Supplementary Material for "Transductive Zero-Shot Learning for 3D Point Cloud Classification"

In this supplementary material, we further assess our proposed method with additional quantitative and qualitative evaluations. In the quantitative evaluation section, we evaluate **(1)** the effect of the batch size on 3D Zero-Shot Learning (ZSL) using ModelNet10, **(2)** the effect of using a different point cloud architecture, EdgeConv [12], and **(3)** the effect of using the experimental protocol for Generalized Zero-Shot Learning (GZSL) proposed by Song *et al.* [8]. In the qualitative evaluation section, we show success and failure cases on unseen classes from ModelNet10.

1. Additional Quantitative Evaluation

1.1. Batch Size

In this experiment, we evaluate the effect of the batch size on the accuracy of our proposed method for the 3D ModelNet10 dataset. As can be seen in Figure 1, the size of the batch has a significant impact on the performance, with the best performance on this dataset being achieved at a batch size of 32.

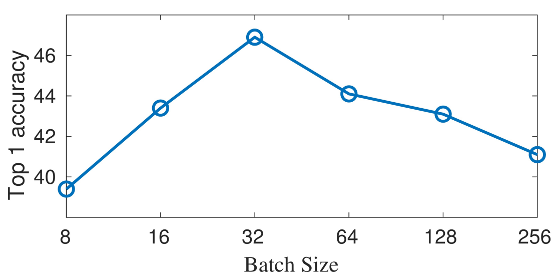

Figure 1. Top-1 accuracy on the ModelNet10 dataset as the batch size varies.

1.2. Point Cloud Architecture

In this paper, we used PointNet [4] as the backbone point cloud architecture in our 3D experiments. However, while PointNet is one of the first works that has been proposed for point cloud classification using deep learning, there are many other methods [4, 5, 12, 2, 15, 11, 14, 6, 9, 1] which were introduced later and tend to achieve better performance for supervised 3D point cloud classification. Here, we compare PointNet with EdgeConv [12] to study the effect of using a more advanced point cloud architecture for the task of 3D ZSL classification. In supervised 3D point cloud classification, EdgeConv achieves 92.2% accuracy on ModelNet40 while PointNet achieves 89.2%. In this additional experiment, we use ModelNet10 as the unseen set to compare those two methods. As shown in Table 1, both PointNet and EdgeConv achieve similar performance. We would expect to see some improvement when using EdgeConv since it works better in the case of supervised classification. In Figure 2, it can be seen however that both PointNet and EdgeConv cluster unseen point cloud features similarly and imperfectly. This again shows the difficulty of the ZSL task on 3D data where there are a lack of good pretrained models.

Method	ModelNet10	McGill	SHREC2015
PointNet [4]	**46.9**	**21.7**	**13.0**
EdgeConv [12]	45.2	20.6	**13.0**

Table 1. ZSL results on the 3D ModelNet10 [13], McGill [7], and SHREC2015 [3] datasets using different point cloud architecture, PointNet and EdgeConv.

1.3. QFSL's Generalized ZSL Evaluation Protocol

In this experiment, we evaluate the effect of using a different evaluation protocol for the GZSL experiments, as proposed by Song *et al.* [8]. Under this protocol, the unlabeled data, which consists of seen and unseen instances, is divided into halves, and two models are trained. In each model, half of unlabeled data is used for training and the other half for testing. The final performance is calculated by averaging the performance of these two models. The authors suggest that this allows for fairer evaluation, although it is an imperfect solution. Nonetheless, we show in Table 2 for the ModelNet10 dataset that our method performs better than QFSL with respect to all accuracy measures under both this protocol and the original protocol from our paper. In fact, both methods perform better under this different protocol, which suggests that splitting the unlabeled data in this way makes the task easier. As a result, we use our more conservative GZSL evaluation protocol in the main paper.

