WILEY | Hindawi

Research Article
Exploiting Interslice Correlation for MRI Prostate Image Segmentation, from Recursive Neural Networks Aspect

Qikui Zhu,[1] **Bo Du** [iD]**,**[1] **Baris Turkbey,**[2] **Peter Choyke,**[2] **and Pingkun Yan**[3]

[1]*School of Computer, Wuhan University, Wuhan, China*
[2]*National Cancer Institute, National Institutes of Health, Bethesda, MD, USA*
[3]*Biomedical Engineering Department, Rensselaer Polytechnic Institute, Troy, NY 12180, USA*

Correspondence should be addressed to Bo Du; remoteking@whu.edu.cn

Segmentation of the prostate from Magnetic Resonance Imaging (MRI) plays an important role in prostate cancer diagnosis. However, the lack of clear boundary and significant variation of prostate shapes and appearances make the automatic segmentation very challenging. In the past several years, approaches based on deep learning technology have made significant progress on prostate segmentation. However, those approaches mainly paid attention to features and contexts within each single slice of a 3D volume. As a result, this kind of approaches faces many difficulties when segmenting the base and apex of the prostate due to the limited slice boundary information. To tackle this problem, in this paper, we propose a deep neural network with bidirectional convolutional recurrent layers for MRI prostate image segmentation. In addition to utilizing the intraslice contexts and features, the proposed model also treats prostate slices as a data sequence and utilizes the interslice contexts to assist segmentation. The experimental results show that the proposed approach achieved significant segmentation improvement compared to other reported methods.

1. Introduction

Accurately segmenting the prostate from Magnetic Resonance Imaging (MRI) provides very useful information for clinical applications like computer aided diagnosis and image guided interventions [1]. However, it is a very challenging task due to the lack of clear boundary definition and the significant variation of shape and texture across images from different patients [2] as shown in Figure 1.

According to the guidance information used in the model, we can classify the existing prostate MRI segmentation methods into four kinds: region based, shape prior based, contour based, and classification methods [3–5]. Toth et al. [6] presented an Active Shape Model (ASM) initialization scheme for prostate segmentation which leverages multimodal information to initialize ASM. Samiee et al. [7] proposed a model using shape prior of prostate to refine the prostate boundary; Klein [8] proposed an automatic segmentation model which is based on manually matching segmented atlas images.

In the past several years, deep learning based techniques, especially fully convolutional neural networks (FCNs), have proved very effective on image segmentation [9–11], including biomedical image segmentation. Zhu et al. [12] proposed a deeply supervised CNN that utilized the residual information to accurately segment the prostate MRI. Bao and Chung [13] introduced a multiscale structured FCN model for brain MRI segmentation by capturing discriminative features from input patch. Other examples on introducing deep learning into biomedical image segmentation can be found in [14–17].

However, a straightforward extension of those 2D image segmentation methods to 3D may not yield satisfactory performance, due to the anisotropic nature of many medical imaging modalities. To tackle this problem, Chen et al. [18] proposed a method to combine fully convolutional neural networks with the extended Convolutional Long Short-Term Memory (C-LSTM), which improved 3D medical image segmentation performances by simultaneously leveraging the abstraction capabilities of both FCNs and RNNs. Nevertheless, their model relies on the so-called U-Net [18] to extract
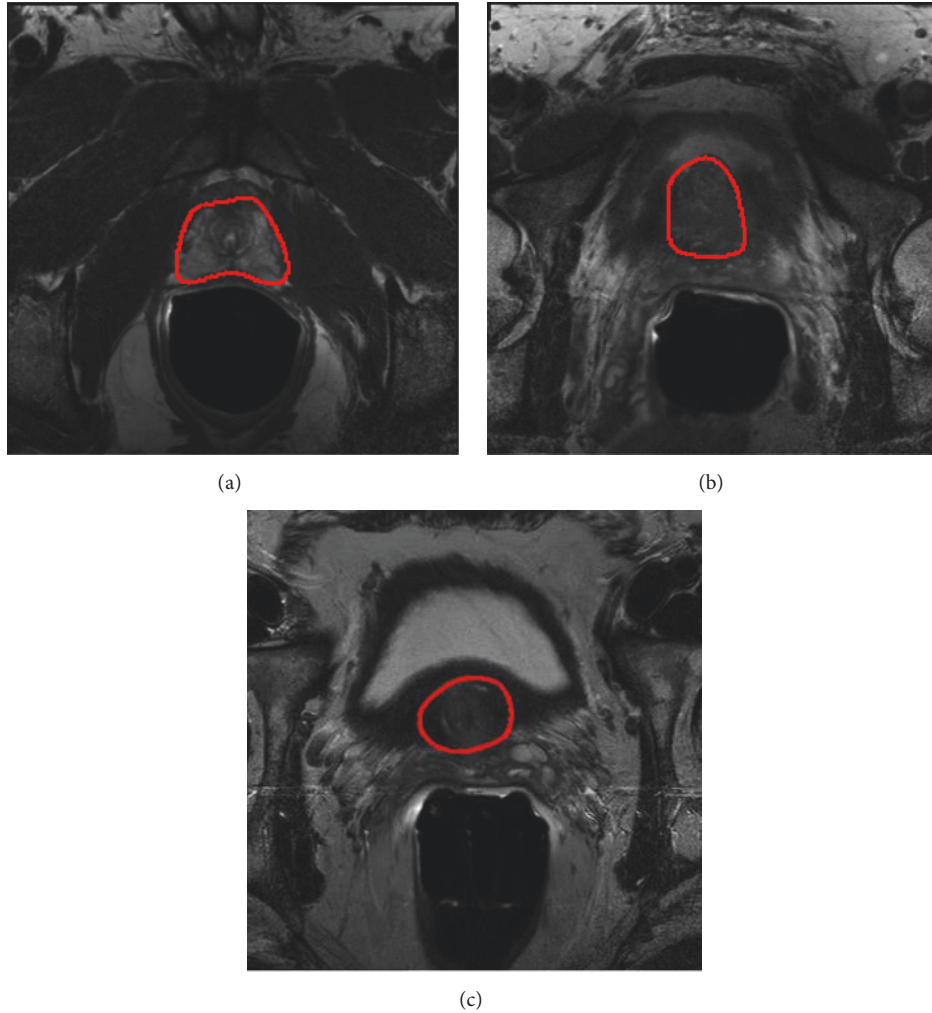
(a)

(b)

(c)

FIGURE 1: Challenges in segmenting prostate in MR images. (a) Imaging artifacts inside the bladder. (b) Weak boundary. (c) Complex intensity distribution.

image features. The following Bidirectional Convolutional LSTMs (BDC-LSTMs) only work on the extracted features. Thus, useful information for assisting image segmentation may be lost.

