Fully Automatic Left Atrial Segmentation from MRI Sequences Using Deep CNN and UKF

UNIVERSITY OF ALBERTA

Xiaoran Zhang[1, 2], Glynn Martin[1], Michelle Noga[1], Kumaradevan Punithakumar[1]

[1]Department of Radiology and Diagnostic Imaging, Servier Virtual Cardiac Centre, Mazankowski Alberta Heart Institute, University of Alberta [2]Department of Automation, Beijing Institute of Technology, Beijing China

Introduction

Motivation

Left atrial anatomical knowledge is clinically important for atrial fibrillation ablation guidance, fibrosis quantification and biophysical modeling. Over the last few decades, a rapid development of segmenting left atrium has been made using Magnetic Resonance Imaging (MRI). The advances in cardiac MRI imaging have also provided large amount of data with an increasingly high level of quality, which contributes significantly to the imaging based diagnostics.

Difficulties

1) The size of the left atrium is relatively small as compared to left or right ventricles in cardiac MRI images.
2) Boundaries are not clearly defined when the blood pool of the left atrium goes into preliminary veins.
3) The shape variability of the left atrium is large between different patients.
4) Delineation is time consuming and much more complex compared to segmenting CT images.

Our Method

We propose a modified U-net structure by adding a step activation function after training three separate U-net model with 5-fold cross-validation. After analyzing images for high intensity variance, we notice that U-net model fails to produce temporally consistent predictions. Therefore, we use Unscented Kalman Filter by proposing a dynamic model for temporal periodicity to process the image masks predicted by modified U-net model.

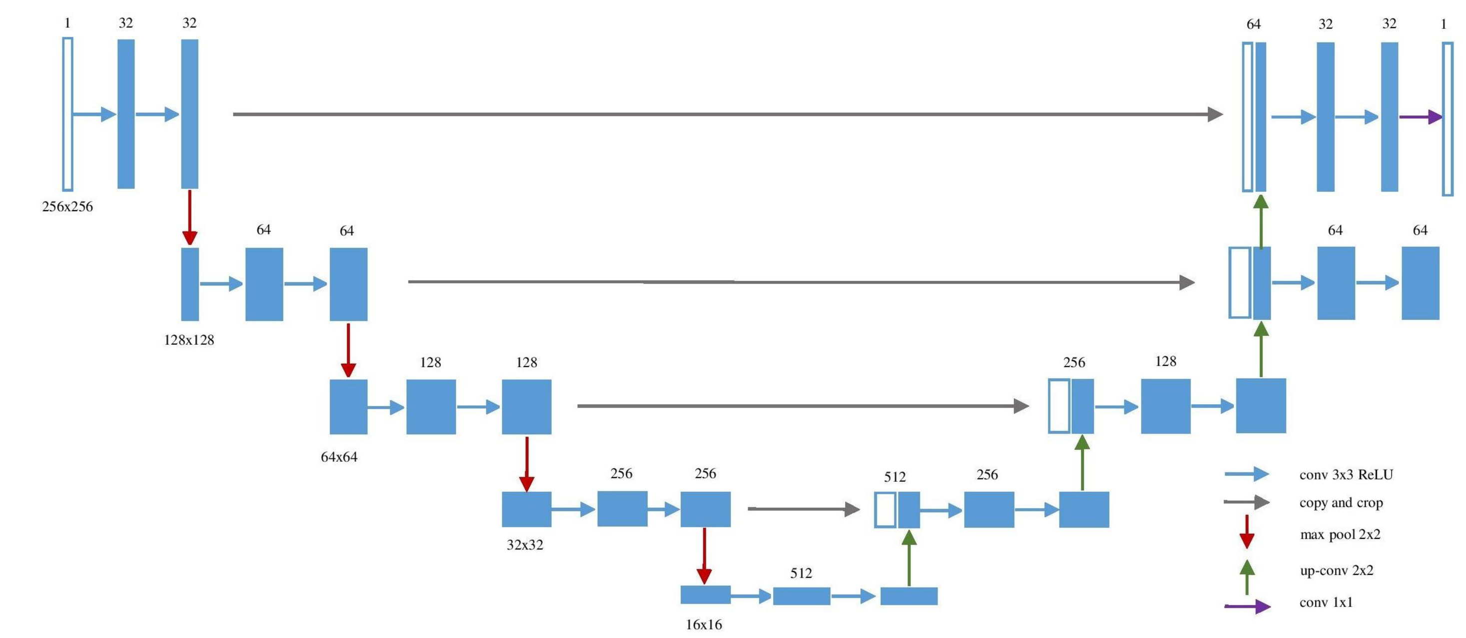

Fig.1 U-net architecture. Raw input images are MRI sequences and each image is resized to 256x256 pixels using padding method.

