Improving Graph Representation for Point Cloud Segmentation via Attentive Filtering

Supplementary Material

1. Implementation Details

GCN Block and GAF Block. We construct the GCN block and the GAF block by combining a GCN/GAF and a separated MLP. Besides, we adopt residual connections in the GCN/GAF blocks. For a better understanding, we give the formulation of our Graph Attentive Filter again and give a detailed description as follows:

$$\mathbf{F}^{l+1} = \sigma(\mathrm{FFN}(\mathcal{R} \cdot \hat{\mathbf{A}})) \diamond \Phi(\mathbf{F}^{l}), \tag{1}$$

where FFN is a feed-forward network to further learn the relation, $\cdot$ represents element-wise multiplication and the symmetric adjacency matrix $\hat{\mathbf{A}}$ is repeated to fit the correlation matrix $\mathcal{R} \in \mathbb{R}^{N \times N \times C}$, where C is the dimension of the point features $\mathbf{F}^{l}$. Φ represents the non-linear feature transformation and σ is a normalization operator. $\diamond$ is the graph aggregation operation. Our GAF estimate channel-wise edge relations $\tilde{\mathcal{R}} = \sigma(\mathrm{FFN}(\mathcal{R} \cdot \hat{\mathbf{A}}))$, and we give a detailed formulation of the channel-wise graph aggregation as follows:

$$\mathbf{F}_{ij}^{l+1} = \underset{k}{\mathrm{Pooling}}(\tilde{\mathcal{R}}_{ikj}\, \Phi(\mathbf{F}^{l})_{kj}). \tag{2}$$

In practice, we use the softmax function as normalization. For computational efficiency, the channel-wise estimated correlations are shared for every 8 channels.

Network Architectures. Figure 1 illustrates the detailed design of our segmentation heads for part segmentation and semantic segmentation. The part segmentation head is constructed following CurveNet [40].

2. More Experimental Results

We report the results on Toronto-3D [33] in table 1. Toronto-3D is a large-scale dataset for outdoor scene segmentation, which covers about 1KM of urban roadways with 8 categories. 78.3 million points are scanned by mobile LiDAR systems. Compared with the indoor datasets, Toronto-3D contains more noise. We split Toronto-3D into 5m×5m blocks and sampled 2048 points from each block following previous works [19,20]. We normalize each point cloud into the unit block, and the initial radius r is set to 0.05m. The sampling rate in each stage is set to 2.

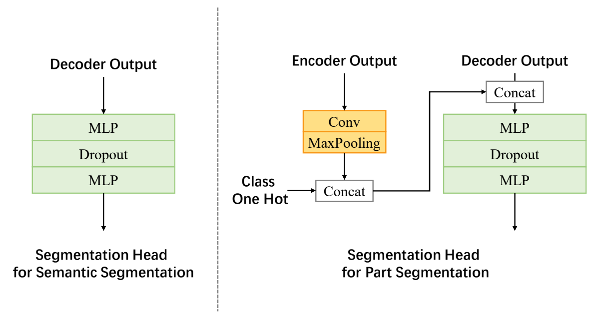

Figure 1. Architectures of segmentation heads.

We report per-class results and compare more methods on S3DIS [1] in Table 3 and Table 4. Our methods outperform previous methods in both Area-5 and 6-fold cross-validation. Compared with the previous SOTA Stratified Transformer [14], our AF-GCN is 22.2% faster in inference and 10× faster in training. We also report per-class results on ScanNetV2 [5]. Compared with the S3DIS dataset, points in ScanNetV2 are relatively sparse and the voxel-based methods usually obtain better performance. As shown in Table 5, our method outperforms recent point-based methods. We report per-class results on ShapeNet-Part [43] in Table 2. Our method outperforms others both in category mIoU and instance mIoU. More visualization is shown in Figure 2.

We also conduct experiments for object classification on ScanObjectNN [37]. We reported the best performance we obtained (OA:88.2, mAcc:86.2), with no significant improvement from baselines. Since the classification task does not require a decoder, and the input scale is so small (1024) that the global features are less affected by ”distant” neighbors, we did not discuss it in the main text.