As to prostate segmentation task, an insurmountable challenge is the image slice specifically at the apex and base always loses clear boundaries and necessary information. This phenomenon brings the main difficulty to prostate segmentation. However, shape prior is an effective way to resolve this challenge. For instance, Qin et al. [19] proposed an adaptive shape prior constrained directional level set model (ASPDLS) to segment the inner and outer boundaries of the bladder wall and achieved accurate segmentation results. Motivated by the fact that the acquired MRI images typically have a high intraslice resolution and there exists a high spatial dependence between slices from the same patient, we utilize interslice as shape prior to guide the process of feature extraction and explore necessary information from interslice to alleviate information loss as shown in Figure 2. Besides the architecture of RNNs has superiority performances in modeling sequential data [17, 20, 21]. To improve the performance of prostate segmentation, in this paper, we propose
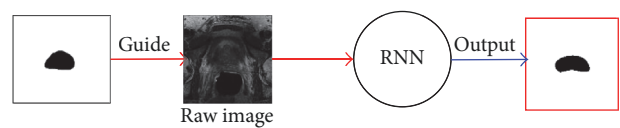

FIGURE 2: Interslice as shape prior to guide the process of feature extraction.

a network, called UR-Net, which treats prostate slices as a data sequence, utilizing the intraslice contexts and features to assist segmentation.

There are two main contributions of this paper. First, we treat prostate slices as a data sequence and utilize interslice as shape prior to guide the process of feature extraction and explore necessary information from interslice. Second, we explore the power of RNNs rather than the traditional CNNs to extract image feature. The experimental results demonstrate that the use of RNNs can substantially improve the performance of prostate segmentation.

The rest of the paper is organized as follows. The architecture of Recurrent Neural Network and the details of proposed network architecture are described in Section 2.
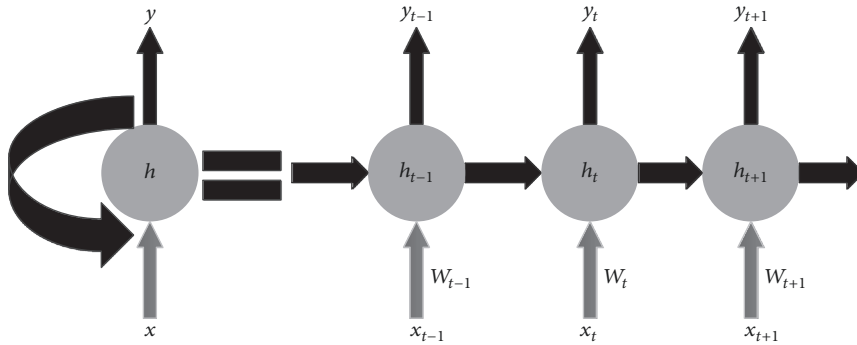
FIGURE 3: Architecture of the Recurrent Neural Networks and unfolding Recurrent Neural Networks into a full network. x_t is the input at time step t. h_t is the hidden state at time step t, which is the "memory" of the network. y_t is the output at time step t.

Section 3 presents the experimental results and performance evaluation. The conclusions are provided in Section 4.

2. UR-Net

In this section, we first review the classic Recurrent Neural Networks (RNNs) and then move on to describe the extension of RNNs to Long Short-Term Memory (LSTM) [22] and Convolutional Long Short-Term Memory (CLSTM) [18] which are specific Recurrent Neural Networks. After that, the proposed UR-Net is presented in detail.

2.1. Recurrent Neural Networks (RNNs). Recurrent Neural Network (RNN) has a long history in the artificial neural network community which was designed to model temporal sequences. The architecture of typical RNNs is shown in Figure 3. This model has shown great promise in many tasks, such as NLP [23], non-Markovian control, and text tasks [24]. The idea behind RNNs is to make use of sequential information with the output being dependent on the previous computation. RNNs have a memory, which can remember the information about what has been calculated so far. In theory [25], RNNs can remember the information in arbitrarily long sequences and make use of the previous computations, but in practice they are limited to looking back only a few steps, because of the problem of vanishing gradient.

At each time step t, the RNNs utilize the input data x_t and the previous hidden state h_{t-1} to calculate the next hidden state h_t and output y_t by applying the following recursive operation:

$$h_t = f_h \left(W_h x_t + U h_{t-1} + b_h \right),$$
$$y_t = f_o \left(W_o h_t + b_o \right), \tag{1}$$

where f is an element-wise nonlinearity function; $W_h, U,$ and b_h are the parameters of hidden state; W_o and b_o are output parameters.

During the last decade, several methods have been explored for training RNNs, such as backpropagation through time (BPTT) [26], real-time recurrent learning (RTRL) [27], and extended Kalman filtering based techniques (EKF) [28]. Though those training methods can help us train RNNs, they suffer from the vanishing gradient problem.
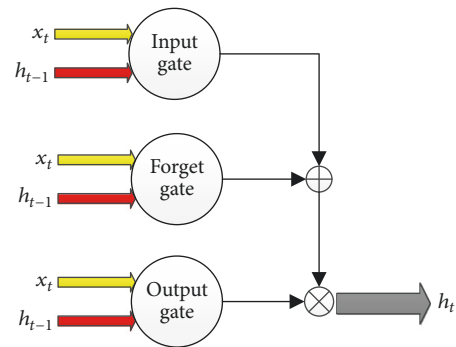

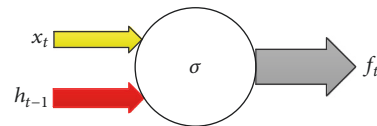FIGURE 4: Architecture of the Long Short-Term Memory.

FIGURE 5: Architecture of the forget gate in LSTM.