U-net Structure

U-net structure [1] contains convolutional, rectifier, max pool and up convolutional layer. The loss function for the model is implemented as the negative of Dice coefficient illustrated as follows.

$$loss = -\frac{2|X \cap Y|}{|X|+|Y|}$$

where X is the prediction set during training and Y is the desired label.

Variance Computation

After obtaining raw output images from U-net, we compute self variance for each image. We notice that images with high variance cannot be predicted properly by the modified U-net structure. Thus, we set a step activation function to select those images that need to be processed, where a is the self variance value for each image and y is the binary label indicating the selection decision.

$$y = \begin{cases} 1 & a > 10000 \\ 0 & 0 \le a \le 10000 \end{cases}$$

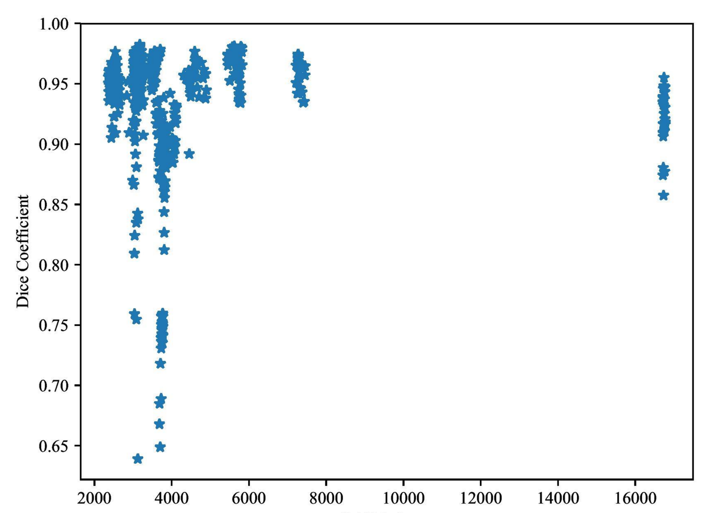

Fig.2 Self variance versus Dice Coefficient for Train 80 scenario.

Reliability Analysis

In order to evaluate the impact of training samples on prediction performance, we calculate dice metric versus reliability [2]. The ideal curve lies in the right upper corner. The reliability is illustrated as follows. From the figure, we can find that there is a significant improvement from Train 20 to 60 while the improvement is minor from Train 60 to 80.

$$S(t) = P(\{T > t\}) = \int_t^{\infty} f(u)du = 1 - F(t)$$

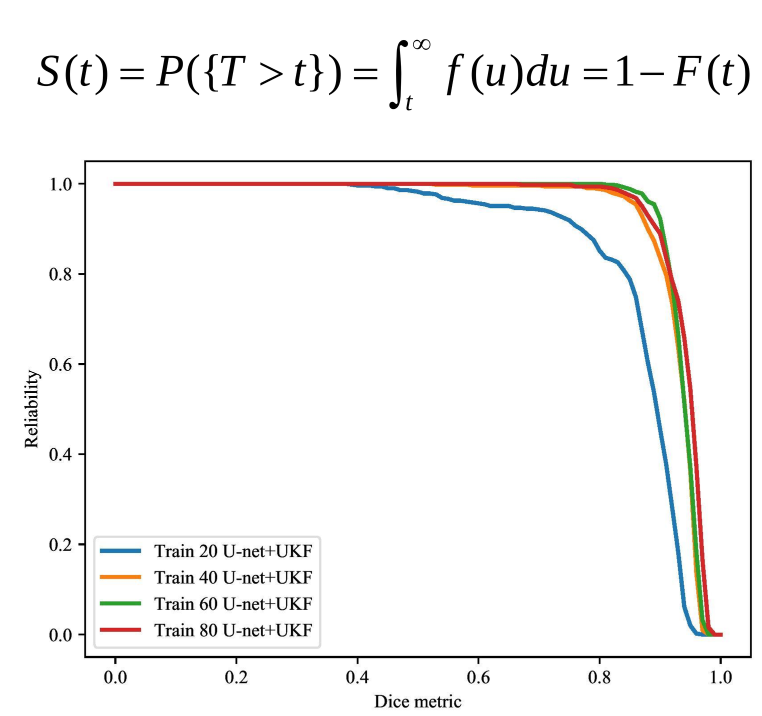

Fig.3 Reliability versus dice metric

Unscented Kalman Filter

Unscented Kalman Filter is implemented after the intensity variance computation and selection based on a state-space model proposed in [3] by adding an angular time-varying frequency. The input to Unscented Kalman Filter is the contour matrix in Cartesian coordinates obtained from prediction image masks.

