

GTCaR: Graph Transformer for Camera Re-localization

Supplementary Material

Xinyi Li¹ and Haibin Ling²

¹ Magic Leap, Sunnyvale, CA, USA xinli@magicleap.com

² Stony Brook University, Stony Brook, NY, USA hling@cs.stonybrook.edu

1 Ablation Study

In this section we provide the ablation experiment results on the 7-Scenes Dataset [?], the Cambridge Dataset [21] and the Oxford Robotcar Dataset [26], respectively. As introduced in the main paper, in all the ablation studies, we use GTCaR[*removed-module*] to denote the network without the specific module. **Ablations on the 7-Scenes Dataset.**

Table 1: Ablations on the 7-Scenes Dataset.

Scene Scene scale	Chess 3 x 2m ²	Fire 2.5 x 1m ²	Heads 2 x 0.5m ²	Office 2.5 x 2m ²	Pumpkin 2.5 x 2m ²	Kitchen 4 x 3m ²	Stairs 2.5 x 2m ²	Avg.
GTCaR	0.09m, 1.94°	0.27m, 8.45°	0.14m, 9.34°	0.12m, 2.41°	0.15m, 2.13°	0.21m, 2.73°	0.26m, 8.92°	0.18m, 5.13°
GTCaR[temporal]	0.22m, 5.16°	0.29m, 10.02°	0.21m, 11.24°	0.31m, 5.29°	0.31m, 3.26°	0.32m, 6.43°	0.33m, 13.11°	0.28m, 7.79°
GTCaR[graph]	0.12m, 2.99°	0.31m, 9.36°	0.17m, 10.57°	0.15m, 4.85°	0.17m, 2.28°	0.21m, 4.95°	0.32m, 10.22°	0.21m, 6.46°
GTCaR[MPNN]	0.14m, 5.21°	0.26m, 11.24°	0.19m, 12.20°	0.15m, 6.48°	0.20m, 2.13°	0.18m, 4.95°	0.29m, 8.87°	0.20m, 7.30°

Ablations on the Cambridge Dataset.

Table 2: Ablations on the Cambridge Dataset.

Scene Scene scale	College 5.6x10 ³ m ²	Shop 8.8x10 ³ m ²	Church 4.8x10 ³ m ²	Hospital 2.0x10 ³ m ²	Court 8.0x10 ³ m ²	Street 5.0x10 ³ m ²	Avg.
GTCaR	0.42m, 0.52°	0.64m, 1.56°	1.55m, 2.56°	1.32m, 1.97°	5.62m, 2.17°	10.27m, 19.88°	3.30m, 4.78°
GTCaR[temporal]	0.69m, 1.22°	0.82m, 2.03°	1.72m, 3.51°	1.66m, 2.49°	6.95m, 2.95°	15.76m, 22.63°	4.6m, 5.81°
GTCaR[graph]	0.45m, 0.98°	0.93m, 5.24°	2.02m, 4.55°	2.14m, 3.07°	7.34m, 4.84°	25.67m, 30.85°	6.43m, 8.26°
GTCaR[MPNN]	0.48m, 0.66°	0.69m, 1.54°	1.53m, 3.02°	1.65m, 1.97°	5.98m, 2.38°	10.15m, 24.39°	3.41m, 5.66°

Ablations on the RobotCar Dataset.

Table 3: Ablations on the RobotCar Dataset.

Scene	LOOP	FULL
Scene scale	1120m	9562m
GTCaR	5.46m, 1.98°	14.37m, 3.68°
GTCaR[temporal]	8.25m, 3.48°	42.87m, 9.47°
GTCaR[graph]	9.03m, 3.55°	32.16m, 8.29°
GTCaR[MPNN]	7.95m, 2.64°	22.59m, 6.61°

2 Pose Error Distribution

In this section we report the cumulative distributions of the translation and rotation errors on the LOOP and FULL dataset from the Oxford RobotCar Dataset. The performance is compared against CNN+GNN [47], LsG [46], MapNet [4] and PoseNet [21] in Fig. 1.

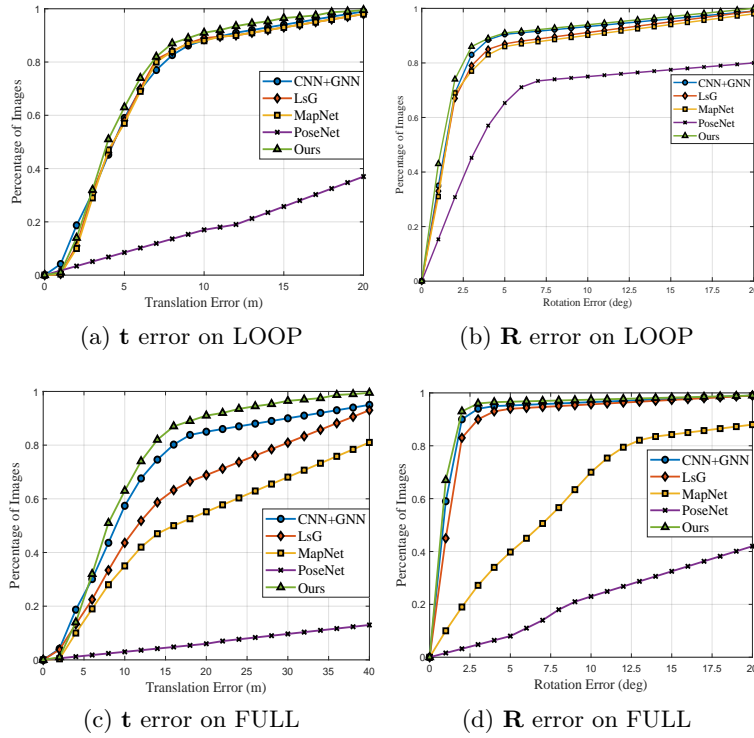


Fig. 1: Cumulative distributions of the translations errors and rotation errors on the two datasets. The x-axis represents the error and the y-axis denotes the percentage of image frames with error lower than the x-value.