2.2. Long Short-Term Memory (LSTM). To address the problem of vanishing gradient and long-term dependency residing in RNNs [29], a special implementation of RNN, Long Short-Term Memory (LSTM), was introduced by Hochreiter and Schmidhuber [22]. The architecture of LSTM is shown in Figure 4.

One LSTM unit consists of an input gate (i), a forget gate (f), an output gate (o), and a memory cell (c) which possess the ability of remembering or forgetting the information over potentially long periods of time. The input gate puts and controls the input data into the memory cell. The forgetting gate decides what information we are going to throw away form the memory cell. The output gate decides which parts of data in the memory cell are going to output and simultaneously controls the output data flow into the rest of the network.

The step of LSTM can be described as follows: (1) the first step in the LSTM model corresponds to (2) which is to decide what information should be discarded from input data. This decision is made by forget gate; the architecture of forget gate is shown in Figure 5. The forget gate possesses a forgetting
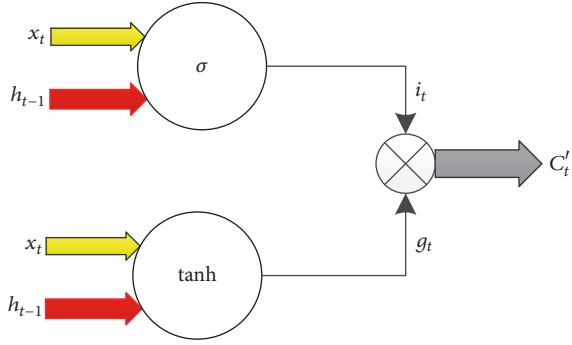
FIGURE 6: Architecture of the input gate in LSTM.

FIGURE 7: Architecture of the output gate in LSTM.

layer which consists of a sigmoid function. When getting the data from previous stage h_{t-1} and input data x_t, the sigmoid function outputs a number 0 or 1 for each data in the cell state C_{t-1}, the number 0 represents completely throwing it away, while 1 represents storing it.

$$f_t = \sigma\left(W^f x_t + U^f h_{t-1} + b^f\right). \tag{2}$$

(2) The second step is to decide what new information we should store in the cell. The input gate consists of a sigmoid function and a tanh function as shown in Figure 6. When the input gate receives a new data, the sigmoid function decides what values will be updated and the tanh function creates a candidate values g_t, those operations corresponding to (3). At last, the input gate controls the candidate values to update the cell state.

$$\begin{aligned} i_t &= \sigma\left(W^i x_t + U^i h_{t-1} + b^i\right), \\ g_t &= \tanh\left(W^g x_t + U^g h_{t-1} + b^g\right). \end{aligned} \tag{3}$$

(3) When we got the information coming from forget gate and input gate, we can utilize (4) to update the cell state. This operation can drop some useless information.

$$c_t = f_t \circ c_{t-1} + i_t \circ g_t. \tag{4}$$

(4) The fourth step is to decide what information we should output which is based on cell state. The final result consists of two parts, the first part comes from cell state selected by a tanh function. The second part comes from input data which will be selected by a sigmoid function. The output gate is shown in Figure 7 and the computing methods are

$$\begin{aligned} o_t &= \sigma\left(W^o x_t + U^o h_{t-1} + b^o\right), \\ h_t &= o_t \circ \tanh\left(c_t\right). \end{aligned} \tag{5}$$
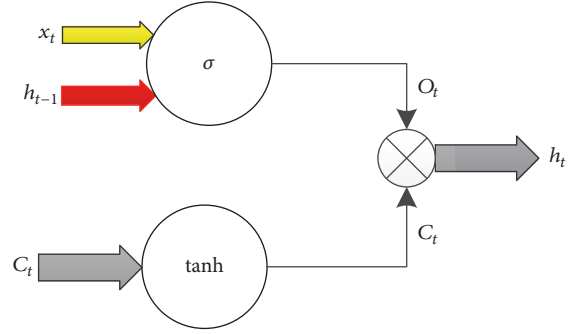
Putting those together, the gates at discrete time t ($t = 1, 2, \ldots$) are computed as follows:

$$\begin{aligned} i_t &= \sigma\left(W^i x_t + U^i h_{t-1} + b^i\right), \\ f_t &= \sigma\left(W^f x_t + U^f h_{t-1} + b^f\right), \\ o_t &= \sigma\left(W^o x_t + U^o h_{t-1} + b^o\right), \\ g_t &= \tanh\left(W^g x_t + U^g h_{t-1} + b^g\right), \\ c_t &= f_t \circ c_{t-1} + i_t \circ g_t, \\ h_t &= o_t \circ \tanh\left(c_t\right). \end{aligned} \tag{6}$$

The standard LSTM architecture is just designed for one-dimensional data. It cannot be directly applied to 2D image data. To apply LSTM in prostate image, in our model, we apply Convolutional LSTM (CLSTM) as convolutional layer. This can be achieved by using a convolution operator to replace the matrix multiplication. The core equations of CLSTM are presented in

$$\begin{aligned} i_t^d &= \sigma\left(W_i^d * x_t^d + U_i^d h_{t-1}^d + b_i^d\right), \\ f_t^d &= \sigma\left(W_f^d * x_t^d + U_f^d h_{t-1}^d + b_f^d\right), \\ o_t^d &= \sigma\left(W_o^d * x_t^d + U_o^d h_{t-1}^d + b_o^d\right), \\ g_t^d &= \tanh\left(W_g^d * x_t^d + U_g^d h_{t-1}^d + b_g^d\right), \\ c_t^d &= f_t^d \circ c_{t-1}^d + i_t^d \circ g_t^d, \\ h_t^d &= o_t \circ \tanh\left(c_t\right), \end{aligned} \tag{7}$$

where ($*$) denotes convolution operator and h is the output of the layer; d denotes the fact that the CLSTM works slice by slice in certain direction.