We select mean position, velocity, current position in x-y coordinate and time-varying angular velocity as state variable illustrated as follows. $S_k = [\bar{x}_k \; x_k \; \dot{x}_k \; \bar{y}_k \; y_k \; \dot{y}_k \; \omega_k]$

where k indicates the frame number. In our case, the maximum value of the k is equal to 25 or 30. The discrete-time dynamic model is given as follows.

$$S_{k+1} = f_k(S_k) + v_k$$

where $f_k(S_k)$ is the predication equation and v_k denotes a Gaussian process noise sequence.

$$f_k(S_k) = \begin{bmatrix} F_k & 0_{3\times3} & 0_{3\times1} \\ 0_{3\times3} & F_k & 0_{3\times1} \\ 0_{3\times3} & 0_{3\times3} & 1 \end{bmatrix} \quad F_k = \begin{bmatrix} 1 & 0 & 0 \\ 1-\cos(\omega_k\Delta T) & \cos(\omega_k\Delta T) & \frac{1}{\omega_k}\cos(\omega_k\Delta T) \\ \omega_k\sin(\omega_k\Delta T) & -\omega_k\sin(\omega_k\Delta T) & \cos(\omega_k\Delta T) \end{bmatrix}$$

where ω_k denotes the time-varying angular frequency and ΔT is the time interval in a cardiac cycle.

$$\omega_k = \frac{2\pi \times HeartRate}{60} \quad \Delta T = \frac{60}{HeartRate \times k} \quad v_k = \begin{bmatrix} Q_k & 0_{3\times3} & 0_{3\times1} \\ 0_{3\times3} & Q_k & 0_{3\times1} \\ 0_{3\times3} & 0_{3\times3} & 1 \end{bmatrix}$$

where Q_k denotes the covariance of process noise. The elements of Q_k given by:

$$q_{11} = q_1^2\Delta T \quad q_{12} = q_{21} = \frac{q_1^2(\omega_k\Delta T - \sin(\omega_k\Delta T))}{\omega_k} \quad q_{13} = q_{31} = q_1^2(1-\cos(\omega_k\Delta T))$$

$$q_{22} = \frac{q_1^2{\omega_k}^2(3\omega_k\Delta T - 4\sin(\omega_k\Delta T) + \cos(\omega_k\Delta T)) + q_2^2\sin^2(\omega_k\Delta T)}{\omega_k}$$

$$q_{23} = q_{32} = \frac{q_1^2{\omega_k}^2(1-2\cos(\omega_k\Delta T) + \cos^2(\omega_k\Delta T)) + q_2^2\sin^2(\omega_k\Delta T)}{2{\omega_k}^2}$$

$$q_{33} = -\frac{q_1^2\cos(\omega_k\Delta T)\sin(\omega_k\Delta T) - q_2^2(\cos(\omega_k\Delta T)\sin(\omega_k\Delta T) - \omega_k\Delta T)}{2\omega_k}$$

The measurement equation, which is also update equation, is given by $z_k = H_k S_k + \eta_k$. In order to extract variables for UKF output, the measurement matrix is given as follows to extract the position in x and y coordinate. The η_k indicates a zero-mean Gaussian noise where r=1e-2.

$$H_k = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix} \quad \eta_k = \begin{bmatrix} r & 0 \\ 0 & r \end{bmatrix}$$

The initialization of Unscented Kalman Filter is also important as the model has no prior knowledge of initial value. Therefore, a two point-difference method is implemented. The initial value in x coordinate in the state vector is given by:

$$\hat{x}_1 = z_{1,i} \quad \hat{\dot{x}}_1 = \frac{z_{2,j} - z_{1,j}}{\Delta T} \quad \hat{\bar{x}}_1 = \frac{1}{k}\sum_{j=1}^{k} z_{j,i}$$

The corresponding initial covariance is given by

$$P_1 = \begin{bmatrix} \phi_1 & 0_{3\times3} & 0_{3\times1} \\ 0_{3\times3} & \phi_1 & 0_{3\times1} \\ 0_{3\times3} & 0_{3\times3} & 1 \end{bmatrix} \quad \phi_1 = \begin{bmatrix} r & \frac{r}{k} & \frac{r}{k\Delta T} \\ \frac{r}{k} & r & \frac{r}{\Delta T} \\ \frac{r}{k\Delta T} & \frac{r}{k\Delta T} & \frac{2r}{\Delta T^2} \end{bmatrix}$$

Results

Our system was trained and tested on 100 patients clinical MRI images obtained by Mazankowski Alberta Heart Institute. We classify our dataset according to the number of chamber and choose 2-chambers group in this experiment. Each patient has more than 25 frames per each heart beat, which allows us to take advantage of time-serial data information by applying Unscented Kalman Filter. In order to evaluate the impact of training samples on our system, we design several scenarios by increasing the number of patients in the training data sets from 20 to 80.

