PointNorm: Dual Normalization is All You Need for Point Cloud Analysis

Shen Zheng, Jinqian Pan, Changjie Lu, Gaurav Gupta

Carnegie Mellon University, New York University, Wenzhou-Kean University

INTERNATIONAL NEURAL NETWORK SOCIETY

IJCNN2023

INTERNATIONAL JOINT CONFERENCE ON NEURAL NETWORKS

18–23 JUNE 2023 | GOLD COAST CONVENTION AND EXHIBITION CENTRE, QUEENSLAND, AUSTRALIA

Outline

Background

Algorithm-Level Applications

- Shape classification
- Object detection
- Semantic segmentation

System-Level Applications

- Autonomous driving
- Robotics
- VR & AR

Challenges

Point Cloud Irregularity

- Uneven distribution
- Difficult surface representation

Computational Complexity

- Complex feature extractors
- Large-scale point cloud analysis

Related Works: Point Cloud Analysis

Intermediate Voxels

- Voxnet [1], Voxelnet [2]
- Loss of fine-grained details

Raw Point Cloud

- PointNet [3], PointNet++ [4]
- Better detail preservation

Irregularity Unsolved -> Poor Accuracy

Related Works: Point Cloud Geometry

Local Geometry

- PointConv [5]: density functions
- EdgeConv [6]: graph nodes

Global Geometry

- PointASNL [7]: local-nonlocal module
- CurveNet [8]: curve-based guided walk

Complex Optimization → Poor Latency

Workflow

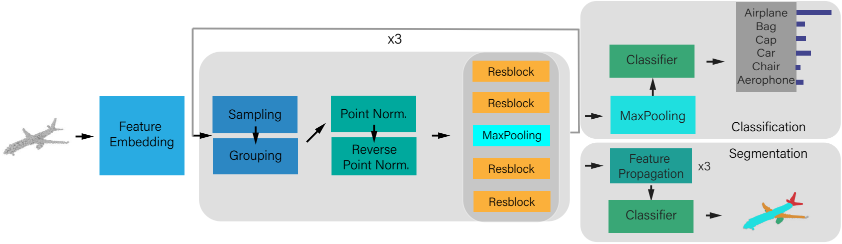


Figure 1: The workflow of PointNorm for shape classification and part segmentation.

DualNorm

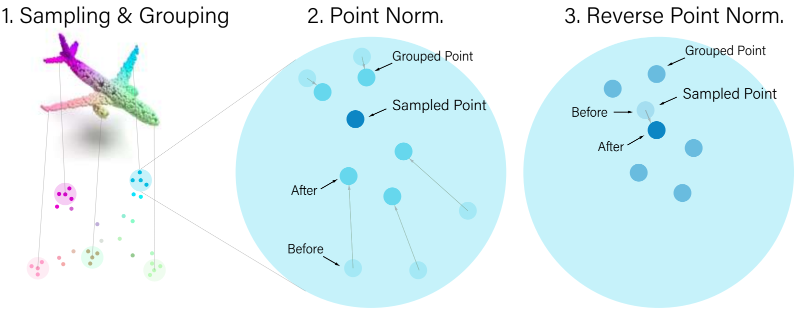


Figure 2: Overview of DualNorm (Point Normalization + Reverse Point Normalization).

Standard Deviation Analysis

Suppose α_i is a parameter, σ_i is the standard deviation, and $\Delta = \frac{\alpha_i}{\sigma_i}$.

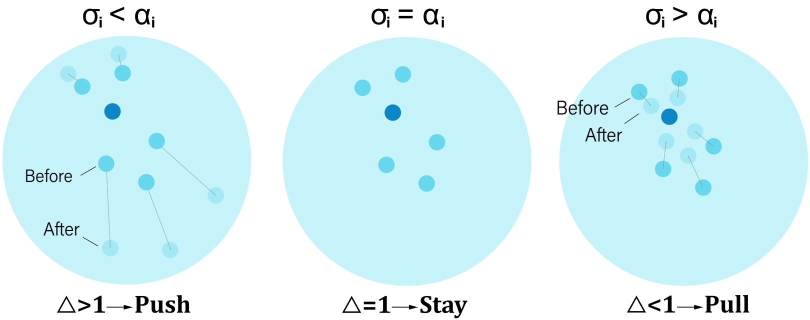


Figure 3: PointNorm's 'push-and-pull' strategy for optimizing the point cloud density.

