problem.

Commonly, there are several schemes to dig the potential of the limited training data. The first is harnessing the data augmentation methods (Krizhevsky, Sutskever, and Hinton 2012). Such a scheme usually has performance gains by transforming the training data. However, the information added by these augmented data is limited and we need more efficient methods to further improve the performance. The second scheme is employing skip connections in ConvNet architecture to boost the information propagation within the ConvNet in order to achieve more efficient training with limited training data. Recently, a special skip connection, namely residual connection has been demonstrated as an effective mechanism to train very deep ConvNets and led to a series of breakthrough on some challenge datasets, such as ImageNet and MS COCO dataset (He et al. 2015). It has been demonstrated that the residual connections can improve the information flow within the networks and hence accelerate the convergence speed and improve the performance (He et al. 2016; Szegedy, Ioffe, and Vanhoucke 2016; Drozdzal et al. 2016).

In this paper, we distill the residual connections into a volumetric ConvNet and propose a novel learning architecture with mixed long and short residual connections for automated prostate segmentation from 3D MR images. Our volumetric ConvNet adopts fully convolutional architecture (Long, Shelhamer, and Darrell 2015) and can be trained end-to-end to perform efficient, precise and volume-to-volume prediction. Compared with previous works that just leverage the residual connections within local residual blocks, we extend the residual connections across residual blocks and promote the information exchange between the down-sampling path and up-sampling path in the fully convolutional architecture. By incorporating mixed long and short residual connections into our volumetric network, the information can be smoothly propagated throughout the network, which enhances the discriminative capability of networks and improves the training efficiency. Particularly, the long residual connections can recover the spatial information loss caused by down-sampling operations of the network and leverage the location information propagated from earlier layers of the network to achieve better segmentation results. We evaluated our method on the open MICCAI PROMISE12 challenge dataset; it ranked first in the challenge, outperforming other methods by a large margin.

Related Work
Prostate segmentation Previous automated prostate segmentation methods mainly include multi-atlas based methods (Klein et al. 2008) and deformable methods (Toth and Madabhushi 2012). For example, Klein et al. (2008) proposed an automated segmentation method based on atlas matching. In this method, several template images with corresponding segmentations are registered to the target image using a non-rigid registration method, and then the aligned segmentations are fused to obtain the final results. Toth et al. (2012) proposed an extension of active appearance model (AAM) to capture shape information with a multi-feature landmark-free framework. On the other hand, various graph cut based methods were also proposed to segment the prostate from MR images. Tian et al. (2016) proposed a superpixel-based 3D graph cut algorithm to obtain the prostate surface. Many successful approaches were proposed to use feature-based machine learning methods, such as K-nearest-neighbor, random forest classifier and marginal space learning (Zheng and Comaniciu 2014). Recently, with the impressive performance achieved by deep learning methods, some researchers proposed to utilize deep learning techniques to learn representation features for automated prostate segmentation. For example, Liao et al. (2013) proposed a deep learning framework with an independent subspace analysis network and Cheng et al. (2016) combined CNN and AAM methods for accurate prostate segmentation.

Deep learning for volumetric data processing In the field of medical image computing, many imaging modalities are volumetric, such as 3D Computed Tomography (CT) and MR Images. A lot of effort has been dedicated to employing CNNs to process volumetric data. Some of them employed variants of 2D CNNs to exploit adjacent slices (Chen et al. 2015), orthogonal planes (Prasoon et al. 2013; Roth et al. 2014) or multi-view planes (Setio et al. 2016) to aggregate 3D contextual feature in the model. However, these methods cannot effectively make full use of the 3D spatial information. Some studies started to employ 3D CNN to cope with detection and segmentation problems in medical volumetric data (Li et al. 2014; Brosch et al. 2016; Kamnitsas et al. 2016; Chen et al. 2016a). These networks adopted some relatively simple architectures and hence suffered from limited representation capability. By borrowing the design principles of U-Net (Ronneberger, Fischer, and Brox 2015), some researchers proposed 3D fully convolutional networks for efficient segmentation tasks (Ciçek et al. 2016; Milletari, Navab, and Ahmadi 2016; Merkow et al. 2016). Although these networks improved the segmentation performance, there was still much room to dig the potential of CNNs by effectively training the networks under limited training data. Our work is also related to Chen et al. (2016b), which is a 3D extension of residual networks for the segmentation task. But in our work, we extend the short residual
connections across residual blocks and seamlessly combine
the long and short residual connections, which greatly im-
proves the training efficiency and discriminative capability
of our network under limited training data.

