Supplementary Document One shot PACS: Patient specific Anatomic Context and Shape prior aware recurrent registration-segmentation for longitudinal thoracic cone beam CTs

Jue Jiang and Harini Veeraraghavan

I. Additional Results

1) PlanCT to CBCT deformation: Example registrations from planning CT to CBCT using PACS-aware method are shown in Fig. 1. The registration in the recurrent steps are also shown. The top two rows correspond to one example and the bottom two rows correspond to a different example. The warped pCT images through the recurrent steps are also shown (row 2 and row 4 Fig. 1).

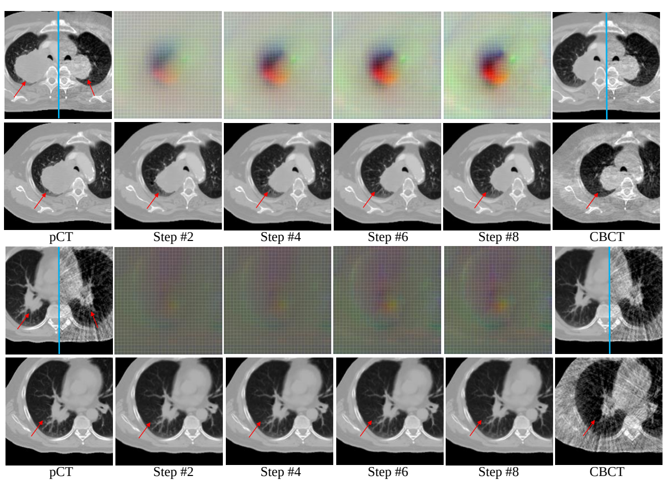

Fig. 1. Example registrations with progressive deformations for two representative cases. The tumor location is indicated with a red arrow.

2) Simple diagram of CDS, RNN and (C)LSTM: Fig. 2 shows the comparison between classical dynamic system(CDS), basic RNN and (convolutional) LSTM.

3) Esophagus segmentation on weekly cone beams: Fig .3 shows esophagus segmentation from weekly cone beams.

4) Segmentation accuracy versus recurrent steps: Fig. 4 shows the segmentation accuracy expressed using Dice similarity coefficient (DSC) with increasing number of recurrent steps. The computing times for those steps are also shown.

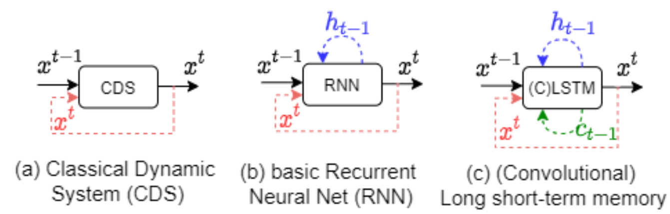

Fig. 2. Simple diagram of simple dynamic system, basic recurrent network and (C)LSTM.

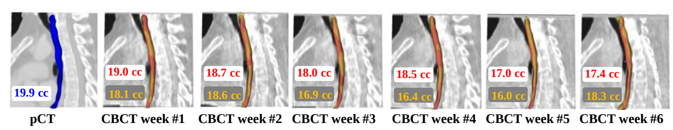

Fig. 3. Esophagus segmentation (red) and manual contour (yellow) longitudinally visualization from PlanCT (blue). The esophagus volume are shown in the figure. We only show the esophagus ROI region for better visualization.

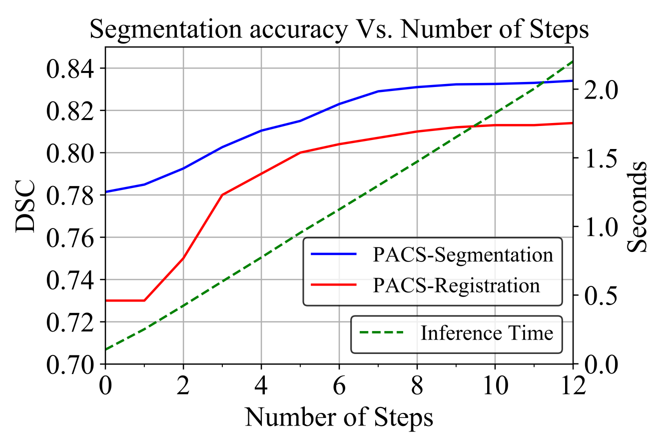

Fig. 4. Segmentation accuracy and the computation times with increasing number of recurrent steps.

5) Fusion strategies studied for combining the anatomic context and shape prior with CBCT for segmentation: Fig. 5 shows the two different strategies used for computing the CBCT segmentation, namely the early fusion (Fig. 5(a)) and intermediate fusion(Fig. 5(b)). The early fusion method combines the anatomic context (pCT) and shape prior (pCT delineation) after progressive warping using the recurrent registration into the recurrent units placed in the encoder of the CBCT segmentor. The intermediate fusion combines the encoded features of the final warped pCT and its delineation with the CBCT encoded features in the decoder of the CBCT segmentor.

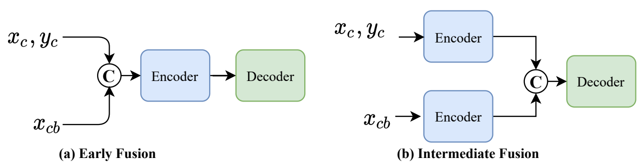

Fig. 5. Early vs intermediate fusion for CBCT segmentation.

6) Descending aorta contour during registration: Fig. 6 shows the descending aorta contour from pCT (a) deformed to (b) and overlaid on the CBCT image(c).