3. Limitation and Future Work

Limitation. Our method obtains competitive performance in multiple point cloud segmentation datasets. However, compared with the voxel-based methods or the point-based methods using voxel-like processing techniques, our method obtains a relatively lower performance in deal-

Methods (time order)	OA	mIoU	Road	Rd mrk.	Natural	Building	Util. line	Pole	Car	Fence
PointNet++ [30]	92.56	59.47	92.90	0.00	86.13	82.15	60.96	62.81	76.41	14.43
DGCNN [38]	94.24	61.79	93.88	0.00	91.25	80.39	62.40	62.32	88.26	15.81
MS-PCNN [27]	90.03	65.89	93.84	3.83	93.46	82.59	67.80	71.95	91.12	22.50
TGNet [19]	94.08	61.34	93.54	0.00	90.83	81.57	65.26	62.98	88.73	7.85
KPConv [36]	95.39	69.11	94.62	0.06	**96.07**	91.51	**87.68**	81.56	85.66	15.72
MS-TGNet [33]	95.71	70.50	94.41	17.19	95.72	88.83	76.01	73.97	**94.24**	23.64
diffConv [20]	-	76.73	83.31	51.06	69.04	79.55	80.48	**84.41**	76.19	**89.83**
Ours	**97.06**	**79.76**	**97.42**	**69.56**	94.79	**94.96**	78.21	83.35	91.54	28.21

Table 1. Quantitative results on Toronto-3D [33] dataset for semantic segmentation. We compare with different methods in terms of overall point accuracy (OA), mean per-class IoU (mIoU) and per-class mIoU.

Methods	Cat. mIoU	Ins. mIoU	aero	bag	cap	car	chair	earphone	guitar	knife	lamp	laptop	motor	mug	pistol	rocket	skateboard	table
PointNet [29]	80.4	83.7	83.4	78.7	82.5	74.9	89.6	73.0	91.5	85.9	80.8	95.3	65.2	93.0	81.2	57.9	72.8	80.6
PointNet++ [30]	81.9	85.1	82.4	79.0	87.7	77.3	90.8	71.8	91.0	85.9	83.7	95.3	71.6	94.1	81.3	58.7	76.4	82.6
PointCNN [18]	84.6	86.1	84.1	86.5	86.0	80.8	90.6	79.7	92.3	88.4	85.3	96.1	77.2	95.3	84.2	64.2	80.0	83.0
RS-CNN [24]	84.0	86.2	83.5	84.8	88.8	79.6	91.2	81.1	91.6	88.4	86.0	96.0	73.7	94.1	83.4	60.5	77.7	83.6
DGCNN [17]	82.3	85.2	84.0	83.4	86.7	77.8	90.6	74.7	91.2	87.5	82.8	95.7	66.3	94.9	81.1	63.5	74.5	82.6
KPConv [36]	85.1	86.4	84.6	86.3	87.2	81.1	91.1	77.8	92.6	88.4	82.7	96.2	78.1	95.8	85.4	69.0	82.0	83.6
DensePoint [23]	84.2	86.4	84.0	85.4	90.0	79.2	91.1	81.6	91.5	87.5	84.7	95.9	74.3	94.6	82.9	64.6	76.8	83.7
3D-GCN [22]	82.1	85.1	83.1	84.0	86.6	77.5	90.3	74.1	90.0	86.4	83.8	95.6	66.8	94.8	81.3	59.6	75.7	82.8
PAConv [41]	84.6	86.1	84.3	85.0	90.4	79.7	90.6	80.8	92.0	88.7	82.2	95.9	73.9	94.7	84.7	65.9	81.4	84.0
Ours	**85.3**	**87.0**	85.3	87.3	89.1	82.3	92.2	80.5	92.3	88.5	85.2	96.1	78.5	96.1	85.2	64.5	78.9	83.7

Table 2. Quantitative results on ShapeNetPart [43].

ing with some sparse point cloud datasets such as ScanNetV2 [5]. Besides, the way we estimate the feature correlation is relatively primitive.

Future work. In the future, we will try to implement our approach on large outdoor datasets such as SemanticKITTI. We will explore the voxel-like processing techniques in graph construction to deal with the sparse point cloud better. To further improve the graph representation, we will try to explore a more delicate hybrid structure and design a more efficient way to estimate correlations between points.

