Localizing the Recurrent Laryngeal Nerve via Ultrasound with a Bayesian Shape Framework

Haoran Dou* , Luyi Han*, Yushuang He, Jun Xu, Nishant Ravikumar, Ritse Mann, Alejandro F. Frangi, Pew-Thian Yap, and Yunzhi Huang✉

(* equally contributed to this work)

Background

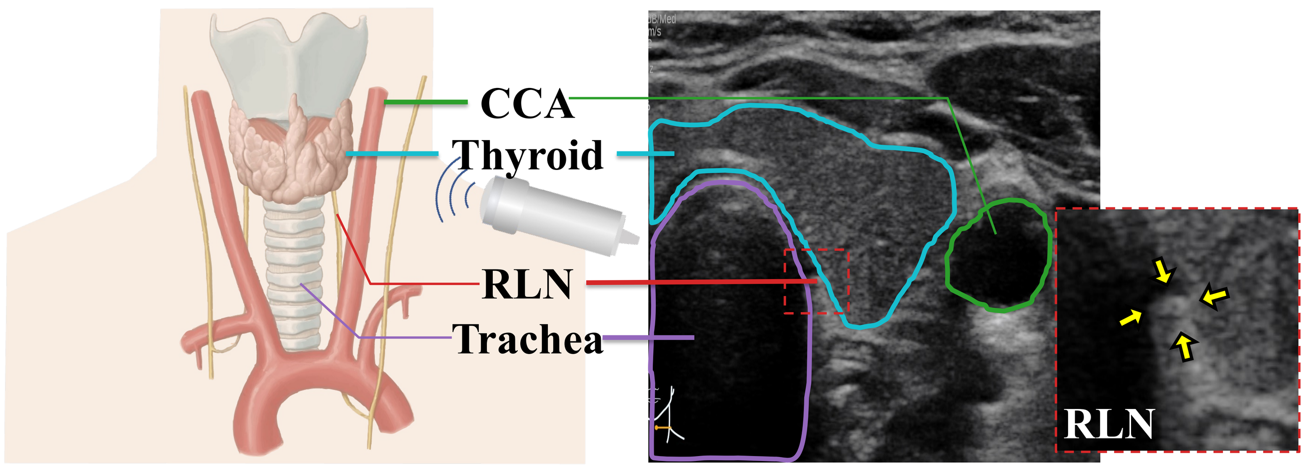

Fig. 1. Ultrasound imaging of bilateral RLNs.

We propose the first learning-based framework to identify the RLN from a US image for pre-operative assessment of contraindication for robotic thyroidectomy. To avoid missing the tiny RLN in the midst of significantly larger structures, we introduce a method based on anatomical prior knowledge for RLN detection.

Highlights and Materials

Highlights

- We introduce Bayesian shape alignment for geometrical constraints, allowing the utilization of spatial prior knowledge in determining an ROI enclosing the RLN;
- We introduce Locate-Net, a dual-path network that uses both local and global information to refine the localization of the RLN centroid.

Materials

465 patients diagnosed with thyroid cancer by preoperative biopsy and enrolled for thyroidectomy participated in this study. Each patient has both left and right scans of the RLN. Each scan contains a variable number of qualified US frames, ranging from 1 to 4. 325, 46 and 94 subjects were randomly selected for training, validation, testing, respectively.

Motivation and Method

Clinically, surgeons recognize the RLN based on its surrounding organs (SO), including the common carotid arteries (CCA), thyroid, and trachea. The spatial relationships of these anatomical structures are typically consistent, but not entirely identical, across individuals. Here, we incorporate the spatial relationship into the Bayesian inference with the following mathematical model for a given image I:

