Frugal 3D Point Cloud Model Training via Progressive Near Point Filtering and Fused Aggregation

Donghyun Lee
eudh1206@snu.ac.kr

Yejin Lee
yejinlee@snu.ac.kr

Jae W. Lee
jaewlee@snu.ac.kr

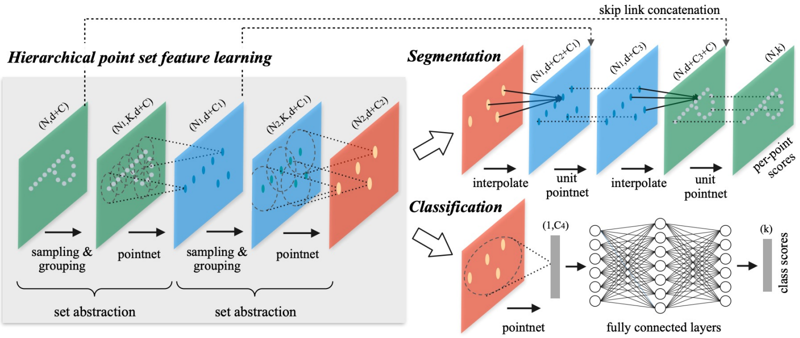
Hongil Yoon
hongilyoon@google.com

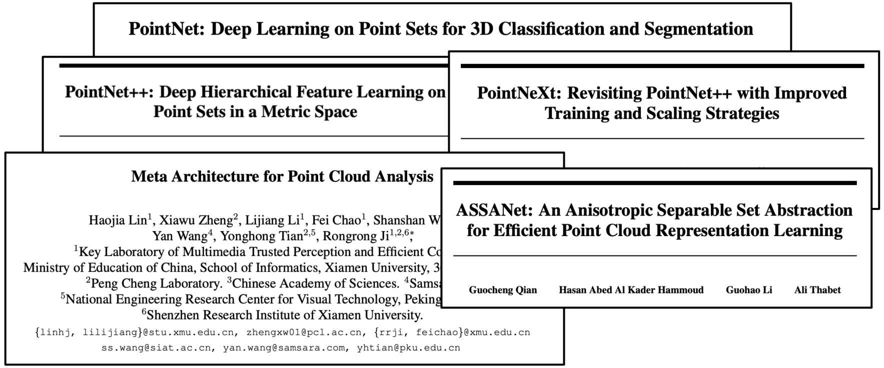
Department of CSE
Seoul National University

European Conference on Computer Vision (ECCV 2024), Milan, Italy

Deep Neural Networks for 3D Point Cloud

- PointNet and PointNet++ are the first to apply DNN to raw 3D point cloud without preprocessing.
- Various ideas have been proposed, continuously enhancing the model performance and computational efficiency.




PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation

PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space

PointNeXt: Revisiting PointNet++ with Improved Training and Scaling Strategies

Meta Architecture for Point Cloud Analysis

Haojia Lin[1], Xiawu Zheng[2], Lijiang Li[1], Fei Chao[1], Shanshan W
Yan Wang[4], Yonghong Tian[2,5], Rongrong Ji[1,2,6]*
[1]Key Laboratory of Multimedia Trusted Perception and Efficient Co
Ministry of Education of China, School of Informatics, Xiamen University, 3
[2]Peng Cheng Laboratory. [3]Chinese Academy of Sciences. [4]Sams
[5]National Engineering Research Center for Visual Technology, Peking
[6]Shenzhen Research Institute of Xiamen University.
{linhj, lilijiang}@stu.xmu.edu.cn, zhengxw01@pcl.ac.cn, {rrji, feichao}@xmu.edu.cn
ss.wang@siat.ac.cn, yan.wang@samsara.com, yhtian@pku.edu.cn

ASSANet: An Anisotropic Separable Set Abstraction for Efficient Point Cloud Representation Learning

Guocheng Qian Hasan Abed Al Kader Hammoud Guohao Li Ali Thabet

Figures from Qi, C.R. et al. “PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space”

Challenges in 3D Point Cloud Model Training

Challenges

↑ the size of models & datasets → ↑ training cost for 3D point cloud models

Performance Bottleneck

Farthest Point Sampling (FPS): Takes up average **44.69%** of overall training time.

Aggregation: Takes up average **22.84%** of overall training time.

Our Proposal

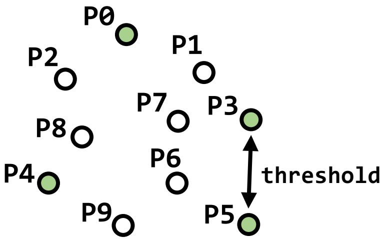
#1. L-FPS: Eliminates **redundant distance calculation** of FPS in the training pipeline.

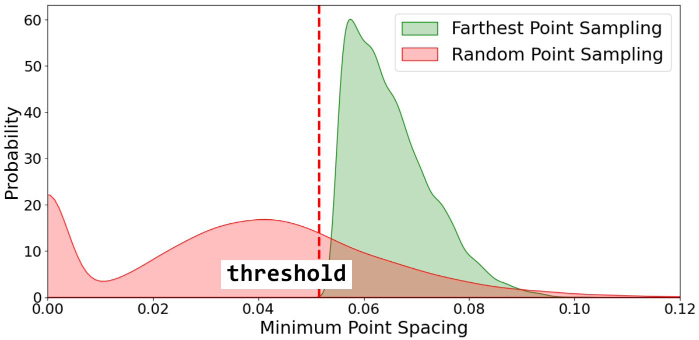
#2. Fused Aggregation: Reduce **redundant memory accesses** during aggregation.