Methods	mIoU	mAcc	OA	ceiling	floor	wall	beam	column	window	door	table	chair	sofa	bookcase	board	clutter
PointNet [29]	41.1	49.0	-	88.8	97.3	69.8	0.1	3.9	46.3	10.8	59.0	52.6	5.9	40.3	26.4	33.2
SegCloud [35]	48.9	57.4	-	90.1	96.1	69.9	0.0	18.4	38.4	23.1	70.4	75.9	40.9	58.4	13.0	41.6
PointCNN [18]	57.3	63.9	85.9	92.3	98.2	79.4	0.0	17.6	22.8	62.1	74.4	80.6	31.7	66.7	62.1	56.7
SPGraph [15]	58.0	66.5	86.4	89.4	96.9	78.1	0.0	42.8	48.9	61.6	84.7	75.4	69.8	52.6	2.1	52.2
HPEIN [13]	61.9	68.3	87.2	91.5	98.2	81.4	0.0	23.3	65.3	40.0	75.5	87.7	58.5	67.8	65.6	49.4
MinkowskiNet [4]	65.4	71.7	-	91.8	98.7	86.2	0.0	34.1	48.9	62.4	81.6	89.8	47.2	74.9	74.4	58.6
KPConv [36]	67.1	72.8	-	92.8	97.3	82.4	0.0	23.9	58.0	69.0	81.5	91.0	75.4	75.3	66.7	58.9
JSENet [11]	67.7	-	-	93.8	97.0	83.0	0.0	23.2	61.3	71.6	89.9	79.8	75.6	72.3	72.7	60.4
RandLA-Net [10]	62.4	71.4	87.2	91.1	95.6	80.2	0.0	24.7	62.3	47.7	76.2	83.7	60.2	71.1	65.7	53.8
CloserLook3D [25]	66.9	72.1	90.0	94.8	98.4	82.5	0.0	25.5	51.3	70.9	92.1	81.9	76.7	70.1	64.5	61.2
PAConv [41]	66.6	73.0	-	94.6	98.6	82.4	0.0	26.4	58.0	60.0	89.7	80.4	74.3	69.8	73.5	57.7
CGA-Net [26]	68.6	-	-	94.5	98.3	83.0	0.0	25.3	59.6	71.0	92.2	82.6	76.4	77.7	69.5	61.5
PCT [9]	61.3	67.7	-	92.5	98.4	80.6	0.0	19.4	61.6	48.0	76.6	85.2	46.2	67.7	67.9	52.3
PointTrans. [46]	70.4	76.5	90.8	94.0	98.5	86.3	0.0	38.0	63.4	74.3	89.1	82.4	74.3	80.2	76.0	59.3
CBL [34]	69.4	75.2	90.6	93.9	98.4	84.2	0.0	37.0	57.7	71.9	91.7	81.8	77.8	75.6	69.1	62.9
FastPointTrans. [28]	68.7	77.1	-	93.8	97.8	85.5	0.6	49.9	60.5	72.9	80.2	88.7	56.0	71.4	78.0	58.1
Stratified Trans. [14]	72.0	78.1	**91.5**	96.2	98.7	85.6	0.0	46.1	60.0	76.8	92.6	84.5	77.8	75.2	78.1	64.0
RepSurf [32]	68.9	76.5	90.2	93.3	98.5	85.7	0.0	38.1	61.5	71.8	80.1	90.4	80.3	71.2	68.2	56.0
HilbertNet [2]	70.9	-	-	94.6	97.8	88.9	0.0	37.6	64.1	73.8	88.4	85.4	73.5	82.7	74.7	60.1
PointMixer [3]	71.4	77.4	-	94.2	98.2	86.0	0.0	43.8	62.1	78.5	90.6	82.2	73.9	79.8	78.5	59.4
Ours	**72.3**	77.9	91.1	95.4	98.6	85.1	0.0	47.3	58.3	80.4	83.9	92.0	80.6	77.3	79.9	61.8
	71.8 ± 0.5	77.5 ± 0.4	91.1 ± 0.1													
Ours†	**73.3**	**79.3**	**91.5**	94.6	98.6	86.3	0.0	51.7	62.8	79.0	84.5	91.8	83.5	77.0	82.4	61.0
	72.5 ± 0.8	78.5 ± 0.6	91.3 ± 0.3													

Table 3. Quantitative results on S3DIS Area 5 dataset [1].

