Supplementary Material: Label-Efficient Learning on Point Clouds using Approximate Convex Decompositions

Anonymous ECCV submission

Paper ID 1148

Abstract. We include the following additional results and discussions supporting the material presented in the main submission:
- Discussion regarding other decomposition methods (Sec. 1)
- Visualizations of ACD pre-trained features using clustering (Sec. 2)
- Per-category results for baseline and self-supervised models on few-shot part segmentation (Sec. 3)

1 On shape decomposition methods

In our current work, we have focused on ACD, which approximates shapes as a collection of (approximately) *convex* parts. There are however various other decomposition methods that might correspond better to notions of semantic parts, and yield a better self-supervisory signal. The current section gives a brief overview of other shape decomposition methods and connections to classical works in early computer vision –

We refer to *shape decomposition* as a procedure that computes a set of simpler proxy-shapes from complex ones. Such procedures have been extensively studied in the geometric processing literature and they operate without resorting to any learning-based technique or human annotated labels – multiple methods have shown that complex shapes can be reliably decomposed relying solely on geometric cues. Biederman's recognition-by-components theory [1] attempts to explain object recognition in humans by the ability to assemble basic shapes such as cylinders and cones, called *geons*, into the complex objects encountered in the visual world. Early work in cognitive science [6] shows that humans are likely to decompose a 3D shape along regions of maximum concavity, resulting in parts that tend to be convex, often referred to as the "*minima rule*". Classical approaches in computer vision have modeled three-dimensional shapes as a composition of simpler primitives, *e.g.* work by Binford [2, 3] and Marr [7]. More recent work in geometric processing has developed shape decomposition techniques that generate different types of primitives which are amenable to tasks like editing, grasping, tracking and animation. Those have explored primitives like 3D curves [4,5,8], cages [11], sphere-meshes [10] and generalized cylinders [12]. Given the promising results from ACD in our experiments, exploring other decomposition methods is a reasonable direction for future research in self-supervised learning on point clouds.

Table 1. Per-category few-shot segmentation on the ShapeNet dataset.

K-shot	Method	mIoU	Aero	Bag	Cap	Car	Chair	Earph	Guitar	Knife	Lamp	Laptop	Motor	Mug	Pistol	Rocket	Skate	Table
1	Baseline	53.14	48.89	44.85	50.42	23.94	69.50	**63.57**	78.31	**67.34**	**50.84**	87.00	19.69	48.60	54.33	33.84	43.72	**65.45**
	ACD	**62.37**	**66.74**	**57.30**	**76.04**	**41.10**	**77.61**	61.16	**86.00**	60.91	46.99	**92.77**	**23.58**	**68.28**	**70.23**	**36.03**	**67.78**	65.36
3	Baseline	58.15	61.59	**60.86**	76.35	32.16	79.20	30.36	86.21	73.15	**54.03**	**93.59**	25.59	48.62	57.15	41.50	43.74	**66.31**
	ACD	**69.56**	**62.49**	52.11	**78.79**	**59.88**	**81.64**	**66.89**	**87.35**	**76.08**	52.57	93.32	**50.79**	**92.94**	**76.29**	**48.84**	**67.88**	65.15
5	Baseline	68.01	62.24	**75.12**	**80.68**	**57.64**	81.34	**63.29**	**86.10**	**76.05**	48.34	94.43	27.41	91.39	74.95	38.09	62.68	68.48
	ACD	**72.66**	**68.63**	68.54	80.03	54.53	**83.03**	61.31	85.75	73.62	**64.76**	**94.34**	**62.87**	**92.52**	**75.72**	**56.69**	**71.17**	**69.09**
10	Baseline	71.37	63.09	74.55	75.98	65.37	82.55	**62.87**	86.34	**80.73**	66.03	94.46	42.40	**92.77**	75.60	46.50	65.66	67.01
	ACD	**74.50**	**67.01**	**76.71**	**80.05**	**68.51**	**82.84**	54.16	**88.81**	73.36	**68.29**	**95.16**	**65.83**	92.59	**77.99**	**57.60**	**66.85**	**76.19**

2 Visualizing point features trained on ACD

We visualize the per-point features learned while training the network to perform ACD using contrastive loss. Refer to Section 3 of the main paper for a detailed explanation. In this experiment, we train a PointNet++ [9] backbone by minimizing the loss in Equation 4. After training, we compute per-point embeddings by concatenating features from the first two set aggregation layers (`SA1` and `SA2`), resulting in a per-point feature of dimensionality 384. Then, for every point cloud, we cluster its points according to those embeddings using Mean Shift. Results are presented in Figure 1. As we can see, the network learns to decompose the shapes in approximately convex parts reasonably well.

3 Per-category results for few-shot segmentation

Part segmentation performance across ShapeNet categories is listed in Table 1. We can see that performance varies widely across categories, indicating some are significantly harder to segment than others. In all cases, for mean IoU averaged across shape categories, including the ACD loss over additional unlabeled data provides a gain in performance over a fully-supervised baseline PointNet++ model. We see consistent gains in performance in categories like *Aeroplane*, *Laptop*, *Motorcycle*, *Pistol*, *Rocket* and *Skateboard*.