2.3. The Proposed Network Architecture. In order to exploit the interslice information effectively, we introduce a Bidirectional Convolutional LSTM (BDC-LSTM) layer into our deep learning network. A BDC-LSTM layer consists of two sets of CLSTMs to extract features as shown in Figure 8. The two CLSTM streams work in two opposite directions. Rather than
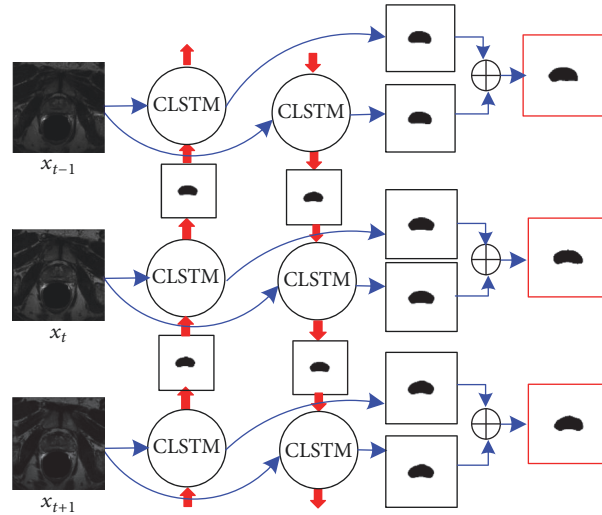
FIGURE 8: The architecture of BDC-LSTM. One BDC-LSTM layer contains two CLSTM layers working in two opposite directions. One from x_{t+1} to x_{t-1} and the other from x_{t-1} to x_{t+1}. The results of two CLSTM layers are connected together as the input of next layer.
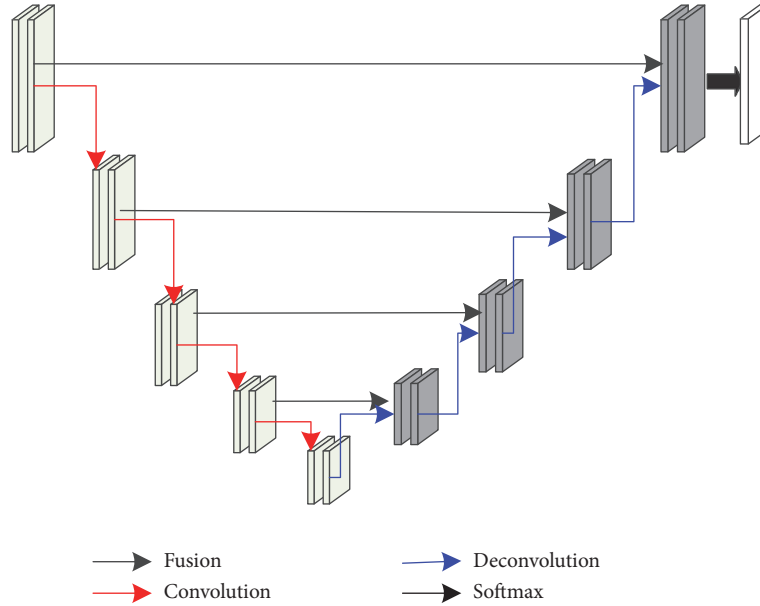

FIGURE 9: Architecture of U-Net.

serializing each prostate image into sequential patches and then leveraging Bidirectional LSTM to segment each patch, in our method, we treat each image as a whole and three adjacent image slices compose a sequence.

When we put one image sequence denoted by x_{t-1}, x_t, x_{t+1} into the BDC-LSTM layer as shown in Figure 8, the layer will utilize interslice and intraslice information to extract prostate features. Firstly, the layer extracts the first slice's x_{t-1} features. And then the result of x_{t-1} and x_t will be treated as a shape prior combined with the later slice x_t and x_{t+1} as input to guide the process of segmentation in turn. Simultaneously the layer will extract features of each slice in opposite directions from x_{t+1} to x_{t-1}, at last, connecting the two different features maps together as the input of next layer.

Our proposed network architecture is shown in Figure 10. The main framework of our proposed method follows the architecture of U-Net [30], since U-Net can successfully extract image features for segmentation with a reasonable network depth. As a matter of fact, U-Net has obtained state-of-the-art performances in many biomedical image processing tasks. For example, Milletari et al. [31] proposed a fully convolutional neural network for volumetric medical image segmentation, called V-Net. This model leverages the power of U-Net to process MRI volumes. The architecture of U-Net is shown in Figure 9.

The proposed network architecture consists of a contracting path on the left, an expansive path on the right, and a classified path on the bottom. Both the contracting path and
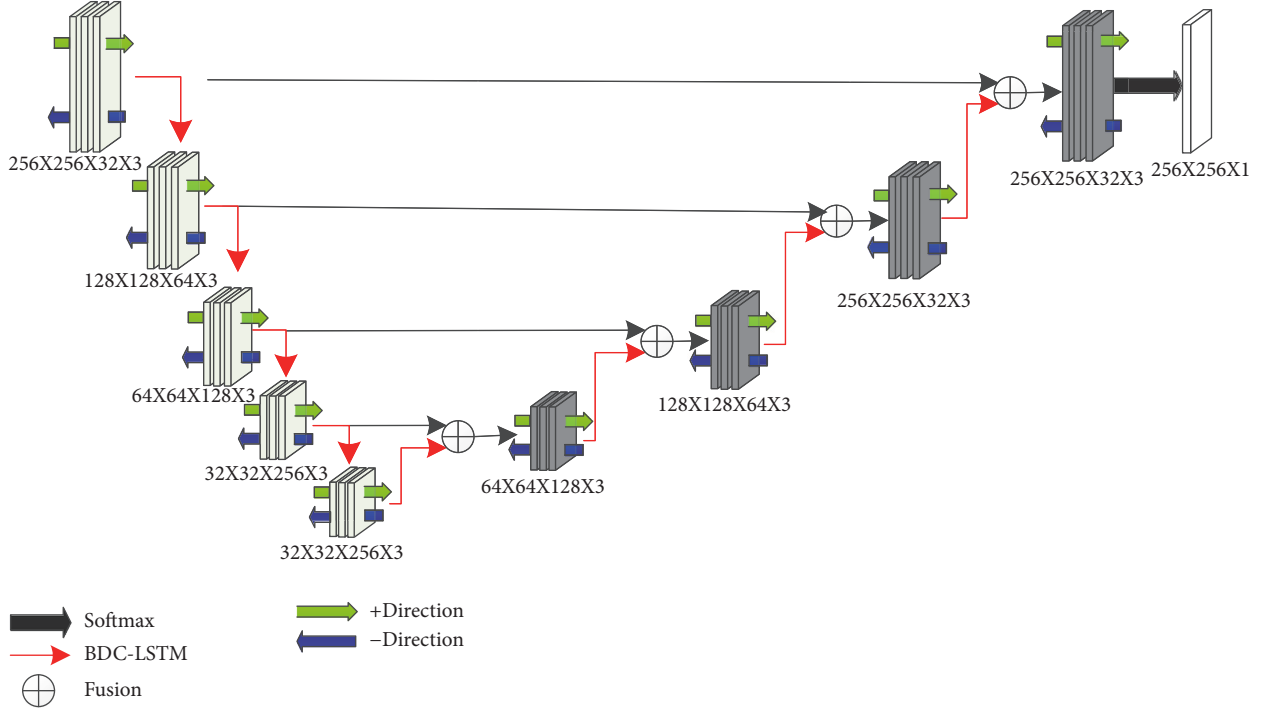
FIGURE 10: Architecture of the proposed network.