Methods	mIoU	mAcc	OA	ceiling	floor	wall	beam	column	window	door	table	chair	sofa	bookcase	board	clutter
PointNet [29]	47.6	66.2	78.6	88.0	88.7	69.3	42.4	23.1	47.5	51.6	54.1	42.0	9.6	38.2	29.4	35.2
RSNet [12]	56.5	66.5	-	92.5	92.8	78.6	32.8	34.4	51.6	68.1	59.7	60.1	16.4	50.2	44.9	52.0
SPGraph [15]	62.1	73.0	86.4	89.9	95.1	76.4	62.8	47.1	55.3	68.4	73.5	69.2	63.2	45.9	8.7	52.9
PointCNN [18]	65.4	75.6	88.1	94.8	97.3	75.8	63.3	51.7	58.4	57.2	71.6	69.1	39.1	61.2	52.2	58.6
PointWeb [45]	66.7	76.2	87.3	93.5	94.2	80.8	52.4	41.3	64.9	68.1	71.4	67.1	50.3	62.7	62.2	58.5
ShellNet [44]	66.8	-	87.1	90.2	93.6	79.9	60.4	44.1	64.9	52.9	71.6	84.7	53.8	64.6	48.6	59.4
KPConv [36]	70.6	79.1	-	93.6	92.4	83.1	63.9	54.3	66.1	76.6	57.8	64.0	69.3	74.9	61.3	60.3
FPConv [21]	68.7	-	-	94.8	97.5	82.6	42.8	41.8	58.6	73.4	81.0	71.0	61.9	59.8	64.2	64.2
RandLA-Net [10]	70.0	82.0	88.0	93.1	96.1	80.6	62.4	48.0	64.4	69.4	69.4	76.4	60.0	64.2	65.9	60.1
SCF-Net [6]	71.6	82.7	88.4	93.3	96.4	80.9	64.9	47.4	64.5	70.1	71.4	81.6	67.2	64.4	67.5	60.9
PAConv [41]	69.3	78.7	-	94.3	93.5	82.8	56.9	45.7	65.2	74.9	59.7	74.6	67.4	61.8	65.8	58.4
BAAF [31]	72.2	83.1	88.9	93.3	96.8	81.6	61.9	49.5	65.4	73.3	72.0	83.7	67.5	64.3	67.0	62.4
PointTrans. [46]	73.5	81.9	90.2	94.3	97.5	84.7	55.6	58.1	66.1	78.2	77.6	74.1	67.3	71.2	65.7	64.8
CBL [34]	73.1	79.4	89.6	94.1	94.2	85.5	50.4	58.8	70.3	78.3	75.7	75.0	71.8	74.0	60.0	62.4
Ours	**77.7**	**85.1**	**91.7**	95.6	97.7	86.1	64.3	66.1	69.8	82.0	77.4	85.0	77.0	72.0	69.3	68.2
Ours†	**78.4**	**86.2**	**91.8**	94.8	97.8	86.7	63.2	69.7	70.5	81.4	76.6	89.4	78.5	71.8	70.6	68.5

Table 4. Quantitative results on S3DIS [1] with 6-fold cross validation.

