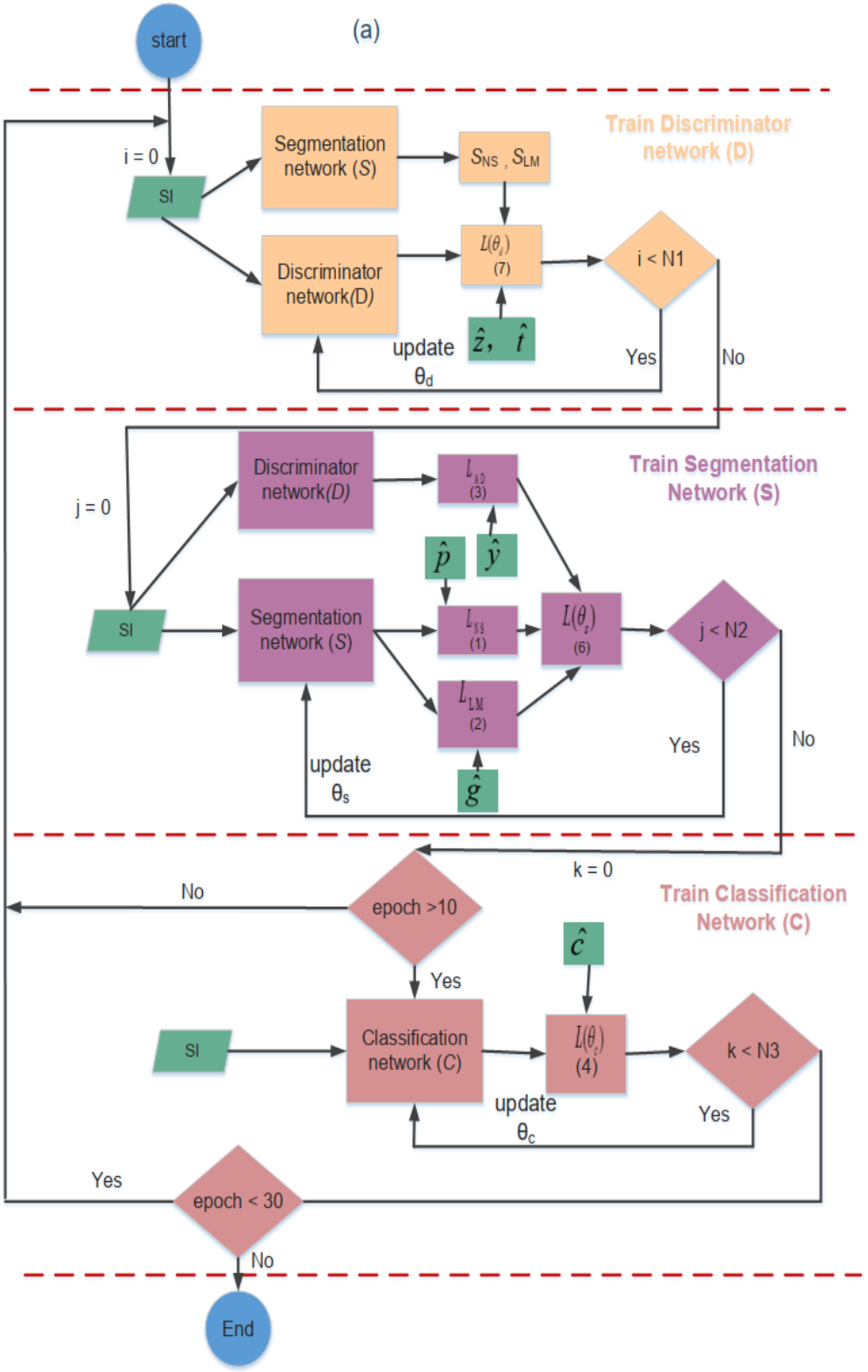
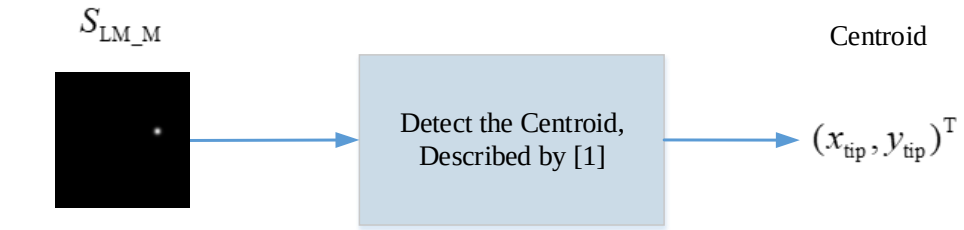
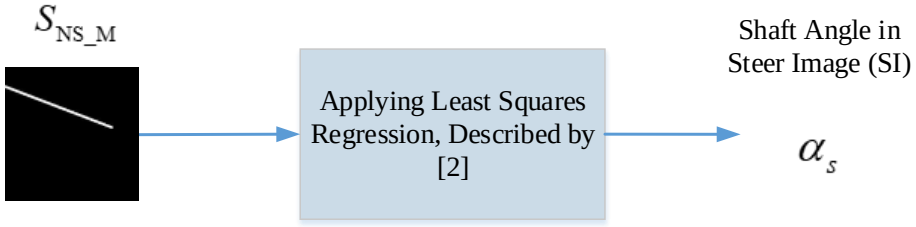
Supplemementary material