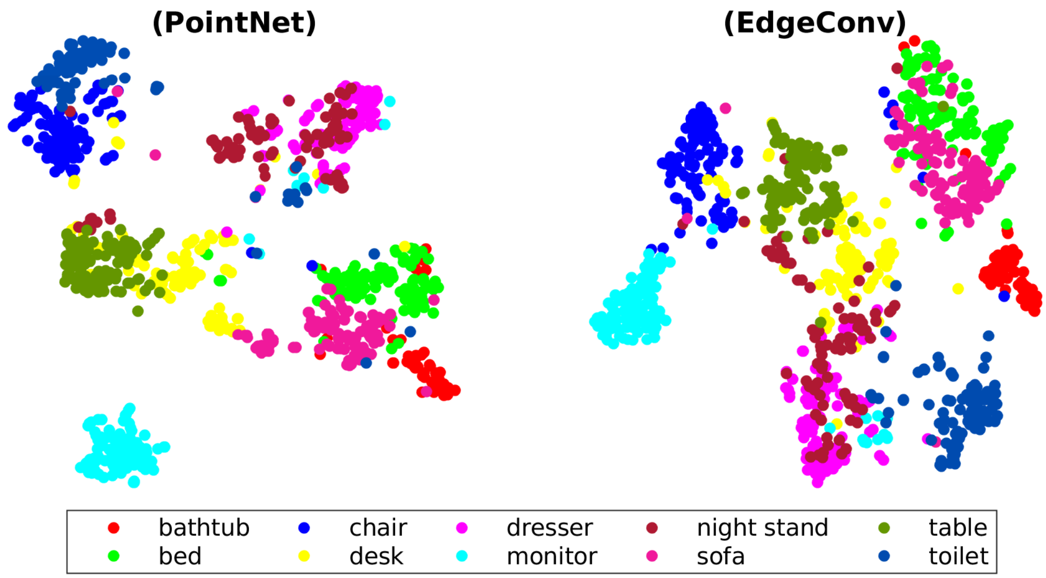

Figure 2. 2D tSNE [10] visualization of unseen point cloud feature vectors (circles) based on (a) PointNet (b) EdgeConv on ModelNet10. The unseen point cloud features are clustered similarly in both PointNet and EdgeConv, despite EdgeConv performing better than PointNet on the task of supervised point cloud classification.

Method	Acc_s	Acc_u	HM
QFSL [8]	58.1 / 68.2	21.8 / 24.3	31.7 / 35.6
Ours	**74.6** / 72.0	23.4 / **29.2**	35.6 / **41.5**

Table 2. GZSL results on the 3D ModelNet10 dataset [13] under evaluation protocols (A) / (B), where (A) is the evaluation protocol from our paper and (B) is the protocol proposed by Song *et al.* [8]. We report the top-1 accuracy (%) on seen classes (Acc_s) and unseen classes (Acc_u) for each method, as well as the harmonic mean (HM) of both measures.

2. Qualitative Evaluation

In this section, we visualize five unseen classes from the ModelNet10 dataset with examples where our method correctly classified the point cloud, shown in Figure 3, and examples where it incorrectly classified the point cloud, shown in Figure 4. The network appears to be providing incorrect predictions for mostly hard examples, those that are quite different from standard examples in that class, or where the classes overlap in their geometry, such as dresser and night stand.

References

[1] A. Cheraghian and L. Petersson. 3dcapsule: Extending the capsule architecture to classify 3d point clouds. In *2019 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pages 1194–1202, Jan 2019. 1

[2] J. Li, B. M. Chen, and G. H. Lee. So-net: Self-organizing network for point cloud analysis. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 9397–9406, 2018. 1

[3] Z. Lian, J. Zhang, S. Choi, H. ElNaghy, J. El-Sana, T. Furuya, A. Giachetti, R. A. Guler, L. Lai, C. Li, H. Li, F. A. Limberger, R. Martin, R. U. Nakanishi, A. P. Neto, L. G. Nonato, R. Ohbuchi, K. Pevzner, D. Pickup, P. Rosin, A. Sharf, L. Sun, X. Sun, S. Tari, G. Unal, and R. C. Wilson. Non-rigid 3D Shape Retrieval. In I. Pratikakis, M. Spagnuolo, T. Theoharis, L. V. Gool, and R. Veltkamp, editors, *Eurographics Workshop on 3D Object Retrieval*. The Eurographics Association, 2015. 1

[4] C. R. Qi, H. Su, K. Mo, and L. J. Guibas. Pointnet: Deep learning on point sets for 3d classification and segmentation. *Proc. Computer Vision and Pattern Recognition (CVPR), IEEE*, 1(2):4, 2017. 1