Optimization Landscape Analysis

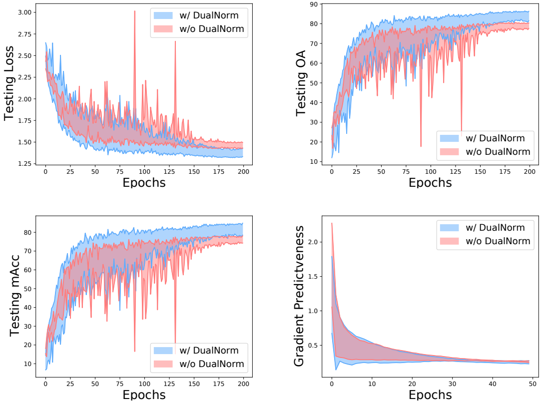


Figure 4: Optimization Landscape Analysis with Loss, OA, mAcc, and Gradient

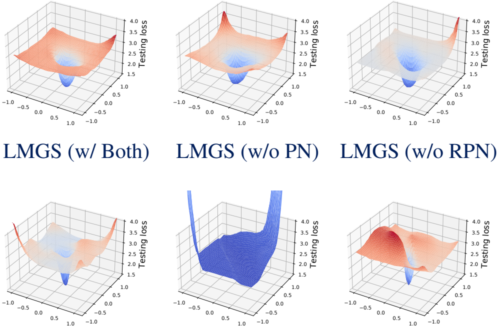
Implementation Details

- AdamW
- Cosine Annealing
- Initial lr = 1e-2
- Final lr = 1e-4
- Batch Size = 32
- Embedding Dimension = 64
- Cross Entropy + Label Smoothing
- Random Rotation + Translation

Ablation Studies: Quantitative

		OA (%)	mAcc (%)	FLOPs	#Params	Train Time	Test Time
Layer	24	86.5	85.2	8.71G	7.30M	99	9
	40	**86.8**	**85.6**	14.59G	12.63M	145	13
	56	86.7	85.0	20.48G	17.95M	190	16
Bottleneck	0.25	85.9	84.5	5.79G	4.65M	99	9
	0.50	86.6	85.4	8.72G	7.31M	111	10
	1.00	**86.8**	**85.6**	14.59G	12.63M	145	13
	2.00	86.6	85.1	26.34G	23.27M	205	17
Local/Global	LMGS	**86.8**	**85.6**	14.59G	12.63M	145	13
	LMLS	81.2	78.3	14.59G	12.63M	145	13
	GMLS	25.1	16.8	14.59G	12.63M	143	13
	GMGS	78.4	75.5	14.59G	12.63M	143	13
Norm.	w/o PN	80.7	77.8	14.59G	12.63M	136	12
	w/o RPN	85.6	84.1	14.59G	12.63M	141	12
	w/ both	**86.8**	**85.6**	14.59G	12.63M	145	13

Ablation Studies: Qualitative



LMGS (w/ Both) LMGS (w/o PN) LMGS (w/o RPN)

LMLS (w/ Both) GMLS (w/ Both) GMGS (w/ Both)

Figure 5: Loss Landscape [9] along two random directions for different PointNorm variants.