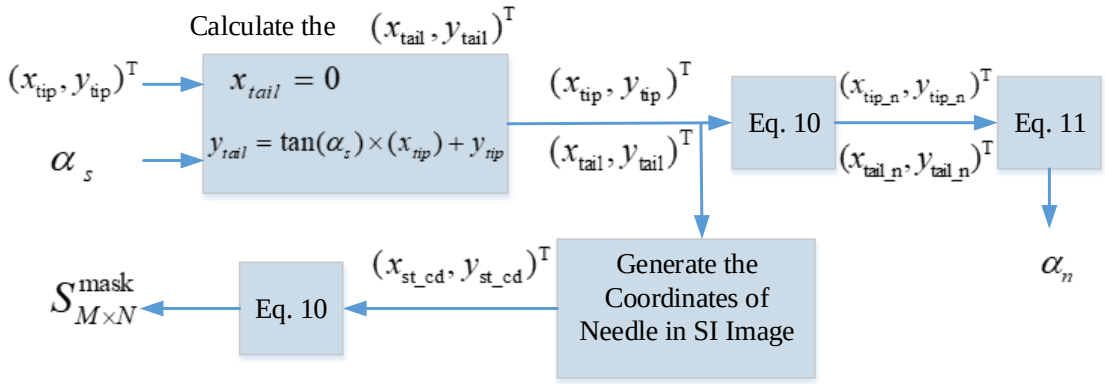




(b)



(c)



(d)

Fig. A1 (a) Training process of the MTL framework. N1, N2, N3 denote the number of iterations of discriminator network, segmentation network and classification network in each epoch respectively. (b) Detect the Centroid in the S_{LM_M}. (c) Estimate the shaft angle α_s in the S_{NS_M}. (d) Generate the needle mask ($S^{mask}_{M\times N}$) and calculate the steer angle α_n.

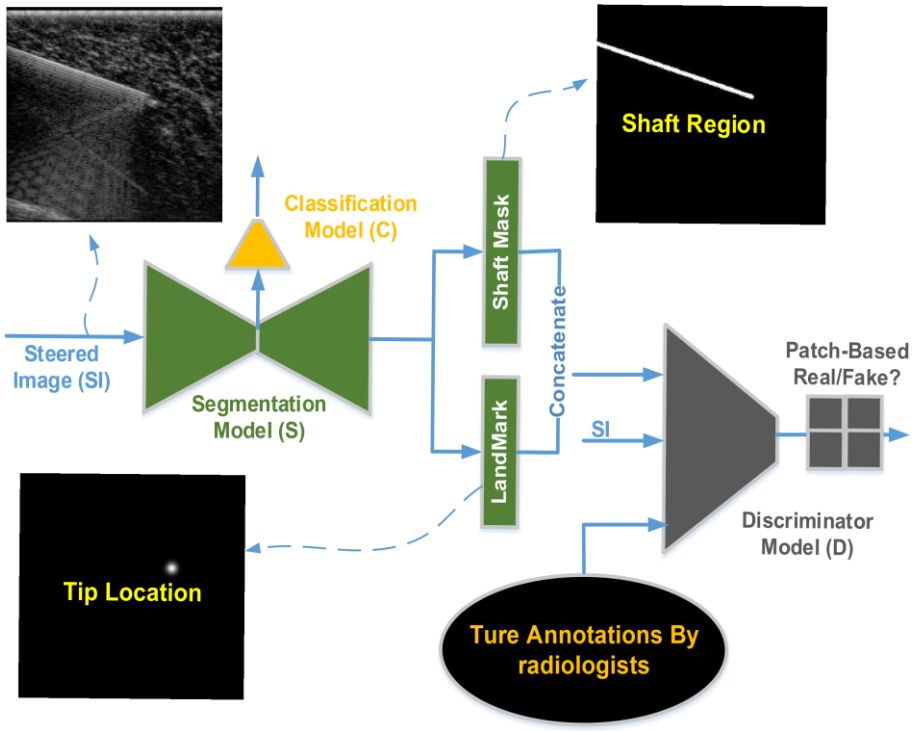


Fig. A2. Architecture of proposed MTL network. The MTL framework consists of a segmentation network (S), a discriminator network (D) and a classification network (C). Discriminator network (D) only be used in the training phase.