Method

Residual Connections

Deep ConvNets with residual connections have achieved
promising results in many challenging natural image pro-
cessing tasks. Residual connection, in principle, is a kind of
skip connection that bypasses the non-linear transformations
with an identity mapping and explicitly reformulates the lay-
ers as learning residual functions with reference to the layer
inputs (He et al. 2015). Formally, the residual connection
can be expressed as:

\[ x_\ell = H_\ell(x_{\ell-1}) + x_{\ell-1}, \tag{1} \]

where \( x_{\ell-1} \) and \( x_\ell \) are input and output of the \( \ell^{th} \) unit,
and \( H_\ell( \cdot ) \) denotes the residual function corresponding to \( \ell^{th} \) unit.

While the residual connection is originally proposed
to address the degradation problem when training a very
deep ConvNet, a few recent studies have demonstrated that
residual connection can also promote information propa-
gation within a ConvNet both forward and backward (He
et al. 2016), and hence accelerate its convergence and im-
prove performance (Szegedy, Ioffe, and Vanhoucke 2016;
Zagoruyko and Komodakis 2016). Hence, from the
down-sampling path, we can only obtain a coarse predic-
tion, which is sufficient for some detection and classifica-
tion tasks but is unfit for voxel-wise semantic segmentation.

In addition, we also integrate a deep supervision mech-
anism (Lee et al. 2015; Dou et al. 2016) in our volumet-
ric ConvNet to accelerate its convergence speed. We exploit
additional supervision injected into some hidden layers via
auxiliary predictions. Specifically, we add one convolutional
layer (kernel size \( 1 \times 1 \times 1 \)) at the end of the network to gen-
erate the main prediction. Besides, we also employ several
convolutional layers (kernel size \( 1 \times 1 \times 1 \)) followed by hid-
den feature maps in the up-sampling path to obtain auxiliary
coe coarse predictions, and then use deconvolutional layers to
get auxiliary dense predictions with the same size of input.

Our Basic Volumetric ConvNet

In order to fully leverage the 3D spatial contextual informa-
tion of volumetric data to dig the potential of learning ca-

Our Volumetric ConvNet with Mixed Residual Connections

In order to improve the segmentation performance of our volumetric ConvNet under limited training data, we incorporate residual connections to our basic ConvNet to enhance the propagation of volumetric context information locally and globally, and by this way to achieve more performance gains from the limited training data. Specifically, we introduce two kinds of residual connections into our volumetric ConvNet. The first kind of residual connections are employed to construct the local residual blocks. It is the same connections with those reported in (He et al. 2015) and we refer them as short residual connections in this paper. The second kind of residual connections are applied to connect the residual blocks with the same resolution in the down-sampling and up-sampling paths. These connections can effectively propagate context and gradient information both forward and backward during the end-to-end training process. We refer these connections as long residual connections. Figure 2 (a) shows the architecture of our volumetric ConvNet with mixed residual connections for prostate segmentation from 3D MR images.

Our volumetric ConvNet consists of several residual blocks (ResBlocks). We design the residual blocks with short residual connections based on the design scheme of He et al. (2015). As shown in Figure 2 (b), it is composed of two convolutional layers and two rectified linear units (ReLU). Previous studies (Simonyan and Zisserman 2014; Tran et al. 2015) have demonstrated that smaller convolutional kernels are more efficient in ConvNet design. Therefore, we adopt small convolution kernels with size of 3×3×3 in convolutional layers. Each convolutional layer is followed by a ReLU as an activation function. We further employ batch normalization (BN) (Ioffe and Szegedy 2015) between each pair of convolution and ReLU, as the BN can reduce the internal covariance shift, and hence accelerate the training process and improve performance.