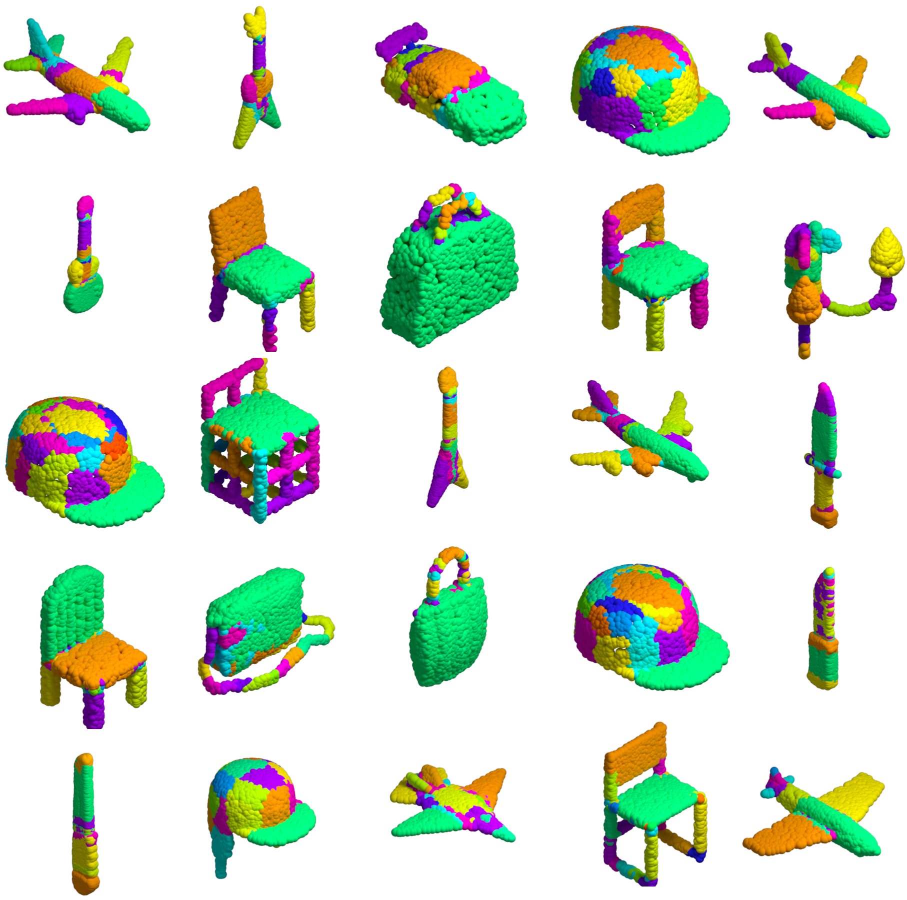

Fig. 1. Visualizing per-point embeddings trained using ACD. For each shape, different colors represent different clusters computed using Mean Shift on the point embeddings from PointNet++ trained to perform ACD.

References

1. Biederman, I.: Recognition-by-components: a theory of human image understanding. Psychological review **94**(2), 115 (1987) 1
2. Binford, I.: Visual perception by computer. In: IEEE Conference of Systems and Control (1971) 1
3. Binford, T.O., Levitt, T.S., Mann, W.B.: Bayesian inference in model-based machine vision. In: Proceedings of the Third Conference on Uncertainty in Artificial Intelligence. pp. 86–97 (1987) 1
4. Gal, R., Sorkine, O., Mitra, N.J., Cohen-Or, D.: iwires: An analyze-and-edit approach to shape manipulation. ACM Transactions on Graphics (Siggraph) **28**(3), #33, 1–10 (2009) 1
5. Gori, G., Sheffer, A., Vining, N., Rosales, E., Carr, N., Ju, T.: Flowrep: Descriptive curve networks for free-form design shapes. ACM Transaction on Graphics **36**(4) (2017). https://doi.org/http://dx.doi.org/10.1145/3072959.3073639 1
6. Hoffman, D.D., Richards, W.: Parts of recognition (1983) 1
7. Marr, D., Nishihara, H.K.: Representation and recognition of the spatial organization of three-dimensional shapes. Proceedings of the Royal Society of London. Series B. Biological Sciences **200**(1140), 269–294 (1978) 1
8. Mehra, R., Zhou, Q., Long, J., Sheffer, A., Gooch, A., Mitra, N.J.: Abstraction of man-made shapes. ACM Transactions on Graphics **28**(5), #137, 1–10 (2009) 1
9. Qi, C.R., Yi, L., Su, H., Guibas, L.: PointNet++: Deep hierarchical feature learning on point sets in a metric space. In: Proc. NIPS (2017) 2
10. Thiery, J.M., Guy, E., Boubekeur, T.: Sphere-meshes: Shape approximation using spherical quadric error metrics. ACM Transaction on Graphics (Proc. SIGGRAPH Asia 2013) **32**(6), Art. No. 178 (2013) 1
11. Xian, C., Lin, H., Gao, S.: Automatic cage generation by improved obbs for mesh deformation. The Visual Computer **28**(1), 21–33 (2012) 1
12. Zhou, Y., Yin, K., Huang, H., Zhang, H., Gong, M., Cohen-Or, D.: Generalized cylinder decomposition. ACM Trans. Graph. **34**(6) (2015) 1