[5] C. R. Qi, L. Yi, H. Su, and L. J. Guibas. Pointnet++: Deep hierarchical feature learning on point sets in a metric space. In *Advances in Neural Information Processing Systems*, pages 5099–5108, 2017. 1

[6] Y. Shen, C. Feng, Y. Yang, and D. Tian. Neighbors do help: Deeply exploiting local structures of point clouds. June 2018. 1

[7] K. Siddiqi, J. Zhang, D. Macrini, A. Shokoufandeh, S. Bouix, and S. Dickinson. Retrieving articulated 3-d models using medial surfaces. *Mach. Vision Appl.*, 19(4):261–275, May 2008. 1

[8] J. Song, C. Shen, Y. Yang, Y. P. Liu, and M. Song. Transductive unbiased embedding for zero-shot learning. *2018*

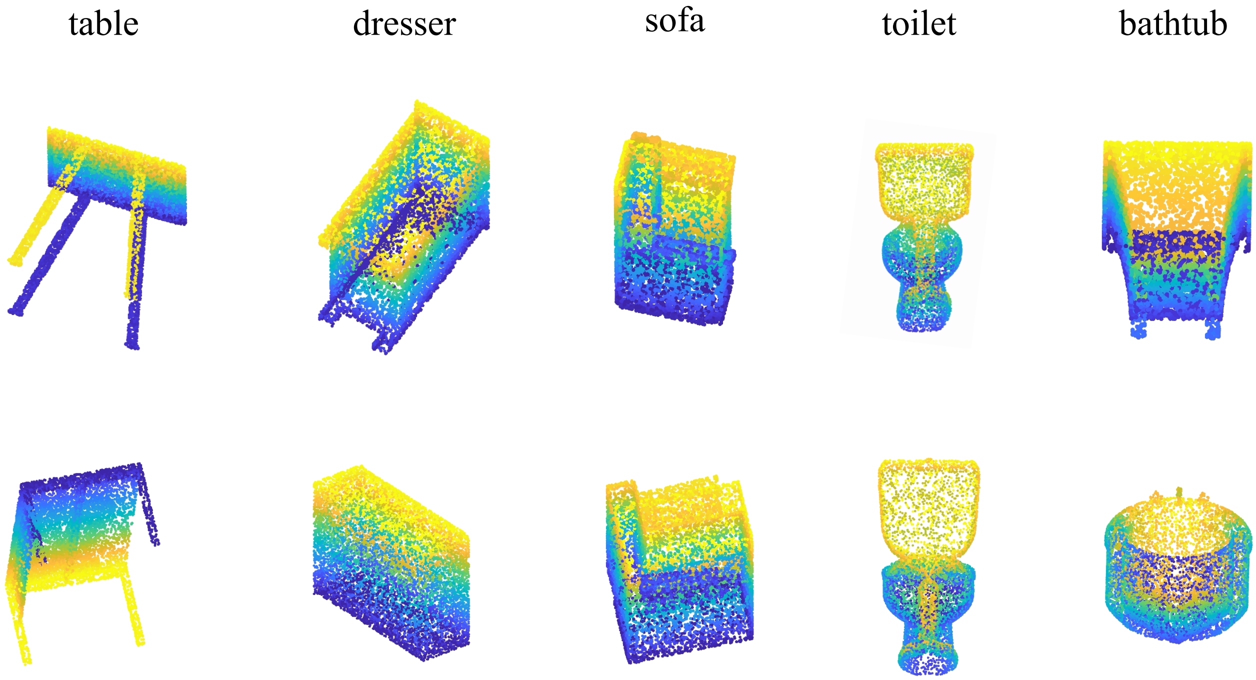

Figure 3. Visualization of five classes from the ModelNet10 dataset with examples of correctly classified point clouds.