Our system was trained and tested using the modified U-net structure with 5-fold of cross validation. Three U-net models were trained separately and the final prediction were determined as the average of these prediction after a step activation function.

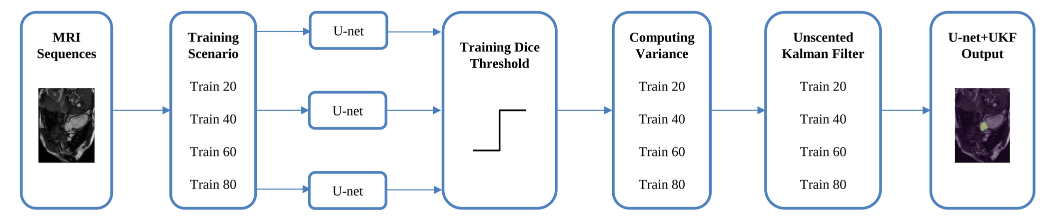

Fig.4 Implementation details with MRI sequences as raw input. Each image is resized to 256x256 pixels. Several scenarios are implemented with different training samples.

Method	Dice Coefficient		Hausdorff Distance		Root Mean Squre Error	
	U-net	U-net+UKF	U-net	U-net+UKF	U-net	U-net+UKF
Train 20	0.825±0.135	0.868±0.097	12.188±8.538	10.228±7.727	10.843±5.870	8.495±5.307
Train 40	0.918±0.076	0.929±0.042	8.603±6.845	7.652±6.187	8.575±5.255	6.944±4.187
Train 60	0.937±0.026	0.937±0.027	6.657±3.512	5.956±2.924	6.638±2.753	5.408±2.330
Train 80	0.933±0.058	0.941±0.037	7.022±4.319	6.020±3.308	6.191±3.474	5.373±3.192

Table.1 Validation results for patients. Dice coefficient illustrates the overlap accuracy between prediction masks with desired labels. Hausdorff distance reports the distance of largest pixel deviation from prediction contours with desired ones. RMSE demonstrates the standard deviation between prediction masks with label images.

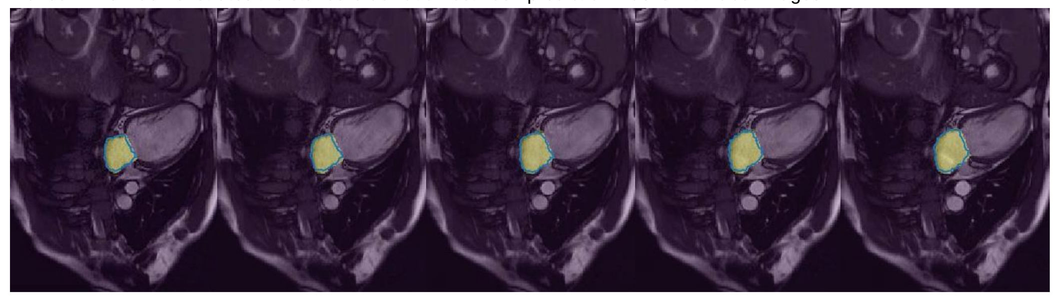

Fig.5 U-net+UKF prediction results in MRI sequences. For each image, the yellow area is the desired label and blue contour is the prediction results.

References

[1] Ronneberger, O., Fischer, P., & Brox, T. (2015, October). U-net: Convolutional networks for biomedical image segmentation. In International Conference on Medical image computing and computer-assisted intervention (pp. 234-241). Springer, Cham.
[2] Ayed, I. B., Punithakumar, K., Li, S., Islam, A., & Chong, J. (2009, September). Left ventricle segmentation via graph cut distribution matching. In International Conference on Medical Image Computing and Computer-Assisted Intervention (pp. 901-909). Springer, Berlin, Heidelberg.
[3] Punithakumar, K., Ayed, I. B., Islam, A., Goela, A., Ross, I. G., Chong, J., & Li, S. (2013). Regional heart motion abnormality detection: An information theoretic approach. Medical image analysis, 17(3), 311-324.

北京理工大學 BEIJING INSTITUTE OF TECHNOLOGY
MAZANKOWSKI ALBERTA HEART INSTITUTE
Mitacs