$$q(RLN, SO|I) \propto p(I|RLN, SO) \times p(RLN|SO) \times p(SO)$$

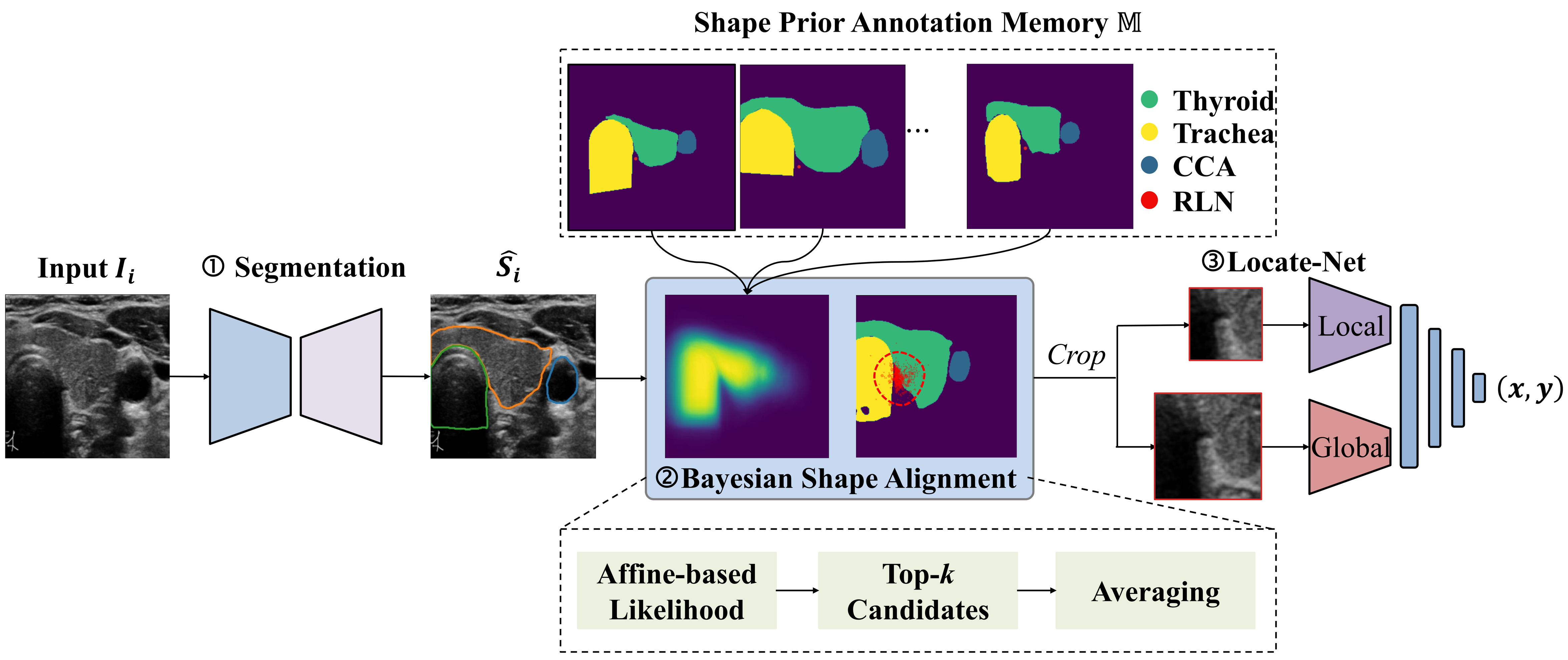

Fig. 2. Overview of the proposed framework.

Fig.2 illustrates the proposed framework for identifying the RLN from a US image using anatomical prior knowledge. Our framework includes three coarse-to-fine sequential modules: (1) Segmentation module; (2) Bayesian shape alignment (BSA) module; and (3) Locate-Net module. The segmentation module obtains the segmentations $\hat{S}$ of organs surrounding the RLN, including the CCA, thyroid, and trachea. These segmentations form posteriors for the BSA module to infer the candidate coordinates of the RLN. Finally, Locate-Net refines the RLN centroid using local details and global contexts based on the patch centered at the inferred candidate coordinates.

