GSNet: Joint Vehicle Pose and Shape Reconstruction with Geometrical and Scene-aware Supervision (Supplementary material)

Lei Ke[1], Shichao Li[1], Yanan Sun[1], Yu-Wing Tai[1,2], and Chi-Keung Tang[1]

[1] The Hong Kong University of Science and Technology
[2] Kwai Inc.

We provide more quantitative analysis and implementation details for GSNet in this document.

Implementation details on 2D/3D keypoints mapping. To build the mapping relation between 2D and 3D semantic keypoints, we counted the hitting times by projecting each vertex on the mesh to image using ground truth pose value to find its nearest neighboring 2D points. Due to the small variance in different car shapes and compact vertex distribution, the same 2D semantic keypoints may be mapped to multiple neighboring vertices on the mesh. For handling the small variance in different car shapes, we adopted weighted voting to represent the importance of each vertex in the neighboring set by its selected times and chose the top 5 vertices.

Influence of the number of keypoints. We study the influence of the number of keypoints on the performance of GSNet by reducing the 66 semantic keypoints for cars defined in [1] to 50 and 33 respectively, and compare the results in Table 1. The minor degradation with only half of the complete keypoints shows that GSNet is still robust when adopting a more sparse keypoint definition.

Shape reconstruction error across vehicle categories. To analyze which car types are more prone to reconstruction errors, we present the shape reconstruction error distribution as shown in Figure 1. Note that only 34 car categories appear in the validation set, where the car types *biaozhi*-508 and *qiya* have the largest shape reconstruction errors.

Table 1. Ablation study on using different numbers of keypoints on GSNet, where we randomly pick 50% and 75% keypoints from the total 66 semantic keypoints.

Number of Keypoints	Rel-mAP	Abs-mAP
33 (using 50%)	19.83	18.61
50 (using 75%)	20.08	18.79
66 (using 100%)	**20.21**	**18.91**

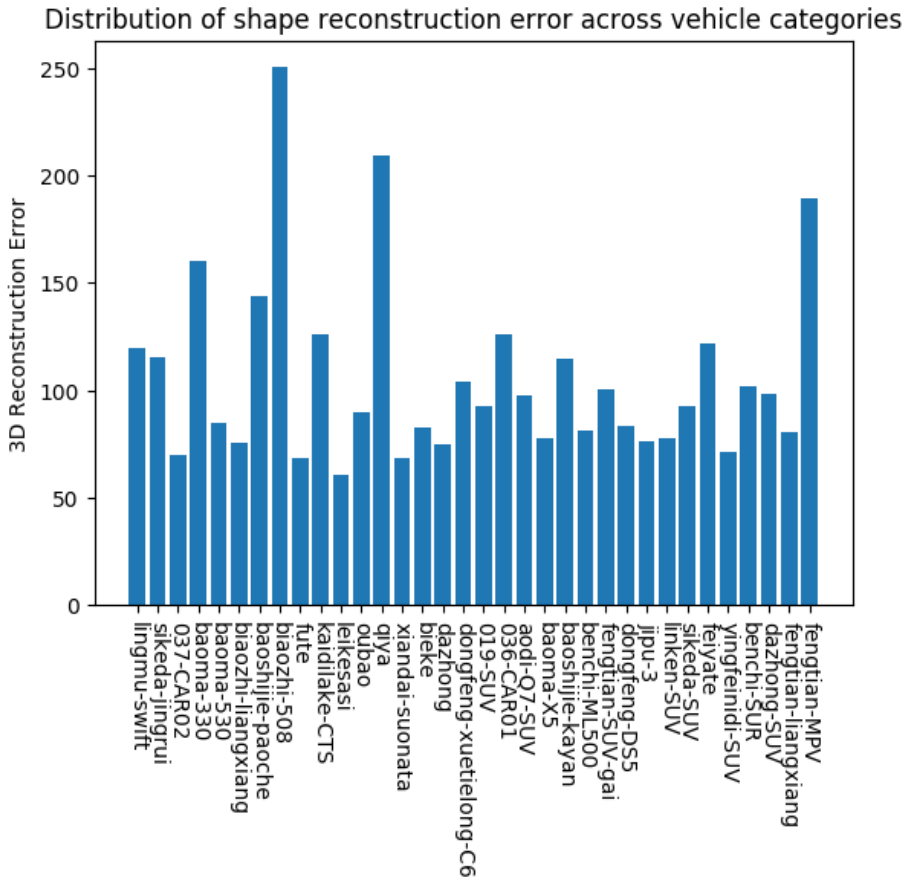

Fig. 1. 3D shape reconstruction error across 34 vehicle categories.

References

1. Song, X., Wang, P., Zhou, D., Zhu, R., Guan, C., Dai, Y., Su, H., Li, H., Yang, R.: Apollocar3d: A large 3d car instance understanding benchmark for autonomous driving. In: CVPR (2019)