Shape Classification: Quantitative

Method	Publication	Input	ModelNet40 [10]		ScanObjectNN [11]		#Params	Train Speed	Test Speed
			OA (%)	mAcc (%)	OA (%)	mAcc (%)			
PointNet [3]	CVPR 2017	1k	89.2	86.0	68.2	63.4	3.47M	-	-
PointNet++ [4]	NeurIPS 2017	1k	90.7	88.4	77.9	75.4	1.48M	**223.8**	308.5
PointCNN [12]	NeurIPS 2018	1k	92.5	88.1	78.5	75.1	-	-	-
DGCNN [6]	TOG 2019	1k	92.9	90.2	78.1	73.6	1.82M	-	-
RS-CNN [13]	CVPR 2019	1k	92.9	-	-	-	2.38M	-	-
PointConv [5]	CVPR 2019	1k	92.5	-	-	-	18.6M	17.9	10.2
KPConv [14]	ICCV 2019	7k	92.9	-	-	-	14.3M	31.0	80.0
PointASNL [7]	CVPR 2020	1k	93.2	-	-	-	10.1M	-	-
Grid-GCN [15]	CVPR 2020	1k	93.1	**91.3**	-	-	-	-	-
DRNet [16]	WACV 2021	1k	93.1	-	80.3	78.0	-	-	-
PAConv [17]	CVPR 2021	1k	93.6	-	-	-	2.44M	-	-
CurveNet [8]	ICCV 2021	1k	93.8	-	-	-	2.04M	20.8	15.0
GDANet [18]	AAAI 2021	1k	93.4	-	-	-	0.93M	26.3	14.0
PRANet [19]	TIP 2021	1k	93.2	90.6	81.0	77.9	-	-	-
PointMLP [20]	ICLR 2022	1k	**94.1**	**91.3**	85.4	83.9	12.60M	47.1	112.0
RepSurf-U [21]	CVPR 2022	1k	-	-	84.6	81.9	1.48M	-	-
PointNorm		1k	**94.1**	**91.3**	**86.8**	**85.6**	12.63M	58.2	140.0
PointNorm-Tiny		1k	93.5	90.6	85.3	83.6	**0.68M**	196.4	**420.0**

Part Segmentation: Quantitative

Method	Inst. mIoU	Cls. mIoU	air-plane	bag	cap	car	chair	aero-phone	guitar	knife	lamp	laptop	motor-bike	mug	pistol	rocket	skate-board	table
PointNet [3]	83.7	80.4	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++ [4]	85.1	81.9	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
PCNN [22]	85.1	81.8	82.4	80.1	85.5	79.5	90.8	73.2	91.3	86.0	85.0	95.7	73.2	94.8	83.3	51.0	75.0	81.8
DGCNN [6]	85.2	82.3	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
PointCNN [12]	86.1	84.6	84.1	86.5	86.0	80.8	90.6	79.7	92.3	88.4	85.3	96.1	77.2	95.2	84.2	64.2	80.0	83.0
RS-CNN [13]	**86.2**	84.0	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
SyncSpecCNN [23]	84.7	82.0	81.6	81.7	81.9	75.2	90.2	74.9	93.0	86.1	84.7	95.6	66.7	92.7	81.6	60.6	82.9	82.1
SPLATNet [24]	85.4	83.7	83.2	84.3	89.1	80.3	90.7	75.5	92.1	87.1	83.9	96.3	75.6	95.8	83.8	64.0	75.5	81.8
SpiderCNN [25]	85.3	82.4	83.5	81.0	87.2	77.5	90.7	76.8	91.1	87.3	83.3	95.8	70.2	93.5	82.7	59.7	75.8	82.8
PAConv [17]	86.1	84.6	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
PointMLP [20]	86.1	84.6	83.5	83.4	87.5	80.5	90.3	78.2	92.2	88.1	82.6	96.2	77.5	95.8	85.4	64.6	83.3	84.3
PointNorm	**86.2**	**84.7**	82.7	84.9	88.9	79.8	90.2	81.9	91.6	87.4	82.9	95.8	78.4	95.5	84.5	65.6	81.4	83.8
PointNorm-Tiny	85.6	84.5	82.9	88.0	89.7	79.3	90.1	79.9	91.6	87.7	82.4	95.8	76.3	95.0	83.5	64.6	81.9	83.5

Table 1: Part Segmentation Result on ShapeNetPart [26].

