

## 1 Supplementary material

Table 1: Errors per model and per view for the standard view navigation experiment. CRL+B ( $N$  Pat) indicates CRL where CPB samples trajectories from  $N$  patients. CRL+BA uses 2 patients per batch.

Goal type	Method	Views	Angle Error (deg)	Position error (mm)
N/A	RL-TEE	ME 4CH	$9.51 \pm 5.91$	$9.05 \pm 5.54$
		ME 2CH	$8.49 \pm 5.52$	$7.14 \pm 5.19$
		ME AV SAX	$14.65 \pm 11.94$	$13.25 \pm 9.25$
		ME LAX	$6.92 \pm 4.56$	$6.85 \pm 4.51$
N/A	SAC	ME 4CH	$8.02 \pm 4.97$	$8.49 \pm 6.23$
		ME 2CH	$9.34 \pm 7.54$	$8.26 \pm 5.71$
		ME AV SAX	$11.76 \pm 15.28$	$9.28 \pm 12.42$
		ME LAX	$10.87 \pm 11.65$	$10.62 \pm 12.08$
Patient	CRL-D	ME 4CH	$20.14 \pm 17.69$	$13.85 \pm 15.21$
		ME 2CH	$14.55 \pm 19.60$	$11.86 \pm 22.74$
		ME AV SAX	$19.30 \pm 19.39$	$11.57 \pm 11.93$
		ME LAX	$18.25 \pm 17.66$	$13.76 \pm 16.40$
Patient	CRL+B (8 Pat)	ME 4CH	$7.27 \pm 7.00$	$5.95 \pm 4.31$
		ME 2CH	$6.19 \pm 3.59$	$3.53 \pm 2.48$
		ME AV SAX	$13.59 \pm 22.07$	$8.35 \pm 10.56$
		ME LAX	$10.03 \pm 11.31$	$6.91 \pm 8.68$
Patient	CRL+B (2 Pat)	ME 4CH	$7.19 \pm 5.12$	$6.89 \pm 4.81$
		ME 2CH	$7.27 \pm 5.99$	$3.64 \pm 2.86$
		ME AV SAX	$13.46 \pm 22.13$	$7.89 \pm 14.88$
		ME LAX	$9.05 \pm 9.14$	$5.47 \pm 4.41$
Patient	CRL+B (1 Pat)	ME 4CH	$8.98 \pm 6.59$	$8.74 \pm 7.41$
		ME 2CH	$6.45 \pm 4.33$	$3.87 \pm 2.84$
		ME AV SAX	$16.86 \pm 23.92$	$9.80 \pm 12.17$
		ME LAX	$13.32 \pm 11.79$	$9.59 \pm 7.88$
Patient	CRL+BA (K=1)	ME 4CH	$8.47 \pm 5.37$	$9.08 \pm 4.73$
		ME 2CH	$6.86 \pm 4.00$	$4.10 \pm 3.07$
		ME AV SAX	$16.43 \pm 20.92$	$10.75 \pm 12.39$
		ME LAX	$12.10 \pm 8.95$	$9.14 \pm 5.80$
Patient	CRL+BA (K=2)	ME 4CH	$6.22 \pm 4.72$	$5.86 \pm 4.11$
		ME 2CH	$7.73 \pm 4.61$	$5.57 \pm 2.96$
		ME AV SAX	$15.35 \pm 16.16$	$7.98 \pm 11.34$
		ME LAX	$9.52 \pm 5.37$	$7.64 \pm 4.19$
Patient	CRL+BA (K=4)	ME 4CH	$9.02 \pm 9.50$	$7.91 \pm 5.90$
		ME 2CH	$6.67 \pm 5.72$	$4.31 \pm 2.49$
		ME AV SAX	$15.05 \pm 19.57$	$8.75 \pm 11.79$
		ME LAX	$13.36 \pm 8.99$	$10.19 \pm 6.39$

Table 2: Perturbation ranges. We sample  $P_N$  perturbations and for each  $P_i$ , we randomly choose a DOF and sample the magnitude of the perturbation using the min/max values indicated below.

Name	Value
Num perturbations sampled $P_N$	8
Translation range (min/max) mm	15/20
In-plane rotation range (min/max) degrees	10/15
Transducer rotation range (min/max) degrees	10/15
Flexions (ante/retro & left/right) range (min/max) degrees	1/5

Table 3: CRL Hyperparameters

Name	Value
Actor & Critic lr	3e-4 for 100M steps then 1e-5 until 200M steps
L2 Norm on SA embeddings	0.01
Logsumexp regularization on $Q_M$	0.01
Temperature scaling	Learnt
Num augmentations $K$	2
Latent dim $H$	128
Batch size	512
Replay buffer size	500K
Episode length	20 steps
Observations size	(96, 96)
Translation magnitude max	5 mm
In-plane rotation magnitude max	10 degrees
Transducer rotation magnitude max	10 degrees
Flexion magnitude max	1 degree