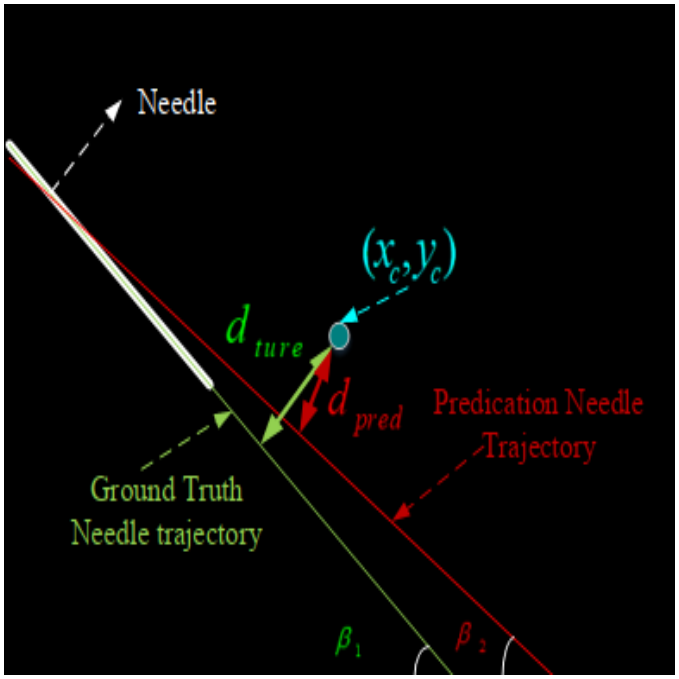


Fig. A3 Illustration of the targeting error (TE) and shaft localization error (SLE).

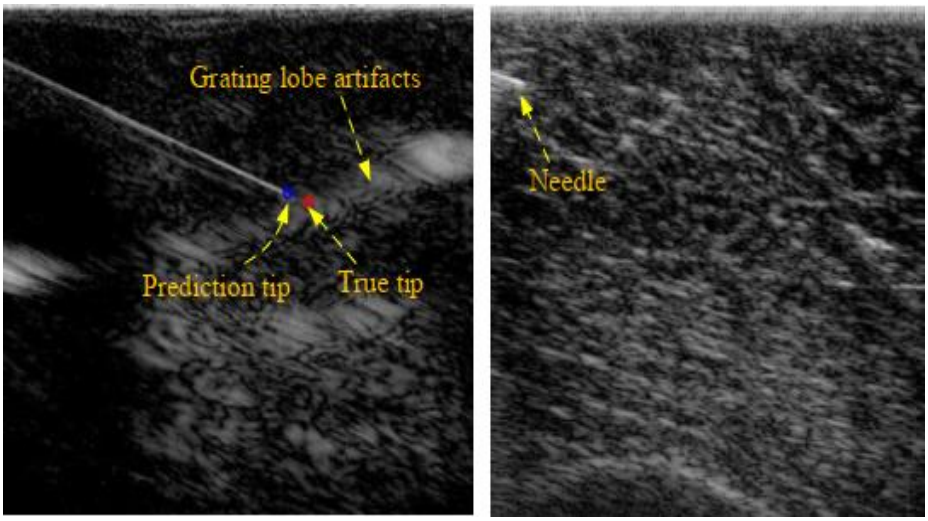


Fig. A4 Illustration of failure cases. The needle tip is contaminated by grating lobe artifacts. The blue circle is the prediction tip location and the red circle is the true tip location.