Methods	Input	Val mIoU	Test mIoU	bath	bed	bksf	cab	chair	cntr	curt	desk	door	floor	othr	pic	ref	show	sink	sofa	tab	toil	wall	wind
PointNet++ [30]	point	53.5	55.7	73.5	66.1	68.6	49.1	74.4	39.2	53.9	45.1	37.5	94.6	37.6	20.5	40.3	35.6	55.3	64.3	49.7	82.4	75.6	51.5
PointCNN [18]	point	-	45.8	57.7	61.1	35.6	32.1	71.5	29.9	37.6	32.8	31.9	94.4	28.5	16.4	21.6	22.9	48.4	54.5	45.6	75.5	70.9	47.5
PointConv [39]	point	61.0	66.6	78.1	75.9	69.9	64.4	82.2	47.5	77.9	56.4	50.4	95.3	42.8	20.3	58.6	75.4	66.1	75.3	58.8	90.2	81.3	64.2
SparseConvNet [8]	voxel	69.3	72.5	64.7	82.1	84.6	72.1	86.9	53.3	75.4	60.3	61.4	95.5	57.2	32.5	71.0	87.0	72.4	82.3	62.8	93.4	86.5	68.3
KPConv [36]	point	69.2	68.4	84.7	75.8	78.4	64.7	81.4	47.3	77.2	60.5	59.4	93.5	45.0	18.1	58.7	80.5	69.0	78.5	61.4	88.2	81.9	63.2
MinkowskiNet [4]	voxel	72.2	73.6	85.9	81.8	83.2	70.9	84.0	52.1	85.3	66.0	64.3	95.1	54.4	28.6	73.1	89.3	67.5	77.2	68.3	87.4	85.2	72.7
SegGCN [16]	point	-	58.9	83.3	73.1	53.9	51.4	78.9	44.8	46.7	57.3	48.4	93.6	39.6	6.1	50.1	50.7	59.4	70.0	56.3	87.4	77.1	49.3
RandLA-Net [10]	point	-	64.5	77.8	73.1	69.9	57.7	82.9	44.6	73.6	47.7	52.3	94.5	45.4	26.9	48.4	74.9	61.8	73.8	59.9	82.7	79.2	62.1
PointASNL [42]	point	63.5	66.6	70.3	78.1	75.1	65.5	83.0	47.1	76.9	47.4	53.7	95.1	47.5	27.9	63.5	69.8	67.5	75.1	55.3	81.6	80.6	70.3
JSENet [11]	point	-	69.9	88.1	76.2	82.1	66.7	80.0	52.2	79.2	61.3	60.7	93.5	49.2	20.5	57.6	85.3	69.1	75.8	65.2	87.2	82.8	64.9
RFCR [7]	point	-	70.2	88.9	74.5	81.3	67.2	81.8	49.3	81.5	62.3	61.0	94.7	47.0	24.9	59.4	84.8	70.5	77.9	64.6	89.2	82.3	61.1
CBL [34]	point	-	70.5	76.9	77.5	80.9	68.7	82.0	43.9	81.2	66.1	59.1	94.5	51.5	17.1	63.3	85.6	72.0	79.6	66.8	88.9	84.7	68.9
Ours†	point	73.4	71.8	84.5	75.1	81.6	71.4	85.1	52.8	81.9	62.3	60.3	95.5	52.9	28.0	65.9	86.8	71.3	78.5	60.1	90.0	84.4	68.1

Table 5. Quantitative results on ScanNetV2 [5].