Results

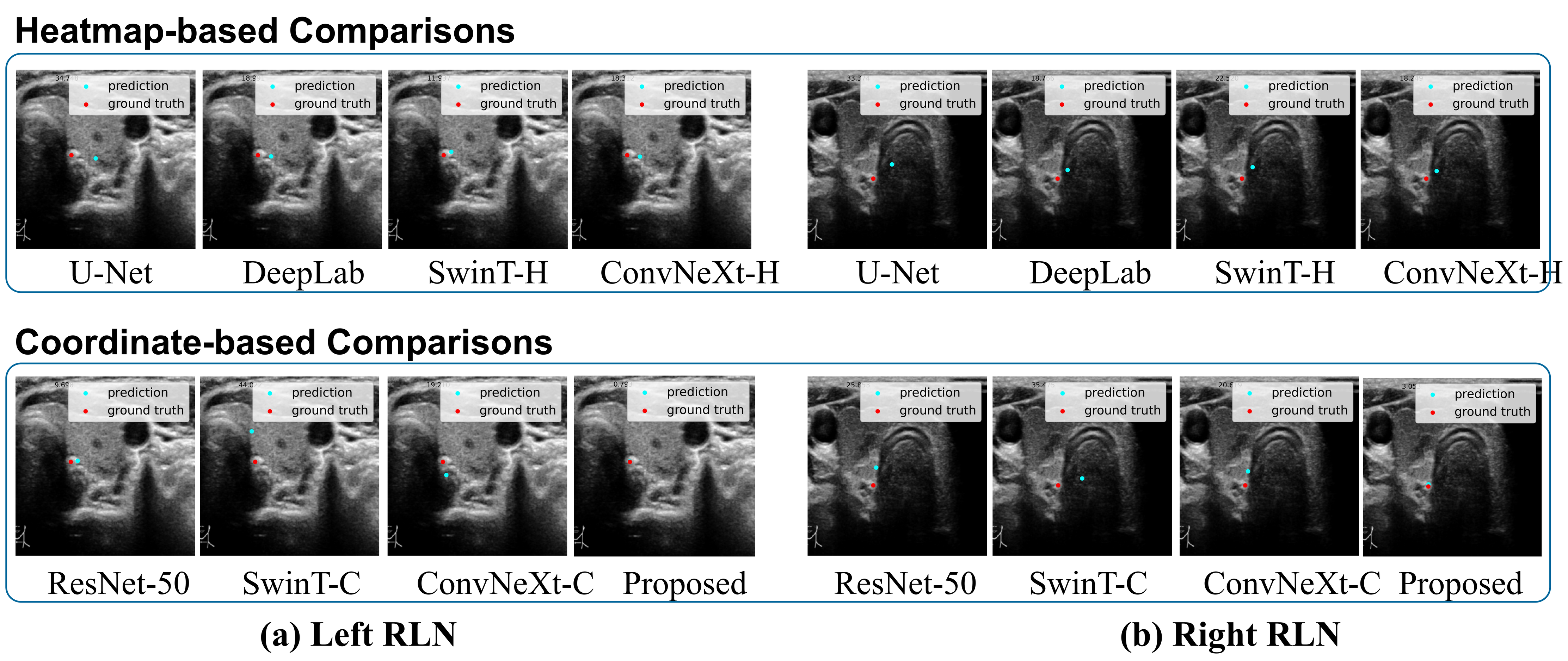

Fig. 3. Example results from competing methods

Comparison between the competitive methods and our method.

	Method	Dist [pix]	Hit Rate [%]
Heatmap-based Comparisons	U-Net[1]	29.3 ± 12.8	11.5
	DeepLab[2]	17.5 ± 8.0	41.2
	SwinT-H[3]	22.7 ± 13.6	30.8
	ConvNeXt-H[4]	12.7 ± 12.1	73.1
Coordinate-based Comparisons	ResNet-50[5]	10.9 ± 9.7	77.5
	SwinT-C[3]	19.7 ± 11.8	42.3
	ConvNeXt-C[4]	16.5 ± 8.2	47.3
	Proposed	**3.5 ± 7.5**	**95.6**

(a) Results of left RLN

	Method	Dist [pix]	Hit Rate [%]
Heatmap-based Comparisons	U-Net[1]	20.9 ± 10.5	31.9
	DeepLab[2]	11.9 ± 7.1	71.3
	SwinT-H[3]	20.2 ± 10.6	35.6
	ConvNeXt-H[4]	13.1 ± 8.7	64.4
Coordinate-based Comparisons	ResNet-50[5]	12.3 ± 8.4	70.0
	SwinT-C[3]	14.4 ± 9.6	59.4
	ConvNeXt-C[4]	14.0 ± 9.5	62.5
	Proposed	**4.6 ± 7.6**	**92.5**

(b) Results of right RLN

Contact

Luyi Han
Department of Radiology,
Netherlands Cancer Institute (NKI),
Email: hanluyi4869@gmail.com
Website: https://fiy2w.github.io/
Phone: +31 620466303

Paper Access

Code Access

References

[1] U-net: Convolutional networks for biomedical image segmentation.
[2] Panoptic-deeplab: A simple, strong, and fast baseline for bottom-up panoptic segmentation. In: Computer Vision and Pattern Recognition (2020)
[3] Swin-tformer: Hierarchical vision transformer using shifted windows.
[4] A ConvNet for the 2020s.
[5] Deep residual learning for image recognition. In: Computer Vision and Pattern Recognition (2016)