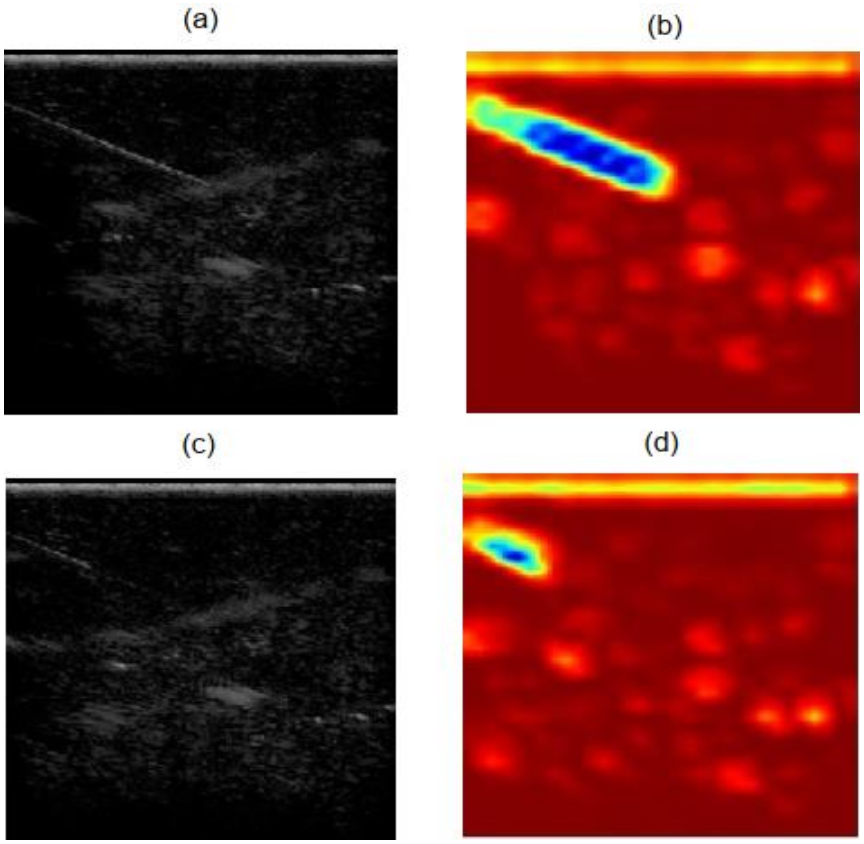


Fig. A5 Illustration of the feature map difference between short and long length samples for needle to insert. (a) Long length needle sample (b) Feature maps of (a), generating from the visualization of the last convolution layer of the classification network (C). (c) Short length needle sample. (d) Feature maps of (c), generating from the visualization of the last convolution layer of the classification network (C).

Table A1. Hyperparameters for training our MTL network

Hyperparameters	Segmentation network(S)	Classification network(C)	Discriminator network(D)
Learning rate	2e-4	5e-4	1e-4
Optimizer	Adam	Adam	Adam
Epoch	30	20	30
Batch size	8	16	8
μ_1	0.5		
μ_2	0.5		
μ_3	0.2		

Table A2. Comparing localization accuracy and overall processing time from the proposed NLEM method versus STL_NLEM, WAT_NLEM and NA_NLEM methods.

Method	1EDLS(%)	ED(mm)	SLE(°)	TE（mm）	OPT(s)
NLEM	95.63 ± 1.00	0.29 ± 0.02	0.27 ±0.02	0.46 ±0.01	0.0149 ±0.0001
STL_NLEM	92.64 ±2.63 (**)	0.31 ±0.22 (**)	0.33 ±0.02 (**)	0.48 ±0.01 (**)	0.0250 ±0.0003 (***)
WAT_NLEM	86.44 ±7.36 (***)	0.38 ±0.09 (***)	0.35 ±0.05 (***)	0.51 ±0.04 (***)	0.0149 ±0.0001
NA_NLEM	41.47 ± 1.51 (***)	0.67 ±0.03 (***)	0.68 ±0.09 (***)	0.78 ±0.15 (***)	0.0149 ±0.0001

Note：The one tailed paired t-test is used to compare related methods with NLEM. **And * ** in the table represent $p < 0.05$ and $p < 0.00001$, respectively. A value of $p < 0.05$ is considered statistically significant.

Table A3. Comparing classification accuracy from the proposed NLEM method versus STL_NLEM method.

Method	Accuracy (%)	Precision (%)	Sensitivity (%)	Specificity (%)
NLEM	97.37 ±0.71	97.70 ±1.01	99.71 ±0.31	96.82 ±1.25
STL_NLM	94.63 ±3.11 (**)	98.49 ±2.39	93.60 ±4.38 (**)	97.36 ±3.57 (**)

Note：The one tailed paired *t*-test is used to compare STL_NLEM method with NLEM. ** in the table represent $p < 0.05$. A value of $p < 0.05$ is considered statistically significant.

[1] Mohammad H Jafari, Hany Girgis, Nathan Van Woudenberg, Zhibin Liao, Robert Rohling, Ken Gin, Purang Abolmaesumi, Terasa Tsang. Automatic biplane left ventricular ejection fraction estimation with mobile point-of-care ultrasound using multi-task learning and adversarial training. International Joural of Computer Assisted Radiology and Surgery. 2019. doi:10.1007/s11548-019-01954-w

[2] Wu W, Song L, Yang Y, Wang J, Liu H and Zhang L. Exploring the dynamics and interplay of human papillomavirus and cervical tumorigenesis by integrating biological data into a mathematical model. *BMC Bioinformatics*, 2020, 21: 1-8. doi: 10.1186/s12859-020-3454-5