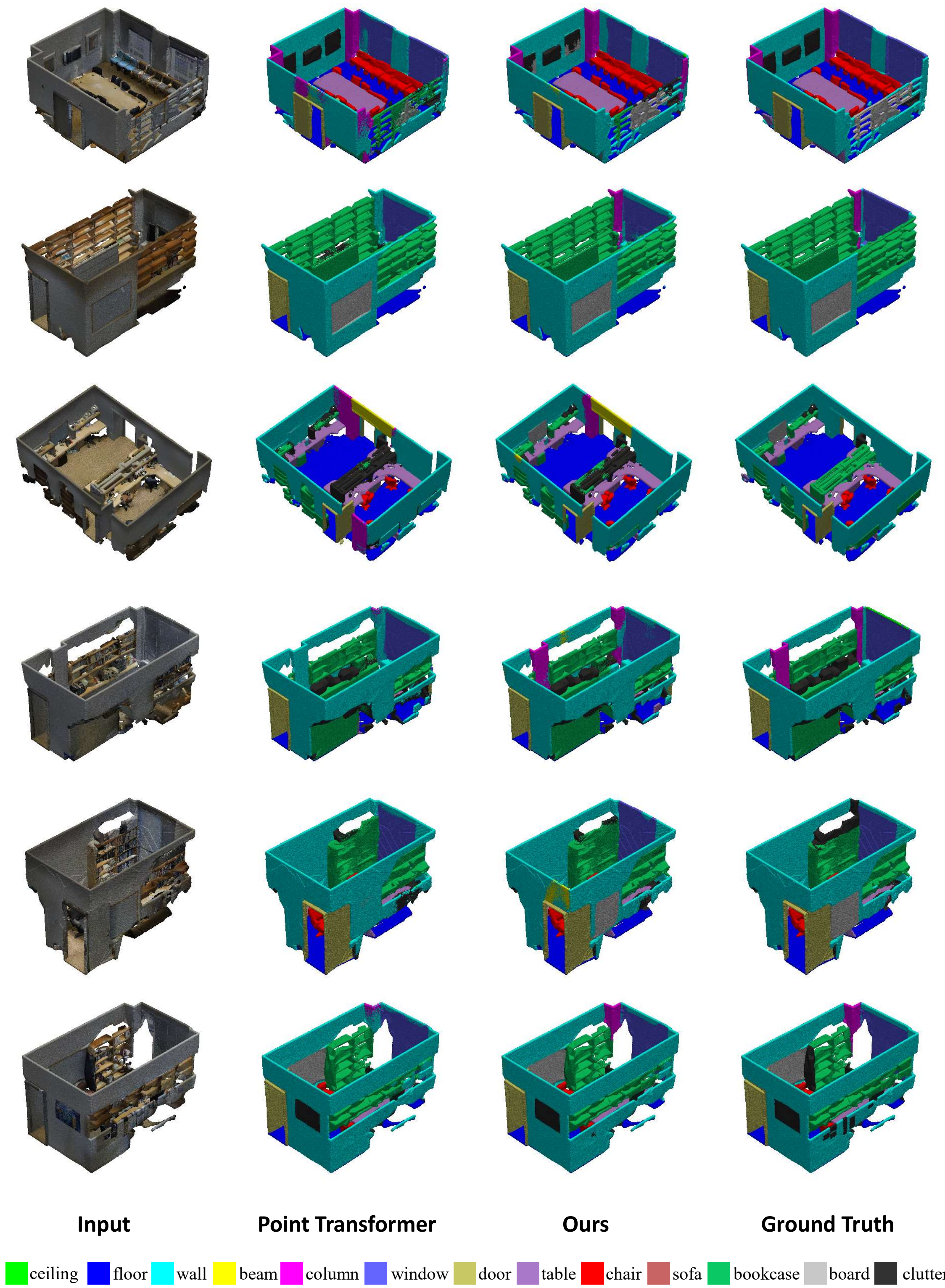

Figure 2. More visualization compared with PointTransformer [46]. Zoom-in for a better view.

References

[1] Iro Armeni, Ozan Sener, Amir R. Zamir, Helen Jiang, Ioannis Brilakis, Martin Fischer, and Silvio Savarese. 3d semantic parsing of large-scale indoor spaces. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2016. 1, 3

[2] Wanli Chen, Xinge Zhu, Guojin Chen, and Bei Yu. Efficient point cloud analysis using hilbert curve. In *European Conference on Computer Vision*, pages 730–747. Springer, 2022. 3

[3] Jaesung Choe, Chunghyun Park, Francois Rameau, Jaesik Park, and In So Kweon. Pointmixer: Mlp-mixer for point cloud understanding. In *European Conference on Computer Vision*. Springer, 2022. 3

[4] Christopher Choy, JunYoung Gwak, and Silvio Savarese. 4d spatio-temporal convnets: Minkowski convolutional neural networks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2019. 3

[5] Angela Dai, Angel X. Chang, Manolis Savva, Maciej Halber, Thomas Funkhouser, and Matthias Niessner. Scannet: Richly-annotated 3d reconstructions of indoor scenes. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, July 2017. 1, 2, 3

[6] Siqi Fan, Qiulei Dong, Fenghua Zhu, Yisheng Lv, Peijun Ye, and Fei-Yue Wang. Scf-net: Learning spatial contextual features for large-scale point cloud segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 14504–14513, June 2021. 3

[7] Jingyu Gong, Jiachen Xu, Xin Tan, Haichuan Song, Yanyun Qu, Yuan Xie, and Lizhuang Ma. Omni-supervised point cloud segmentation via gradual receptive field component reasoning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 11673–11682, June 2021. 3

[8] Benjamin Graham, Martin Engelcke, and Laurens van der Maaten. 3d semantic segmentation with submanifold sparse convolutional networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2018. 3

[9] Meng-Hao Guo, Jun-Xiong Cai, Zheng-Ning Liu, Tai-Jiang Mu, Ralph R. Martin, and Shi-Min Hu. Pct: Point cloud transformer. *Computational Visual Media*, 7(2):187–199, Apr 2021. 3