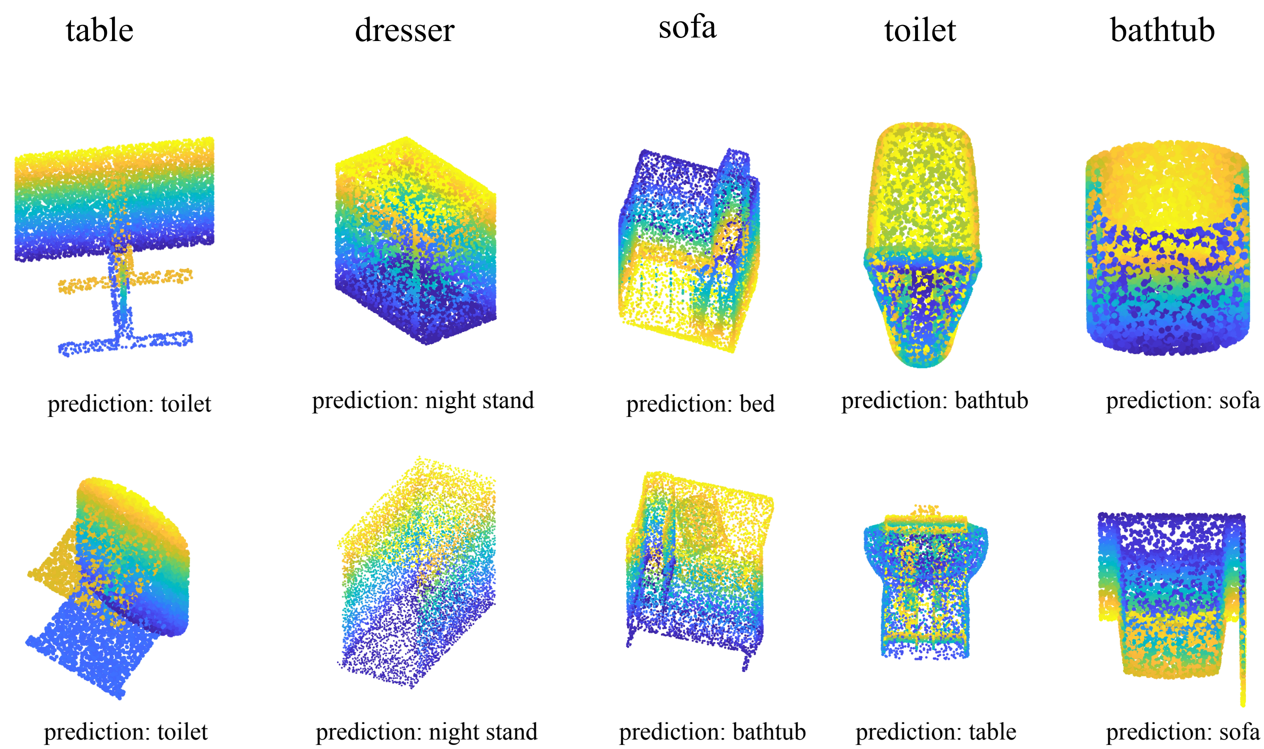

Figure 4. Visualization of five classes from the ModelNet10 dataset with examples of incorrectly classified point clouds. The predicted classes are shown below each model.

IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 1024–1033, 2018. 1, 2

[9] H. Su, V. Jampani, D. Sun, S. Maji, E. Kalogerakis, M.-H. Yang, and J. Kautz. SPLATNet: Sparse lattice networks for point cloud processing. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 2530–2539, 2018. 1

[10] L. Van Der Maaten. Accelerating t-sne using tree-based algorithms. *Journal of machine learning research*, 15(1):3221–3245, 2014. 2

[11] C. Wang, B. Samari, and K. Siddiqi. Local spectral graph convolution for point set feature learning. *arXiv preprint arXiv:1803.05827*, 2018. 1

[12] Y. Wang, Y. Sun, Z. Liu, S. E. Sarma, M. M. Bronstein, and J. M. Solomon. Dynamic graph cnn for learning on point clouds. *arXiv preprint arXiv:1801.07829*, 2018. 1

[13] Z. Wu, S. Song, A. Khosla, F. Yu, L. Zhang, X. Tang, and J. Xiao. 3d shapenets: A deep representation for volumetric shapes. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1912–1920, 2015. 1, 2

[14] S. Xie, S. Liu, Z. Chen, and Z. Tu. Attentional shapecontextnet for point cloud recognition. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2018. 1

[15] Y. Xu, T. Fan, M. Xu, L. Zeng, and Y. Qiao. Spidercnn: Deep learning on point sets with parameterized convolutional filters. *arXiv preprint arXiv:1803.11527*, 2018. 1