The down-sampling path contains one convolutional layer and three residual blocks, as shown in Figure 2 (a). Three 2×2×2 max pooling layers with stride of 2 are applied between them. As for the up-sampling path, we also employ three residual blocks. Three deconvolutional layers with stride of 2 are employed to restore the feature map size. Note that in the down-sampling path, the input and output of residual blocks have different numbers of feature maps. We add 1×1×1 convolutional layers in the short residual connections in the down-sampling path to match the dimensions of input and output (He et al. 2015).

To further boost the information exchange among different layers, we implement long residual connections to our volumetric ConvNet by connecting the corresponding residual blocks with the same resolutions in the down-sampling path and the up-sampling path using the same residual learning scheme as the short residual connections. These long residual connections are illustrated as the solid-line arrows in Figure 2 (a). These skip connections can explicitly propagate two kinds of important information within the ConvNet. First, they can propagate the spatial location information forward to the up-sampling path in order to recover the spatial information loss caused by down-sampling operations for more accurate segmentation. Second, as we employ summation operations to construct the skip connections, our architecture can more smoothly propagate the gradient flow backward, and hence improve the training efficiency and network performance (He et al. 2016; Wang et al. 2016). In addition, such connections can also implicitly promote the information exchange between feature extraction and prediction and thus offer guidance for the training of earlier layers in the ConvNet, which can be considered as a kind of intermediate supervision to enhance the training. Overall, our volumetric ConvNet with mixed residual connections can be trained end-to-end and achieve efficient volume-to-volume prediction by improving the information propagation within the ConvNet.

Experiments and Results

Dataset and Pre-processing

We performed extensive experiments to evaluate our method on MICCAI Prostate MR Image Segmentation (PROMISE12) challenge dataset (Litjens et al. 2014), an ongoing benchmark for evaluating algorithms for segmentation of the prostate from MR images. The training dataset contains 50 transversal T2-weighted MR images of the prostate and corresponding segmentation ground truth. The testing dataset consists of 30 MR images and the ground truth is held out by the organizer for independent evaluation. These images are acquired in different hospitals, using different equipments and different acquisition protocols, and display the maximum variations of MR images acquired in clinical setting: there are variations in voxel size, dynamic range, position, field of view and anatomic appearance. Different from previous works using complex pre-processing steps, like N4 bias field correction, we simply resized all MR volumes into a fixed resolution of 0.625×0.625×1.5 mm and then normalized them as zero mean and unit variance. We also utilized simple data augmentation strategy to enlarge the training dataset. The augmentation operations included rotation (90, 180 and 270 degrees) and flip in axial plane.

Implementation

Our volumetric ConvNet was implemented based on a modified Caffe library (Jia et al. 2014) supporting 3D operations. All the trainings and experiments were conducted on a workstation equipped with a NVIDIA TITAN X GPU. The networks were trained with Stochastic Gradient Descent (SGD) method with a mini-batch size of 8 due to the limited capacity of GPU memory. The learning rate was set as 0.001 initially and is divided by 10 every 3000 iterations; the models were trained for up to 10000 iterations. We employed a weight decay of 0.0005 and a momentum of 0.9. We utilized two auxiliary predictions in deep supervision scheme and the balancing weights were 0.3 and 0.6, respectively. Due to the limited GPU memory, we randomly cropped 64×64×16 sub-volumes from every sample as input when training the network. In the test phase, we used overlapped sliding windows strategy to crop sub-volumes and then used the av-
average of the probability maps of these sub-volumes to get
the whole volume prediction. The sub-volume size was also
64×64×16 and the stride was 50×50×12. Generally, it took
about 4 hours to train the network and about 12 seconds for
processing one MR images with size of 320×320×60.

Ablation Analysis of Residual Connections

In order to evaluate the effectiveness of the residual connec-
tions in our volumetric ConvNet, we performed a set of abla-
tion experiments on the ConvNet. Because the ground truth
of testing data is held out by organizers and the challenge
organizers only allow resubmission of substantially differ-
ent methods, we conducted experiments via standard 10-fold
cross validation scheme.