[10] Qingyong Hu, Bo Yang, Linhai Xie, Stefano Rosa, Yulan Guo, Zhihua Wang, Niki Trigoni, and Andrew Markham. Randla-net: Efficient semantic segmentation of large-scale point clouds. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2020. 3

[11] Zeyu Hu, Mingmin Zhen, Xuyang Bai, Hongbo Fu, and Chiew-lan Tai. Jsenet: Joint semantic segmentation and edge detection network for 3d point clouds. In *European Conference on Computer Vision*, pages 222–239. Springer, 2020. 3

[12] Qiangui Huang, Weiyue Wang, and Ulrich Neumann. Recurrent slice networks for 3d segmentation of point clouds. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2018. 3

[13] Li Jiang, Hengshuang Zhao, Shu Liu, Xiaoyong Shen, Chi-Wing Fu, and Jiaya Jia. Hierarchical point-edge interaction network for point cloud semantic segmentation. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, October 2019. 3

[14] Xin Lai, Jianhui Liu, Li Jiang, Liwei Wang, Hengshuang Zhao, Shu Liu, Xiaojuan Qi, and Jiaya Jia. Stratified transformer for 3d point cloud segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 8500–8509, June 2022. 1, 3

[15] Loic Landrieu and Martin Simonovsky. Large-scale point cloud semantic segmentation with superpoint graphs. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2018. 3

[16] Huan Lei, Naveed Akhtar, and Ajmal Mian. Seggcn: Efficient 3d point cloud segmentation with fuzzy spherical kernel. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2020. 3

[17] Guohao Li, Matthias Müller, Guocheng Qian, Itzel Carolina Delgadillo Perez, Abdulellah Abualshour, Ali Kassem Thabet, and Bernard Ghanem. Deepgcns: Making gcns go as deep as cnns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2021. 2

[18] Yangyan Li, Rui Bu, Mingchao Sun, Wei Wu, Xinhan Di, and Baoquan Chen. Pointcnn: Convolution on x-transformed points. In *Advances in Neural Information Processing Systems*, volume 31, 2018. 2, 3

[19] Ying Li, Lingfei Ma, Zilong Zhong, Dongpu Cao, and Jonathan Li. Tgnet: Geometric graph cnn on 3-d point cloud segmentation. *IEEE Transactions on Geoscience and Remote Sensing*, 58(5):3588–3600, 2019. 1, 2

[20] Manxi Lin and Aasa Feragen. diffconv: Analyzing irregular point clouds with an irregular view. In *European Conference on Computer Vision*, 2022. 1, 2

[21] Yiqun Lin, Zizheng Yan, Haibin Huang, Dong Du, Ligang Liu, Shuguang Cui, and Xiaoguang Han. Fpconv: Learning local flattening for point convolution. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2020. 3

[22] Zhi-Hao Lin, Sheng-Yu Huang, and Yu-Chiang Frank Wang. Convolution in the cloud: Learning deformable kernels in 3d graph convolution networks for point cloud analysis. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2020. 2

[23] Yongcheng Liu, Bin Fan, Gaofeng Meng, Jiwen Lu, Shiming Xiang, and Chunhong Pan. Densepoint: Learning densely contextual representation for efficient point cloud processing. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, October 2019. 2

[24] Yongcheng Liu, Bin Fan, Shiming Xiang, and Chunhong Pan. Relation-shape convolutional neural network for point cloud analysis. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2019. 2

[25] Ze Liu, Han Hu, Yue Cao, Zheng Zhang, and Xin Tong. A closer look at local aggregation operators in point cloud analysis. In *European Conference on Computer Vision*, pages 326–342. Springer, 2020. 3

[26] Tao Lu, Limin Wang, and Gangshan Wu. Cga-net: Category guided aggregation for point cloud semantic segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 11693–11702, June 2021. 3

[27] Lingfei Ma, Ying Li, Jonathan Li, Weikai Tan, Yongtao Yu, and Michael A. Chapman. Multi-scale point-wise convolutional neural networks for 3d object segmentation from lidar point clouds in large-scale environments. *IEEE Transactions on Intelligent Transportation Systems*, 22(2):821–836, 2021. 2

[28] Chunghyun Park, Yoonwoo Jeong, Minsu Cho, and Jaesik Park. Fast point transformer. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 16949–16958, June 2022. 3