expansive path possess 4 stages, and each stage consists of one BDC-LSTM layer. At the end of networks one softmax layer is added. At the contracting stage, a 2×2 max pooling operation with a stride of 2 is attached for downsampling and the number of feature channels is doubled after each stage. On the opposite side, at the expansive stage, the first step is upsampling, which makes the width and height of feature maps doubled each time until they reach the size of the original images. At the same time, upsampling also halves the number of feature channels. To reduce information loss during convolution, a concatenation from left contracting path to right expansive path is made. The concatenation can provide features extracted from early stage to late stage and also can speed up the convergence of the network. To avoid overfitting, dropout operations have been added at the end of each stage.

2.4. Network Objective Function. For the prostate images, the anatomy of interest usually occupies a very small part of an image. This brings the problem that the networks always ignore the segmentation parts and become biased towards the background. This always led to the learning process trapped in local minima. To overcome this problem, we apply the dice coefficient as the objective function. The dice coefficient function can pay more attention to segmentation parts influences. The dice coefficient (DSC) [19] function between two images can be written as

$$\text{DSC}\left(S_a, S_m\right) = \frac{2\left|S_a \cap S_m\right|}{\left|S_a\right| + \left|S_m\right|}, \tag{8}$$

where S_a denotes the result of automatic segmentation and S_m denotes the result of manual segmentation.

In our work, the ground truth and results of segmentation are binary images, so the dice coefficient DSC between two binary images can be written as

$$\text{DSC} = \frac{2 \sum_i^N p_i q_i}{\sum_i^N p_i^2 + \sum_i^N q_i^2}, \tag{9}$$

N denotes the total number of pixels in the image, and p_i, q_i denote the pixels from ground truth and segmentation, respectively.

This formulation of dice can be differentiated yielding the gradient:

$$\frac{\partial D}{\partial p_j} = 2 \left[\frac{q_i \left(\sum_i^N p_i^2 + \sum_i^N q_i^2\right) - 2 p_j \sum_i^N p_i q_i}{\left(\sum_i^N p_i^2 + \sum_i^N q_i^2\right)^2} \right]. \tag{10}$$

Besides, Milletari et al. [31] have proved that the DSC are much better than the same network trained with a logistic loss for overcoming the network traps in local minima.

3. Experimental Results

3.1. Materials. The MRI prostate images used in our work as shown in Figure 11 were acquired from 80 patients using a Philips 3T MRI scanner with endorectal coil. The in-plane resolution is $0.3 \, \text{mm} \times 0.3 \, \text{mm}$ and interslice distance is 3 mm. Each patient image volume consists of about 26 slices. The dimension of each 2D slice is 512×512 pixels.

3.2. Training Strategy. We randomly selected 76 patients from 80 patients for training and the rest of patients are utilized
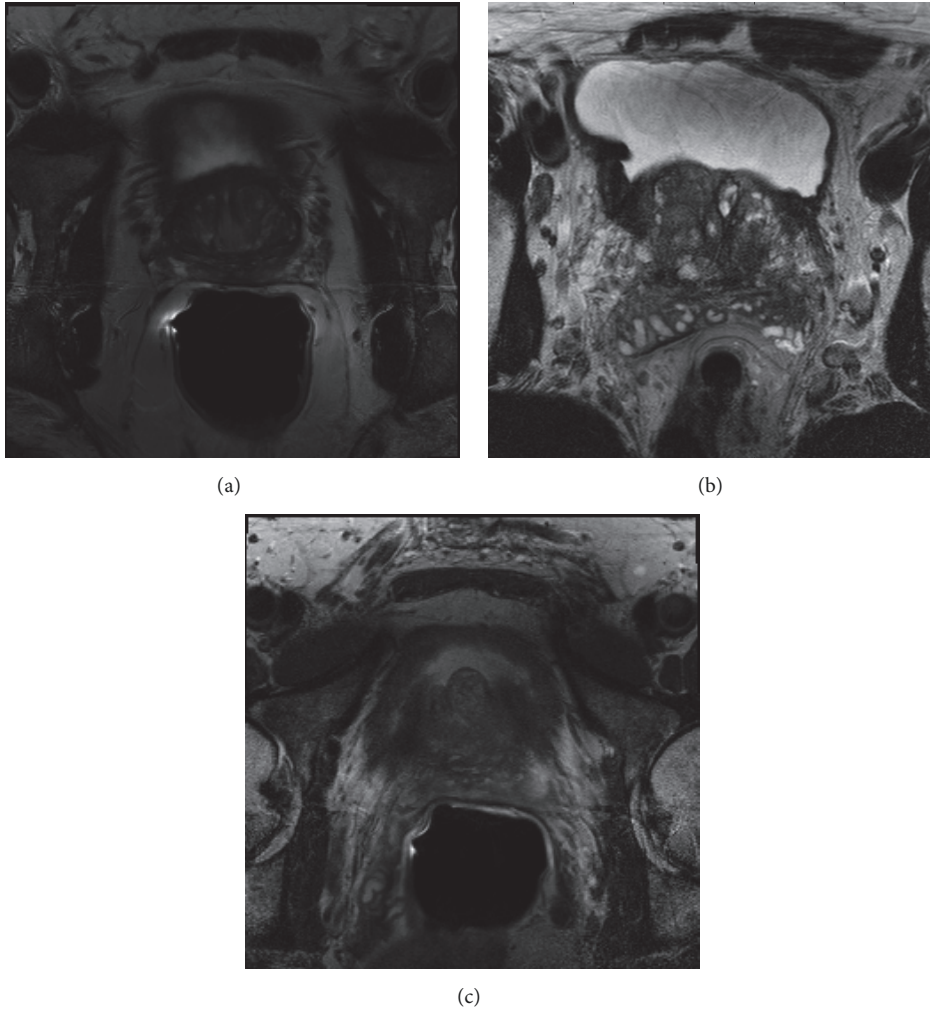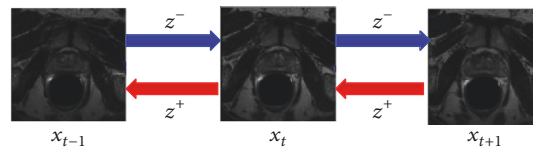
(a)