Part Segmentation: Qualitative

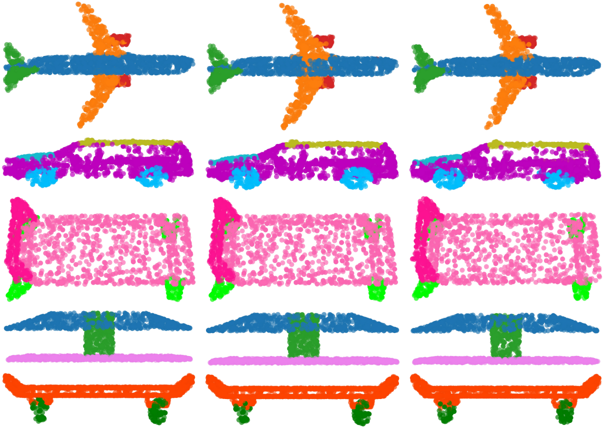


Figure 6: GT, PointNorm, PointNorm-Tiny.

Semantic Segmentation: Quantitative

Method	S3DIS 6-Fold			S3DIS Area-5			#Params	FLOPs
	mIoU	mAcc	OA	mIoU	mAcc	OA		
PointNet [3]	47.6	66.2	78.5	43.2	52.6	77.8	1.7M	4.1G
PointWeb [27]	66.7	76.2	87.3	60.2	66.6	87.0	-	-
KPConv [14]	70.6	**79.1**	-	**67.1**	**72.8**	-	14.9M	-
PointASNL [7]	68.7	79.0	**88.8**	62.6	68.5	87.7	22.4M	19.1G
RPNet	**70.8**	-	-	-	-	-	2.4M	5.1G
DSPoint	63.3	70.9	-	-	-	-	-	-
PointNet++	54.5	67.1	81.0	52.6	63.1	82.3	0.969M	1.00G
PointNet++ (w/ DualNorm)	62.7	73.8	85.7	57.6	68.2	**88.4**	1.006M	1.05G
(w/ DualNorm)	↑8.2	↑6.7	↑4.7	↑5.0	↑5.1	↑6.1	↑0.037M	↑0.05G

Semantic Segmentation: Qualitative

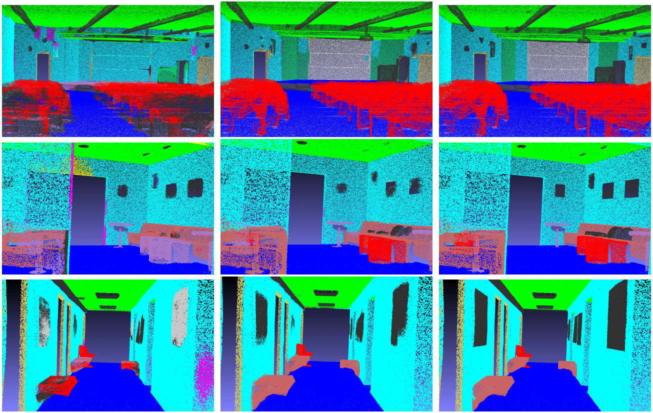


Figure 7: PointNet++, PointNet++ (w/ DualNorm), GT.

Conclusion

Summary

- DualNorm: Point Norm. + Reverse Point Norm.
- Normalization solves irregularity
- Local mean and global std. improves efficiency

Future Works

- Object Detection (SUN RGB-D [28])
- Outdoor Semantic Segmentation (SemanticKITTI [29])

References I

[1] Daniel Maturana and Sebastian Scherer. "Voxnet: A 3d convolutional neural network for real-time object recognition". In: *2015 IEEE/RSJ international conference on intelligent robots and systems (IROS)*. IEEE. 2015, pp. 922–928.

[2] Yin Zhou and Oncel Tuzel. "Voxelnet: End-to-end learning for point cloud based 3d object detection". In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2018, pp. 4490–4499.

[3] Charles R Qi et al. "Pointnet: Deep learning on point sets for 3d classification and segmentation". In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017, pp. 652–660.

