

Camera Relocalization by Computing Pairwise Relative Poses Using Convolutional Neural Network - Supplementary Material -

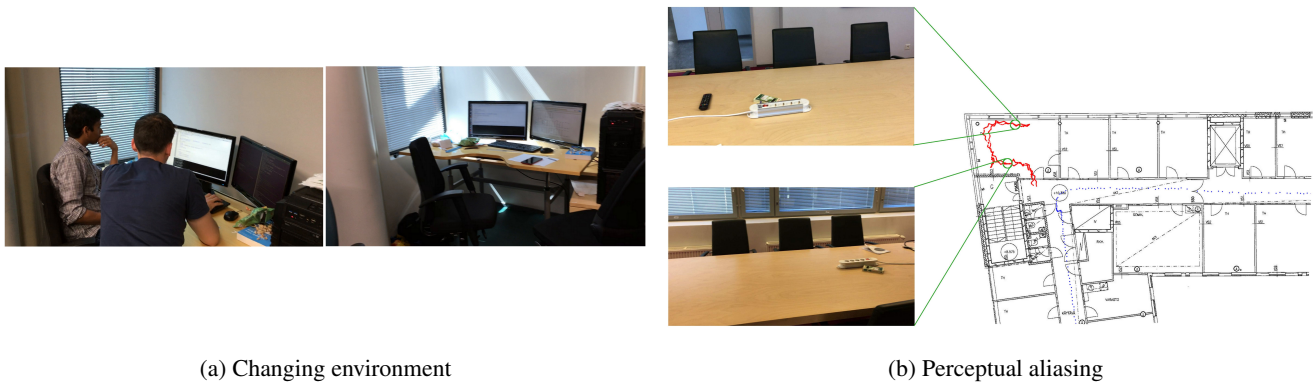


Figure 4: Some challenging cases in *University* dataset.

Scene	Viewpoint 1	Viewpoint 2	Viewpoint 4	Viewpoint 5	Viewpoint 6
Chess	0.17m, 7.50°	0.16m, 7.30°	0.17m, 7.36°	0.17m, 7.30°	0.16m, 7.59°
Fire	0.10m, 6.37°	0.11m, 6.29°	0.10m, 6.44°	0.10m, 6.32°	0.11m, 6.44°
Heads	0.24m, 8.50°	0.25m, 8.82°	0.24m, 8.58°	0.24m, 8.46°	0.25m, 8.83°
Office	0.19m, 11.13°	0.19m, 11.16°	0.19m, 11.11°	0.20m, 11.13°	0.19m, 11.06°
Pumpkin	0.26m, 8.95°	0.26m, 9.27°	0.26m, 9.45°	0.26m, 9.31°	0.25m, 9.31°
Red Kitchen	0.20m, 7.40°	0.20m, 7.45°	0.19m, 7.48°	0.19m, 7.48°	0.19m, 7.47°
Stairs	0.22m, 8.80°	0.23m, 8.15°	0.23m, 8.36°	0.23m, 8.30°	0.23m, 8.37°
Average	0.21m, 8.38°	0.20m, 8.35°	0.20m, 8.40°	0.20m, 8.33°	0.20m, 8.44°

Table 6: Camera relocalization accuracy of the proposed system for different viewpoint changes between the query and the database image. The explanation of **Viewpoint N** notation is provided Section 5.3

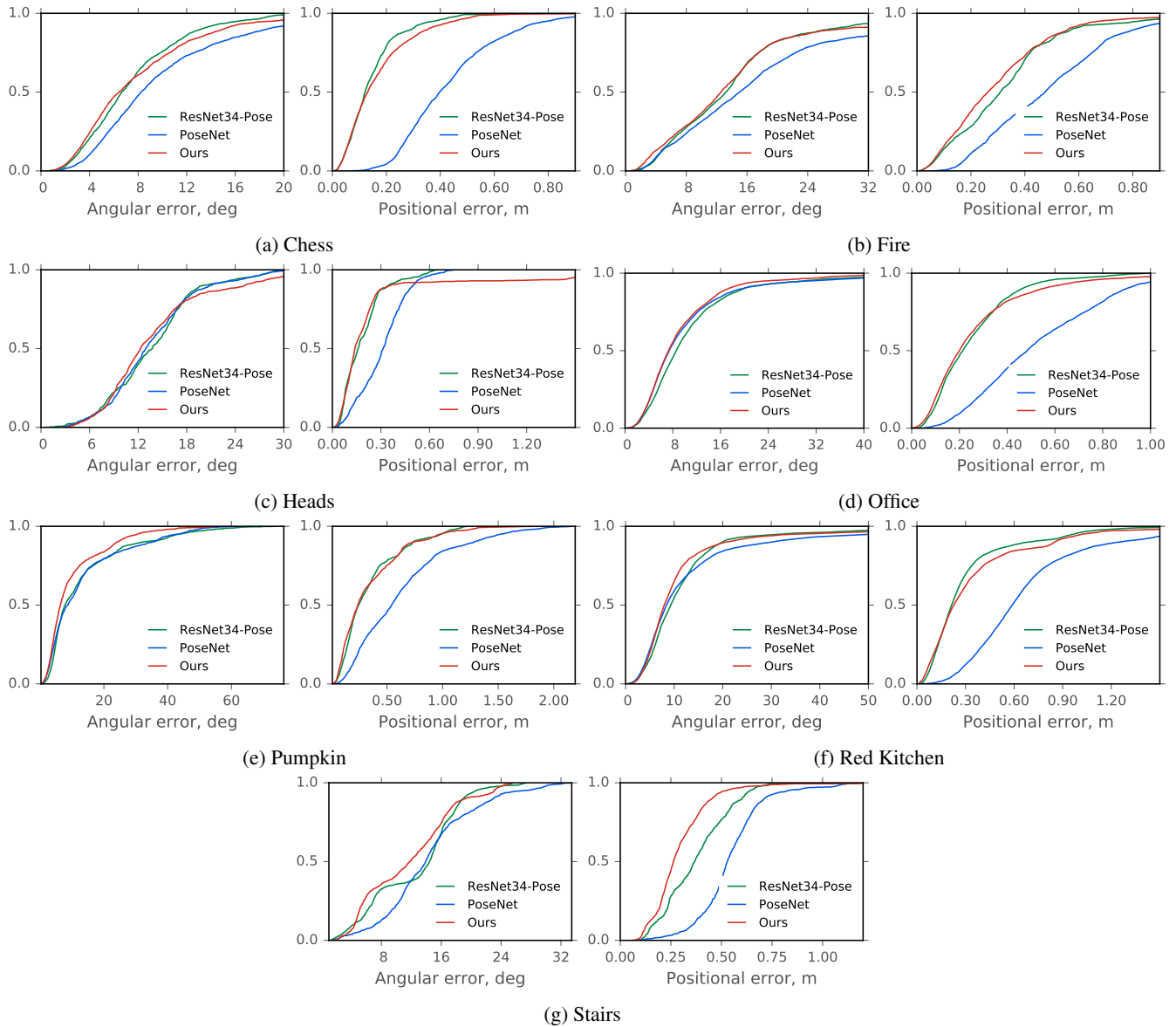


Figure 5: Localization performance of the proposed approach, the baseline model (ResNet34-Pose), and PoseNet presented as normalized cumulative error histograms for all scenes of 7Scenes dataset. It should be noted that the baseline is trained in a scene-specific manner unlike ours. That is, our approach uses the same network for all 7 scenes whereas PoseNet and ResNet34-Pose have a separate network for each scene.