(b)

(c)

FIGURE 11: Parts of MRI prostate images.

x_{t-1} x_t x_{t+1}

FIGURE 12: The z^- direction and z^+ direction.

for testing. During training, we put three sequential slices denoted by x_{t-1}, x_t, x_{t+1} from one patient into the network. And then the BDC-LSTM layers exploit intraslice and interslice contextual information from two directions, one in z^- direction and the other in z^+ direction as shown in Figure 12. Our network is trained end-to-end on the prostate scans dataset. And the network framework is implemented under the open-source deep learning library Keras [32]. Experiments are carried out on GTX1080 GPU with 8 GB of video memory and the CUDA edition is 8.0. In the training phase, the learning rate is set as 0.0001 initially. Due to the limit by the memory, we choose 1 as the mini-batch. And all of the train image and ground truth have been resized to 256 × 256.

3.3. Experiments. To validate whether the deep neural network with RNN layers can significantly improve the segmentation accuracy, we also modify the FCNs by utilizing BDC-LSTM layers to replace the convolutional layers within FCNs. These testing images come with a corresponding ground truth segmentation map which is a binary image and is used to evaluate the performances of automatic segmentation. At last, we compare our model with U-Net, V-Net [31], fully convolutional networks (FCNs), and modified FCNs. Parts of segmentation results of our network are shown in Figure 13.

3.3.1. Qualitative Comparison. From the segmentation results, we selected some representative and challenging images, which have fuzzy boundaries and the pixel intensity
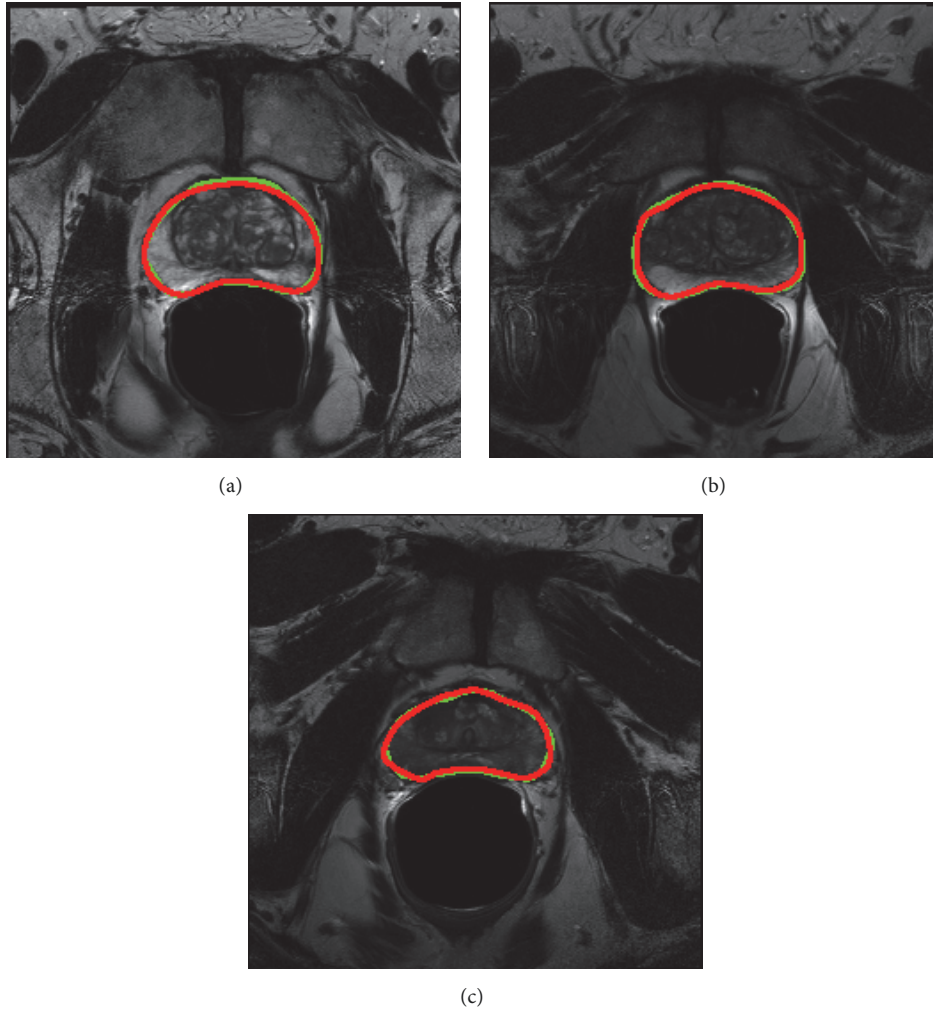
(a)

(b)

(c)

FIGURE 13: Parts of MRI prostate images segmentation results. The red and green contour denote ground truth and the result of UR-Net.

distributions are inhomogeneous both inside and outside. In addition, both prostate and nonprostate regions in those images have similar intensity distributions as shown in Figure 14.

As presented in the third column, FCNs only can detect and segment a part of prostate. And the segmentation results are not accurate, due to the fact that the FCNs model has assigned the labels to a small patch rather than each pixel. Besides, the FCNs ignore the boundaries information. So the FCNs model cannot be directly used in prostate segmentation problem.

As shown in the fourth column, U-Net model has got more accurate segmentation results than FCNs. Because U-Net assigns each label to every pixel and the architecture of U-Net can enhance information propagation through the whole network and improves the network performance, for the slices at the apex and base which lack clear boundary and complete texture, the model cannot segment the prostate accurately.