We first analyze the learning behaviors of our volumet-
ric ConvNet with and without mixed residual connections.
Figure 3 presents the training and validation losses of differ-
ent networks with and without mixed residual connections.
It is observed that the ConvNet with mixed residual con-
nections converges faster and achieves lower validation loss
than the one without mixed residual connections, demon-
strating that residual connections can improve the training
efficiency of the volumetric ConvNet. Table 1 further shows
the performance of our volumetric networks with different
residual connections via cross validation. It is observed that
adding either long or short residual connections can achieve
better Dice performance than networks without residual con-
nections, demonstrating residual connections, as a general
skip connection design strategy, can improve the discrimina-
tive capability of networks by promoting information prop-
agation. The network with only short residual connections
has marginally better performance than that with only long
residual connections. The network with mixed residual connec-
tions achieves the best performance in the ablation ex-
periments, indicating that long and short residual connec-
tions can provide different information propagation locally
and globally in the volumetric ConvNet and combining them
together can further improve the performance.

Comparison with Other Methods

The evaluation metrics used in the PROMISE12 challenge
include the Dice coefficient (DSC), the percentage of the ab-
solute difference between the volumes (aRVD), the average
over the shortest distance between the boundary points of the
volumes (ABD) and the 95% Hausdorff distance (95HD). Note
that all evaluation metrics including both boundary and
volume metrics were calculated in 3D not only for the en-
tire prostate segmentation but also specially for the apex and
base parts of the prostate, which are the most difficult yet
important parts in clinical practice (Litjens et al. 2014). The
organizers calculated a total score incorporating above com-
pletely different but equally important metrics to rank the
submitted methods. Readers can refer to Litjens et al. (2014)
for more evaluation details.

Some qualitative results of our method are shown in Fig-
ure 4. It is observed that our method can produce accu-
rate segmentation results and delineate the clear contours of
prostates in MR images with (case 22) and without (case 4)
endorectal coil. The quantitative results of our method and
our competitors are shown in Table 2. There were totally 21
teams submitting their results until the paper submission
and only top 10 teams are listed in the Table. Note that all the
results reported in this section were obtained directly from

Table 1: Cross validation performance of our volumetric
ConvNets with different configurations.

<table>
<thead>
<tr>
<th>Method</th>
<th>Dice coefficient [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Long-short connections</td>
<td>86.93</td>
</tr>
<tr>
<td>Only short connections</td>
<td>84.68</td>
</tr>
<tr>
<td>Only long connections</td>
<td>84.38</td>
</tr>
</tbody>
</table>
| Without residual connec-
  tions                  | 81.63                |

Figure 3: Training and validation loss of networks with and
without mixed residual connections.

Figure 4: Qualitative segmentation results of case 4 (first
row) and case 22 (second row) at the apex(left), center (mid-
dle) and base (right) of the prostate in testing dataset. The
yellow and red contours indicate the ground truth and our
segmentation results. Note that these results are directly ob-
tained from challenge website.

Figure 1: Qualitative segmentation results of case 4 (first
row) and case 22 (second row) at the apex(left), center (mid-
dle) and base (right) of the prostate in testing dataset. The
yellow and red contours indicate the ground truth and our
segmentation results. Note that these results are directly ob-
tained from challenge website.
<table>
<thead>
<tr>
<th>Method</th>
<th>Type</th>
<th>Whole</th>
<th>Base</th>
<th>Apex</th>
</tr>
</thead>
<tbody>
<tr>
<td>CUMED (ours)</td>
<td>Auto</td>
<td>1.95</td>
<td>5.34</td>
<td>4.29</td>
</tr>
<tr>
<td>Inmorphics</td>
<td>Auto</td>
<td>2.10</td>
<td>5.45</td>
<td>4.73</td>
</tr>
<tr>
<td>Emory</td>
<td>Semi</td>
<td>2.14</td>
<td>5.04</td>
<td>5.31</td>
</tr>
<tr>
<td>ScrAutoProstate</td>
<td>Auto</td>
<td>2.13</td>
<td>5.58</td>
<td>4.93</td>
</tr>
<tr>
<td>CAMP-TUM2</td>
<td>Auto</td>
<td>2.23</td>
<td>5.71</td>
<td>4.62</td>
</tr>
<tr>
<td>ETHZ</td>
<td>Semi</td>
<td>2.25</td>
<td>5.95</td>
<td>5.99</td>
</tr>
<tr>
<td>SIATMIDS</td>
<td>Auto</td>
<td>2.49</td>
<td>6.17</td>
<td>6.09</td>
</tr>
<tr>
<td>CBA</td>
<td>Interactive</td>
<td>2.33</td>
<td>6.57</td>
<td>5.75</td>
</tr>
<tr>
<td>CAMP-TUM</td>
<td>Auto</td>
<td>2.48</td>
<td>7.17</td>
<td>6.31</td>
</tr>
<tr>
<td>SBIA</td>
<td>Auto</td>
<td>2.85</td>
<td>7.33</td>
<td>6.99</td>
</tr>
</tbody>
</table>