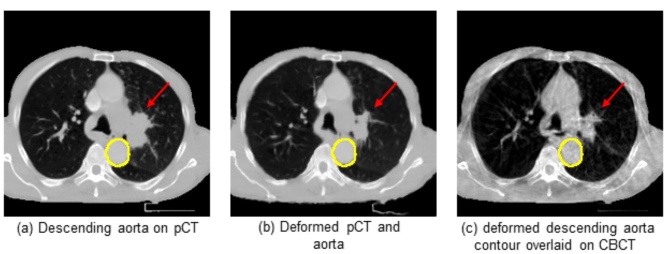

Fig. 6. The descending aorta contour deformation from pCT to CBCT.

7) Ablation experiments results: Fig. 7 shows the segmentations resulting from the various ablation experimental settings on a representative case in the testing set.

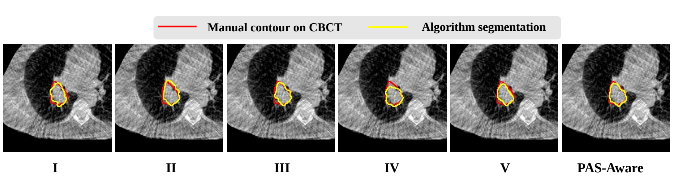

Fig. 7. Computed segmentations using models trained under the ablation experiments setting. (I) w/o shape context prior, (II) w/o the anatomic context, (III) w/o recurrent encoder of the segmentation network, (IV) w/o CLSTM is in segmentation s and (V) w/o OHEM

8) Longitudinal esophagus analysis: Fig. 8 shows longitudinal DSC and HD95 accuracy for segmenting esophagus. The percent slope is also shown.

9) Descending aorta alignment: Fig. 6 shows the alignment of descending aorta in a slice containing the tumor (red arrow) from pCT (a) to CBCT(c).

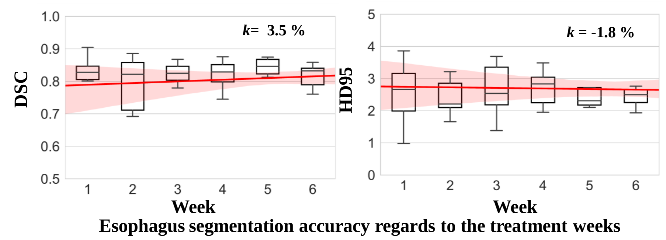

Fig. 8. Esophagus segmentation accuracy computed at different weeks of treatment. Percent slope of accuracy changes is also shown.

II. THE DETAILS OF NETWORK STRUCTURES

The details of each network structure is shown. The registration net g is shown in Table 2. The segmentation network s is shown in Table 3. The CLSTM in our implementation are indicated in blue font.

TABLE I
THE NETWORK ARCHITECTURE USED FOR REGISTRATION. WE USE THE FOLLOWING ABBREVIATION FOR EASE OF PRESENTATION: N=NUMBER OF FEATURES; K=KERNEL SIZE; S=STRIDE SIZE; CLSTM=CONVOLUTIONAL LONG SHORT-TERM MEMORY; VECINC=DIFFEOMORPHIC INTERGRATION LAYER.

Layers	Registration net G	Concatenation
1	CONV-(N16,K3,S2) LeakyReLu	1
2	CLSTM-(N16,K3,S2) LeakyReLu	2
3	CLSTM-(N16,K3,S2) LeakyReLu	3
4	CLSTM-(N16,K3,S2) LeakyReLu	4
5	CONV-(N16,K3,S1)	-
6	CONV-(N32,K3,S1) LeakyReLu	4
7	CONV-(N32,K3,S1) LeakyReLu	3
8	CONV-(N32,K3,S1) LeakyReLu	2
9	CONV-(N32,K3,S1) LeakyReLu	1
10	CONV-(N16,K3,S1) LeakyReLu	-
11	CONV-(N3,K3,S1) LeakyReLu	-
12	VecInC	-

TABLE II
THE UNET NETWORK ARCHITECTURE USED FOR SEGMENTATION S. WE USE THE FOLLOWING ABBREVIATION FOR EASE OF PRESENTATION: N=NUMBER OF FEATURES; K=KERNEL SIZE; S=STRIDE SIZE;CLSTM=CONVOLUTIONAL LONG SHORT-TERM MEMORY;

Layers	Unet	Concatenation
1	CONV-(N32,K3,S1), ReLu	-
2	CLSTM-(N32,K3,S1), ReLu	1
3	Max-Pooling (S2)	-
4	CONV-(N64,K3,S1), ReLu	-
5	CLSTM-(N64,K3,S1), ReLu	2
6	Max-Pooling (S2)	-
7	CONV-(N128,K3,S1), ReLu	-
8	CLSTM-(N128,K3,S1), ReLu	3
9	Max-Pooling (S2)	-
10	CONV-(N256,K3,S1), ReLu	-
11	CLSTM-(N256,K3,S1), ReLu	4
12	Max-Pooling (S2)	-
13	CONV-(N512,K3,S1), ReLu	-
14	CLSTM-(N512,K3,S1), ReLu	-
15	UP-Pooling (S2)	4
16	CONV-(256,K3,S1), ReLu	-
17	CONV-(N256,K3,S1), ReLu	-
18	UP-Pooling (S2)	3
19	CONV-(N128,K3,S1), ReLu	-
20	CONV-(N128,K3,S1), ReLu	-
21	UP-Pooling (S2)	2
22	CONV-(N64,K3,S1), ReLu	-
23	CONV-(N64,K3,S1), ReLu	-
24	UP-Pooling (S2)	1
25	CONV-(N32,K3,S1), ReLu	-
26	CONV-(N2,K1,S1), Softmax	-