The results of modified FCNs are shown in the fifth column. Compared with original FCNs, the segmentation

results of modified FCNs are more accurate. From the results, we can see that the modified FCNs can detect more prostate information under the guidance of previous slice. The improvement of modified FCNs can be attributed to the superiority of the architecture of BDC-LSTM. Compared with the traditional convolutional layer, the BDC-LSTM layer can obtain the losing information from adjacent slices and enhance the performance of network.

The sixth column shows the results of V-Net. Compared with FCNs and U-Net, V-Net can take fully use of the 3D spatial information of the volumetric data. However, due to the limited data and memory, each time, V-Net only can receive local volume; this results in V-Net unable to obtain global information. From Figure 14, we can see that the prostate boundaries lose continuity and curvature.

The results of UR-Net are shown in the seventh column. We can observe that the model achieved the best results on prostate segmentation. It can be attributed to the fact that prostate sequence scans can provide more information than a single slice. And the model utilizes interslice information to aid the segmentation process.
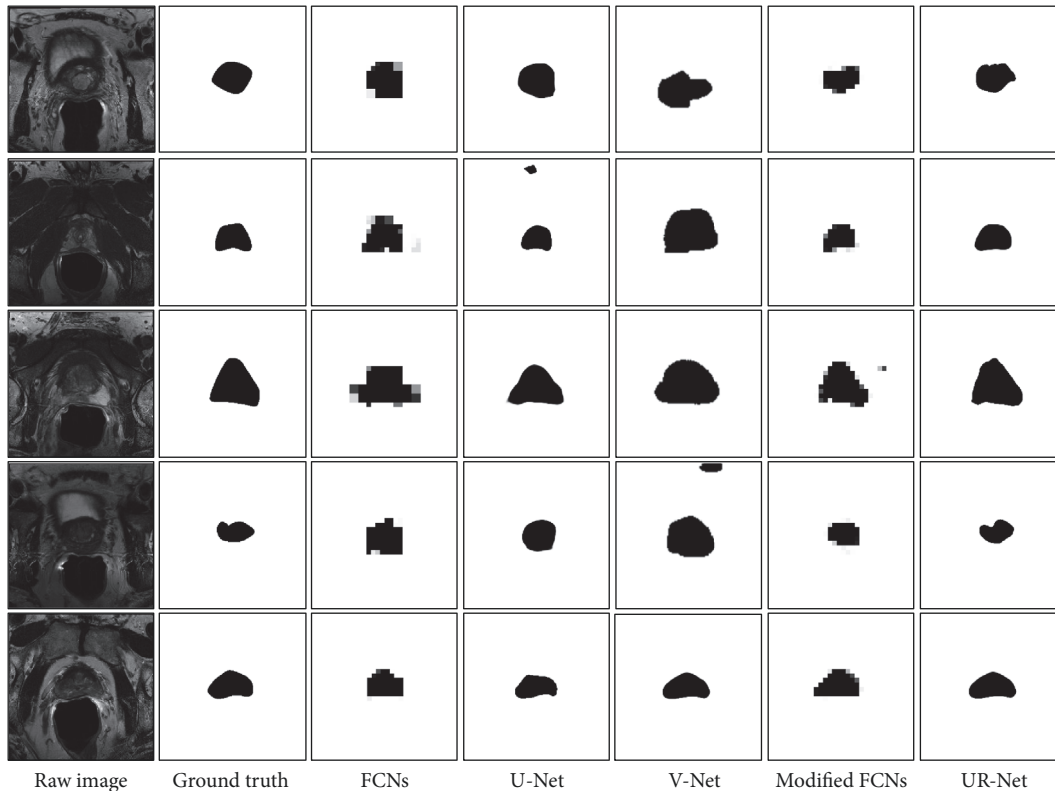
FIGURE 14: Segmentation results. From left to right are raw image, ground truth, the segmentation results of FCNs, U-Net, V-Net, modified FCNs, and UR-Net.

3.3.2. Quantitative Comparison. To quantitatively evaluate the segmentation results, we have computed segmentation results from three aspects as shown in Table 1 including the mean, maximum, and median DSC values. From Table 1, it can be seen that our proposed model obtained the highest scores among all the methods. It shows that the deep neural network with BDC-LSTM layers can obtain promising improvements on prostate MRI images segmentation. Besides, the modified FCNs obtain more accurate segmentation results compared with original FCNs. This improvement should be attributed to the superiority of BDC-LSTM layers, which utilize interslice as shape prior to guide the process of feature extraction and explore necessary information from interslice to alleviate information loss and finally improve the segmentation results.

4. Conclusions

In this paper, we propose a deep neural network with RNNs layers for MRI prostate image segmentation. Different from traditional methods, we treat the prostate scans as sequence data. Except for the local features, we also utilize the interslice information to aid prostate segmentation. In the proposed network, we put three neighboring slices into the network once. And then the network extracts intraslice contexts under the guidance of previous segmentation results from different neighboring slices. Connecting the two different features maps coming from opposite sequential directions together

TABLE 1: Quantitative evaluation results of the segmentation methods.

Methods	DSC values		
	Mean	Maximum	Median
UR-Net	**0.9361**	**0.9772**	**0.9413**
Modified FCNs	0.9047	0.9428	0.9143
V-Net	0.9180	0.9642	0.9380
U-Net	0.9197	0.9684	0.9402
FCNs	0.8489	0.9128	0.8571

can alleviate features lost. Experimental results on extensive MRI prostate image datasets demonstrate that the proposed model achieves better performance than the state-of-the-art convolutional neural networks.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

References

[1] L. Gong, S. D. Pathak, D. R. Haynor, P. S. Cho, and Y. Kim, "Parametric Shape Modeling Using Deformable Superellipses for Prostate Segmentation," *IEEE Transactions on Medical Imaging*, vol. 23, no. 3, pp. 340–349, 2004.

[2] J. Dowling, J. Fripp, P. Greer, S. Ourselin, and O. Salvado, "Automatic atlas-based segmentation of the prostate: a miccai 2009 prostate segmentation challenge entry," in *Medical Image Computing and Computer-Assisted Intervention*, 2009.

[3] J. Wu, Z. Cai, and Z. Gao, "Dynamic K-nearest-neighbor with distance and attribute weighted for classification," in *Proceedings of the 2010 International Conference on Electronics and Information Engineering, ICEIE 2010*, pp. V1356–V1360, Kyoto, Japan, August 2010.