For ABD, 95HD and aRVD, lower values are better; for DSC, higher values are better.

<table>
<thead>
<tr>
<th>Method</th>
<th>Type</th>
<th>Whole</th>
<th>Base</th>
<th>Apex</th>
</tr>
</thead>
<tbody>
<tr>
<td>CUMED (ours)</td>
<td>Auto</td>
<td>1.95</td>
<td>5.34</td>
<td>4.29</td>
</tr>
<tr>
<td>Inmorphics</td>
<td>Auto</td>
<td>2.10</td>
<td>5.45</td>
<td>4.73</td>
</tr>
<tr>
<td>Emory</td>
<td>Semi</td>
<td>2.14</td>
<td>5.04</td>
<td>5.31</td>
</tr>
<tr>
<td>ScrAutoProstate</td>
<td>Auto</td>
<td>2.13</td>
<td>5.58</td>
<td>4.93</td>
</tr>
<tr>
<td>CAMP-TUM2</td>
<td>Auto</td>
<td>2.23</td>
<td>5.71</td>
<td>4.62</td>
</tr>
<tr>
<td>ETHZ</td>
<td>Semi</td>
<td>2.25</td>
<td>5.95</td>
<td>5.99</td>
</tr>
<tr>
<td>SIATMIDS</td>
<td>Auto</td>
<td>2.49</td>
<td>6.17</td>
<td>6.09</td>
</tr>
<tr>
<td>CBA</td>
<td>Interactive</td>
<td>2.33</td>
<td>6.57</td>
<td>5.75</td>
</tr>
<tr>
<td>CAMP-TUM</td>
<td>Auto</td>
<td>2.48</td>
<td>7.17</td>
<td>6.31</td>
</tr>
<tr>
<td>SBIA</td>
<td>Auto</td>
<td>2.85</td>
<td>7.33</td>
<td>6.99</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Method</th>
<th>Type</th>
<th>Whole</th>
<th>Base</th>
<th>Apex</th>
</tr>
</thead>
<tbody>
<tr>
<td>CUMED (ours)</td>
<td>Auto</td>
<td>1.95</td>
<td>5.34</td>
<td>4.29</td>
</tr>
<tr>
<td>Inmorphics</td>
<td>Auto</td>
<td>2.10</td>
<td>5.45</td>
<td>4.73</td>
</tr>
<tr>
<td>Emory</td>
<td>Semi</td>
<td>2.14</td>
<td>5.04</td>
<td>5.31</td>
</tr>
<tr>
<td>ScrAutoProstate</td>
<td>Auto</td>
<td>2.13</td>
<td>5.58</td>
<td>4.93</td>
</tr>
<tr>
<td>CAMP-TUM2</td>
<td>Auto</td>
<td>2.23</td>
<td>5.71</td>
<td>4.62</td>
</tr>
<tr>
<td>ETHZ</td>
<td>Semi</td>
<td>2.25</td>
<td>5.95</td>
<td>5.99</td>
</tr>
<tr>
<td>SIATMIDS</td>
<td>Auto</td>
<td>2.49</td>
<td>6.17</td>
<td>6.09</td>
</tr>
<tr>
<td>CBA</td>
<td>Interactive</td>
<td>2.33</td>
<td>6.57</td>
<td>5.75</td>
</tr>
<tr>
<td>CAMP-TUM</td>
<td>Auto</td>
<td>2.48</td>
<td>7.17</td>
<td>6.31</td>
</tr>
<tr>
<td>SBIA</td>
<td>Auto</td>
<td>2.85</td>
<td>7.33</td>
<td>6.99</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Method</th>
<th>Type</th>
<th>Whole</th>
<th>Base</th>
<th>Apex</th>
</tr>
</thead>
<tbody>
<tr>
<td>CUMED (ours)</td>