[29] Charles R. Qi, Hao Su, Kaichun Mo, and Leonidas J. Guibas. Pointnet: Deep learning on point sets for 3d classification and segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 652–660, July 2017. 2, 3

[30] Charles Ruizhongtai Qi, Li Yi, Hao Su, and Leonidas J Guibas. Pointnet++: Deep hierarchical feature learning on point sets in a metric space. In *Advances in Neural Information Processing Systems*, volume 30, 2017. 2, 3

[31] Shi Qiu, Saeed Anwar, and Nick Barnes. Semantic segmentation for real point cloud scenes via bilateral augmentation and adaptive fusion. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1757–1767, June 2021. 3

[32] Haoxi Ran, Jun Liu, and Chengjie Wang. Surface representation for point clouds. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 18942–18952, June 2022. 3

[33] Weikai Tan, Nannan Qin, Lingfei Ma, Ying Li, Jing Du, Guorong Cai, Ke Yang, and Jonathan Li. Toronto-3d: A large-scale mobile lidar dataset for semantic segmentation of urban roadways. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, June 2020. 1, 2

[34] Liyao Tang, Yibing Zhan, Zhe Chen, Baosheng Yu, and Dacheng Tao. Contrastive boundary learning for point cloud segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 8489–8499, June 2022. 3

[35] Lyne Tchapmi, Christopher Choy, Iro Armeni, JunYoung Gwak, and Silvio Savarese. Segcloud: Semantic segmentation of 3d point clouds. In *International Conference on 3D Vision (3DV)*, pages 537–547. IEEE, 2017. 3

[36] Hugues Thomas, Charles R. Qi, Jean-Emmanuel Deschaud, Beatriz Marcotegui, Francois Goulette, and Leonidas J. Guibas. Kpconv: Flexible and deformable convolution for point clouds. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, October 2019. 2, 3

[37] Mikaela Angelina Uy, Quang-Hieu Pham, Binh-Son Hua, Duc Thanh Nguyen, and Sai-Kit Yeung. Revisiting point cloud classification: A new benchmark dataset and classification model on real-world data. In *International Conference on Computer Vision (ICCV)*, 2019. 1

[38] Yue Wang, Yongbin Sun, Ziwei Liu, Sanjay E Sarma, Michael M Bronstein, and Justin M Solomon. Dynamic graph cnn for learning on point clouds. *Acm Transactions On Graphics (tog)*, 38(5):1–12, 2019. 2

[39] Wenxuan Wu, Zhongang Qi, and Li Fuxin. Pointconv: Deep convolutional networks on 3d point clouds. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 9621–9630, 2019. 3

[40] Tiange Xiang, Chaoyi Zhang, Yang Song, Jianhui Yu, and Weidong Cai. Walk in the cloud: Learning curves for point clouds shape analysis. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 915–924, October 2021. 1

[41] Mutian Xu, Runyu Ding, Hengshuang Zhao, and Xiaojuan Qi. Paconv: Position adaptive convolution with dynamic kernel assembling on point clouds. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3173–3182, June 2021. 2, 3

[42] Xu Yan, Chaoda Zheng, Zhen Li, Sheng Wang, and Shuguang Cui. Pointasnl: Robust point clouds processing using nonlocal neural networks with adaptive sampling. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2020. 3

[43] Li Yi, Vladimir G Kim, Duygu Ceylan, I-Chao Shen, Mengyan Yan, Hao Su, Cewu Lu, Qixing Huang, Alla Sheffer, and Leonidas Guibas. A scalable active framework for region annotation in 3d shape collections. *ACM Transactions on Graphics (ToG)*, 35(6):1–12, 2016. 1, 2

[44] Zhiyuan Zhang, Binh-Son Hua, and Sai-Kit Yeung. Shellnet: Efficient point cloud convolutional neural networks using concentric shells statistics. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, October 2019. 3

[45] Hengshuang Zhao, Li Jiang, Chi-Wing Fu, and Jiaya Jia. Pointweb: Enhancing local neighborhood features for point cloud processing. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2019. 3

[46] Hengshuang Zhao, Li Jiang, Jiaya Jia, Philip H.S. Torr, and Vladlen Koltun. Point transformer. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 16259–16268, October 2021. 3, 4