[4] J. Wu, Z. Cai, and X. Zhu, "Self-adaptive probability estimation for Naive Bayes classification," in *Proceedings of the 2013 International Joint Conference on Neural Networks, IJCNN 2013*, pp. 1–8, Dallas, TX, USA, August 2013.

[5] J. Wu, S. Pan, X. Zhu, C. Zhang, and P. S. Yu, "Multiple Structure-View Learning for Graph Classification," *IEEE Transactions on Neural Networks and Learning Systems*, vol. PP, no. 99, pp. 1–16, 2017.

[6] R. Toth, P. Tiwari, M. Rosen et al., "A magnetic resonance spectroscopy driven initialization scheme for active shape model based prostate segmentation," *Medical Image Analysis*, vol. 15, no. 2, pp. 214–225, 2011.

[7] M. Samiee, G. Thomas, and R. Fazel-Rezai, "Semi-automatic prostate segmentation of mr images based on flow orientation," *IEEE International Symposium on Signal Processing and Information Technology*, pp. 203–207, 2006.

[8] S. Klein, "Automatic segmentation of the prostate in 3D MR images by atlas matching using localized mutual information," *Medical Physics*, vol. 35, no. 4, pp. 1407–1417, 2008.

[9] Q. Dou, H. Chen, Y. Jin, L. Yu, J. Qin, and P.-A. Heng, "3D deeply supervised network for automatic liver segmentation from CT volumes," in *Medical Image Computing and Computer-Assisted Intervention*, 2016.

[10] H. Fu, Y. Xu, S. Lin, D. W. K. Wong, and J. Liu, "DeepVessel: Retinal Vessel Segmentation via Deep Learning and Conditional Random Field," in *Medical Image Computing and Computer-Assisted Intervention*, 2016.

[11] H. Yang, J. Sun, H. Li, L. Wang, and Z. Xu, "Deep fusion net for multi-atlas segmentation: Application to cardiac MR images," in *Medical Image Computing and Computer-Assisted Intervention*, 2016.

[12] Q. Zhu, B. Du, B. Turkbey, P. L. Choyke, and P. Yan, "Deeply-supervised CNN for prostate segmentation," in *Proceedings of the 2017 International Joint Conference on Neural Networks, IJCNN 2017*, pp. 178–184, Anchorage, AK, USA, May 2017.

[13] S. Bao and A. C. S. Chung, "Multi-scale structured CNN with label consistency for brain MR image segmentation," *Computer Methods in Biomechanics and Biomedical Engineering: Imaging and Visualization*, pp. 1–5, 2016.

[14] K. Fritscher, P. Raudaschl, P. Zaffino, M. F. Spadea, G. C. Sharp, and R. Schubert, "Deep neural networks for fast segmentation of 3D medical images," in *Medical Image Computing and Computer-Assisted Intervention*, 2016.

[15] H. Huang, X. Hu, Y. Zhao et al., "Modeling Task fMRI Data via Deep Convolutional Autoencoder," *IEEE Transactions on Medical Imaging*, 2017.

[16] J. Lv, B. Lin, W. Zhang et al., "Task fmri data analysis based on supervised stochastic coordinate coding," *Medical Image Analysis*, vol. 38, pp. 1–16, 2017.

[17] Y. Xie, Z. Zhang, M. Sapkota, and L. Yang, "Spatial clockwork recurrent neural network for muscle perimysium segmentation," in *Medical Image Computing and Computer-Assisted Intervention*, 2016.

[18] J. Chen, L. Yang, Y. Zhang, M. S. Alber, and D. Z. Chen, "Combining fully convolutional and recurrent neural networks for 3D biomedical image segmentation," in *In NIPS Proceedings*, MIT Press, Cambridge, MA, USA, 2016.

[19] X. Qin, X. Li, Y. Liu, H. Lu, and P. Yan, "Adaptive shape prior constrained level sets for bladder MR image segmentation," *IEEE Journal of Biomedical and Health Informatics*, vol. 18, no. 5, pp. 1707–1716, 2014.

[20] R. Dipietro, C. Lea, A. Malpani et al., "Recognizing surgical activities with recurrent neural networks," in *Medical Image Computing and Computer-Assisted Intervention*, 2016.

[21] R. P. K. Poudel, P. Lamata, and G. Montana, "Recurrent fully convolutional neural networks for multi-slice MRI cardiac segmentation," in *Medical Image Computing and Computer-Assisted Intervention*, 2016.

[22] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[23] A. Graves, A.-R. Mohamed, and G. Hinton, "Speech recognition with deep recurrent neural networks," in *Proceedings of the 38th IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '13)*, pp. 6645–6649, May 2013.

[24] J. T. Connor, R. D. Martin, and L. E. Atlas, "Recurrent neural networks and robust time series prediction," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 5, no. 2, pp. 240–254, 1994.

[25] A. Graves, "Generating sequences with recurrent neural networks," 2013, arXiv preprint arXiv:1308.0850.

[26] P. J. Werbos, "Backpropagation through time: what it does and how to do it," *Proceedings of the IEEE*, vol. 78, no. 10, pp. 1550–1560, 1990.

[27] R. J. Williams and D. Zipser, "A learning algorithm for continually running fully recurrent neural networks," *Neural Computation*, vol. 1, no. 2, pp. 270–280, 1989.

[28] J. H. Lee and N. L. Ricker, "Extended Kalman filter based nonlinear model predictive control," *Industrial & Engineering Chemistry Research*, vol. 33, no. 6, pp. 1530–1541, 1994.

[29] A. J. Robinson and F. Fallside, "The utility driven dynamic error propagation network," Technical Report CUED/F-INFENG/TR.1, Cambridge University Engineering Department, 1987.

[30] O. Ronneberger, P. Fischer, and T. Brox, "U-net: convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention*, pp. 234–241, 2015.

[31] F. Milletari, N. Navab, and S.-A. Ahmadi, "V-Net: fully convolutional neural networks for volumetric medical image segmentation," in *Proceedings of the 4th International Conference on 3D Vision (3DV '16)*, pp. 565–571, IEEE, October 2016.

[32] P. Charles, Project title. GitHub repository (2013).