<td>Auto</td>
<td>1.95</td>
<td>5.34</td>
<td>4.29</td>
</tr>
<tr>
<td>Inmorphics</td>
<td>Auto</td>
<td>2.10</td>
<td>5.45</td>
<td>4.73</td>
</tr>
<tr>
<td>Emory</td>
<td>Semi</td>
<td>2.14</td>
<td>5.04</td>
<td>5.31</td>
</tr>
<tr>
<td>ScrAutoProstate</td>
<td>Auto</td>
<td>2.13</td>
<td>5.58</td>
<td>4.93</td>
</tr>
<tr>
<td>CAMP-TUM2</td>
<td>Auto</td>
<td>2.23</td>
<td>5.71</td>
<td>4.62</td>
</tr>
<tr>
<td>ETHZ</td>
<td>Semi</td>
<td>2.25</td>
<td>5.95</td>
<td>5.99</td>
</tr>
<tr>
<td>SIATMIDS</td>
<td>Auto</td>
<td>2.49</td>
<td>6.17</td>
<td>6.09</td>
</tr>
<tr>
<td>CBA</td>
<td>Interactive</td>
<td>2.33</td>
<td>6.57</td>
<td>5.75</td>
</tr>
<tr>
<td>CAMP-TUM</td>
<td>Auto</td>
<td>2.48</td>
<td>7.17</td>
<td>6.31</td>
</tr>
<tr>
<td>SBIA</td>
<td>Auto</td>
<td>2.85</td>
<td>7.33</td>
<td>6.99</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Method</th>
<th>Type</th>
<th>Whole</th>
<th>Base</th>
<th>Apex</th>
</tr>
</thead>
<tbody>
<tr>
<td>CUMED (ours)</td>
<td>Auto</td>
<td>1.95</td>
<td>5.34</td>
<td>4.29</td>
</tr>
<tr>
<td>Inmorphics</td>
<td>Auto</td>
<td>2.10</td>
<td>5.45</td>
<td>4.73</td>
</tr>
<tr>
<td>Emory</td>
<td>Semi</td>
<td>2.14</td>
<td>5.04</td>
<td>5.31</td>
</tr>
<tr>
<td>ScrAutoProstate</td>
<td>Auto</td>
<td>2.13</td>
<td>5.58</td>
<td>4.93</td>
</tr>
<tr>
<td>CAMP-TUM2</td>
<td>Auto</td>
<td>2.23</td>
<td>5.71</td>
<td>4.62</td>
</tr>
<tr>
<td>ETHZ</td>
<td>Semi</td>
<td>2.25</td>
<td>5.95</td>
<td>5.99</td>
</tr>
<tr>
<td>SIATMIDS</td>
<td>Auto</td>
<td>2.49</td>
<td>6.17</td>
<td>6.09</td>
</tr>
<tr>
<td>CBA</td>
<td>Interactive</td>
<td>2.33</td>
<td>6.57</td>
<td>5.75</td>
</tr>
<tr>
<td>CAMP-TUM</td>
<td>Auto</td>
<td>2.48</td>
<td>7.17</td>
<td>6.31</td>
</tr>
<tr>
<td>SBIA</td>
<td>Auto</td>
<td>2.85</td>
<td>7.33</td>
<td>6.99</td>
</tr>
</tbody>
</table>

Acknowledgments

The work is supported by the grants from the National Basic Program of China, 973 Program (Project No. 2015CB351706) and the Research Grants Council of the Hong Kong Special Administrative Region (Project no. CUHK 14202514 and CUHK 14203115).

References


Drozdzal, M.; Vorontsov, E.; Chartrand, G.; Kadoury, S.; and Pal, C. 2016. The importance of skip connec-