[4] Charles Ruizhongtai Qi et al. "Pointnet++: Deep hierarchical feature learning on point sets in a metric space". In: *Advances in neural information processing systems* 30 (2017).

[5] 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*. 2019, pp. 9621–9630.

[6] Yue Wang et al. "Dynamic graph cnn for learning on point clouds". In: *Acm Transactions On Graphics (tog)* 38.5 (2019), pp. 1–12.

[7] Xu Yan et al. "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*. 2020, pp. 5589–5598.

References II

[8] Tiange Xiang et al. “Walk in the cloud: Learning curves for point clouds shape analysis”. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 2021, pp. 915–924.

[9] Hao Li et al. “Visualizing the loss landscape of neural nets”. In: *Advances in neural information processing systems* 31 (2018).

[10] Zhirong Wu et al. “3d shapenets: A deep representation for volumetric shapes”. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2015, pp. 1912–1920.

[11] Mikaela Angelina Uy et al. “Revisiting point cloud classification: A new benchmark dataset and classification model on real-world data”. In: *Proceedings of the IEEE/CVF international conference on computer vision*. 2019, pp. 1588–1597.

[12] Yangyan Li et al. “Pointcnn: Convolution on x-transformed points”. In: *Advances in neural information processing systems* 31 (2018).

[13] Yongcheng Liu et al. “Relation-shape convolutional neural network for point cloud analysis”. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2019, pp. 8895–8904.

[14] Hugues Thomas et al. “Kpconv: Flexible and deformable convolution for point clouds”. In: *Proceedings of the IEEE/CVF international conference on computer vision*. 2019, pp. 6411–6420.

References III

[15] Qiangeng Xu et al. "Grid-gcn for fast and scalable point cloud learning". In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2020, pp. 5661–5670.

[16] Shi Qiu, Saeed Anwar, and Nick Barnes. "Dense-resolution network for point cloud classification and segmentation". In: *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*. 2021, pp. 3813–3822.

[17] Mutian Xu et al. "Paconv: Position adaptive convolution with dynamic kernel assembling on point clouds". In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2021, pp. 3173–3182.

[18] Mutian Xu et al. "Learning geometry-disentangled representation for complementary understanding of 3d object point cloud". In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 35. 4. 2021, pp. 3056–3064.

[19] Silin Cheng et al. "Pra-net: Point relation-aware network for 3d point cloud analysis". In: *IEEE Transactions on Image Processing* 30 (2021), pp. 4436–4448.

[20] Xu Ma et al. "Rethinking network design and local geometry in point cloud: A simple residual mlp framework". In: *arXiv preprint arXiv:2202.07123* (2022).

[21] 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*. 2022, pp. 18942–18952.

References IV

[22] Matan Atzmon, Haggai Maron, and Yaron Lipman. "Point convolutional neural networks by extension operators". In: *arXiv preprint arXiv:1803.10091* (2018).

[23] Li Yi et al. "Syncspeccnn: Synchronized spectral cnn for 3d shape segmentation". In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017, pp. 2282–2290.

[24] Hang Su et al. "Splatnet: Sparse lattice networks for point cloud processing". In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2018, pp. 2530–2539.

[25] Yifan Xu et al. "Spidercnn: Deep learning on point sets with parameterized convolutional filters". In: *Proceedings of the European Conference on Computer Vision (ECCV)*. 2018, pp. 87–102.

[26] Li Yi et al. "A scalable active framework for region annotation in 3d shape collections". In: *ACM Transactions on Graphics (ToG)* 35.6 (2016), pp. 1–12.

[27] Hengshuang Zhao et al. "Pointweb: Enhancing local neighborhood features for point cloud processing". In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2019, pp. 5565–5573.

[28] Shuran Song, Samuel P Lichtenberg, and Jianxiong Xiao. "Sun rgb-d: A rgb-d scene understanding benchmark suite". In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2015, pp. 567–576.

References V

[29] Jens Behley et al. “Semantickitti: A dataset for semantic scene understanding of lidar sequences”. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 2019, pp. 9297–9307.