Farthest Point Sampling - Observations

Observation #1. FPS in the training process incurs a significant number of **redundant distance calculations across epochs.**

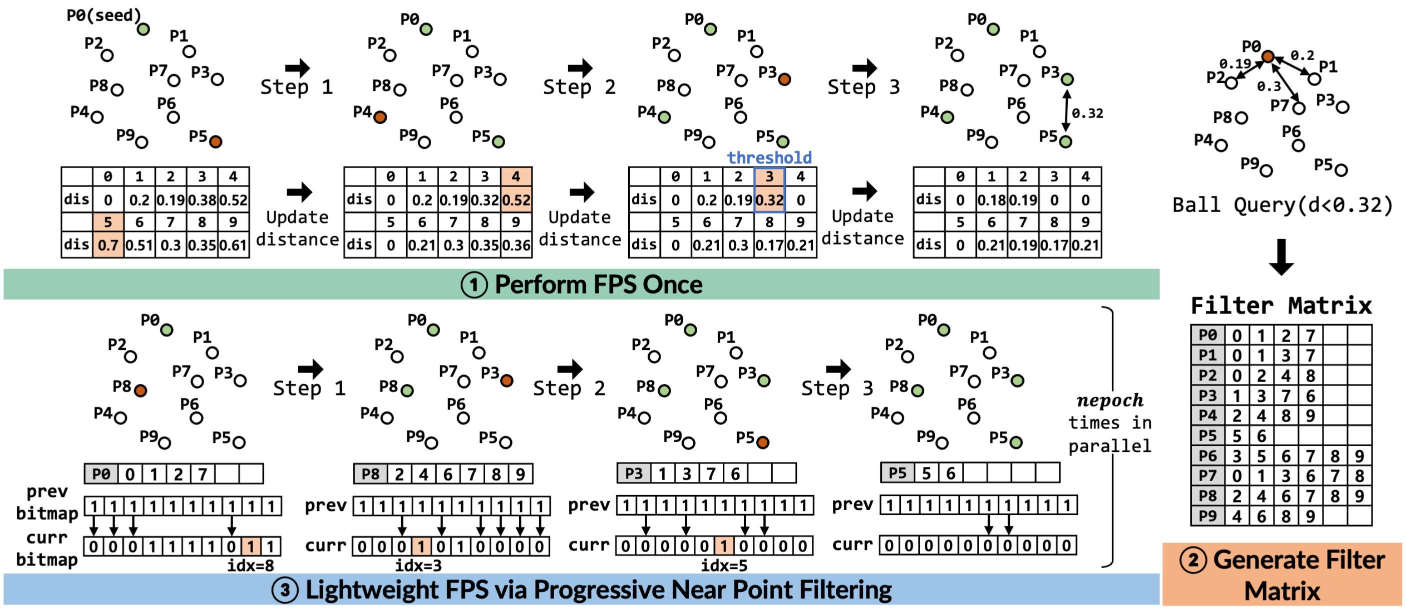
Observation #2. The key factor in achieving high-quality sampling is to ensure a **minimum spacing among the sampled points,** and this information can be obtained in advance, prior to training.





Technique #1. Lightweight FPS (L-FPS)

- We propose ***Lightweight FPS via Progressive Near Point Filtering.***



Good Sampling Quality

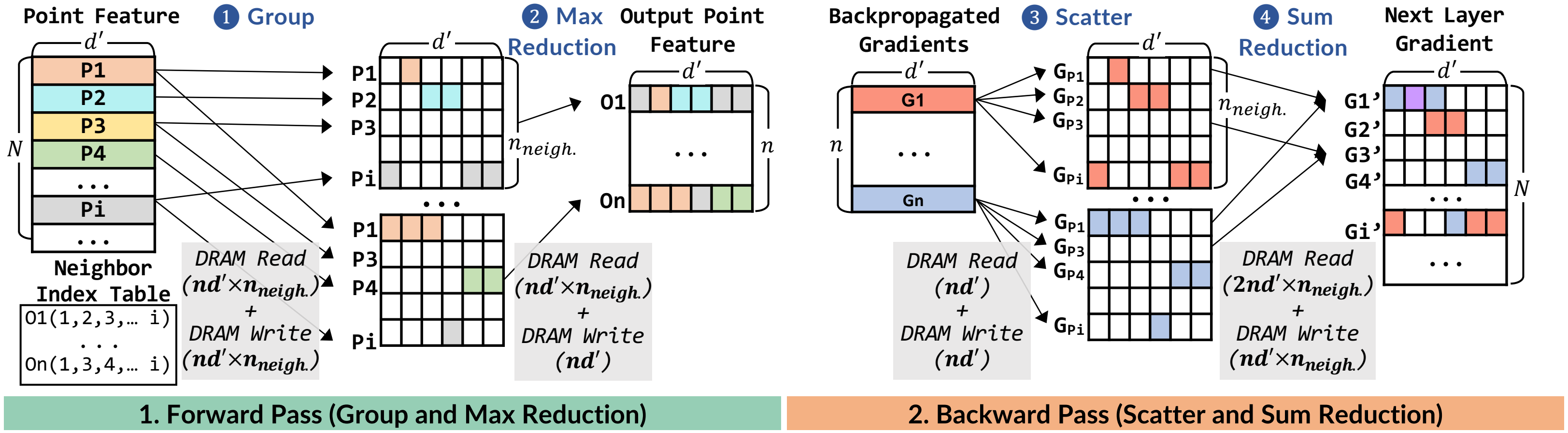
Minimum spacing among the sampled points are ensured.

Enough Randomness

Points are randomly selected every iteration in progressive near point filtering.

Aggregation - Observations

Observation #1. There are **redundant memory accesses** to intermediate values in forward and backward passes.

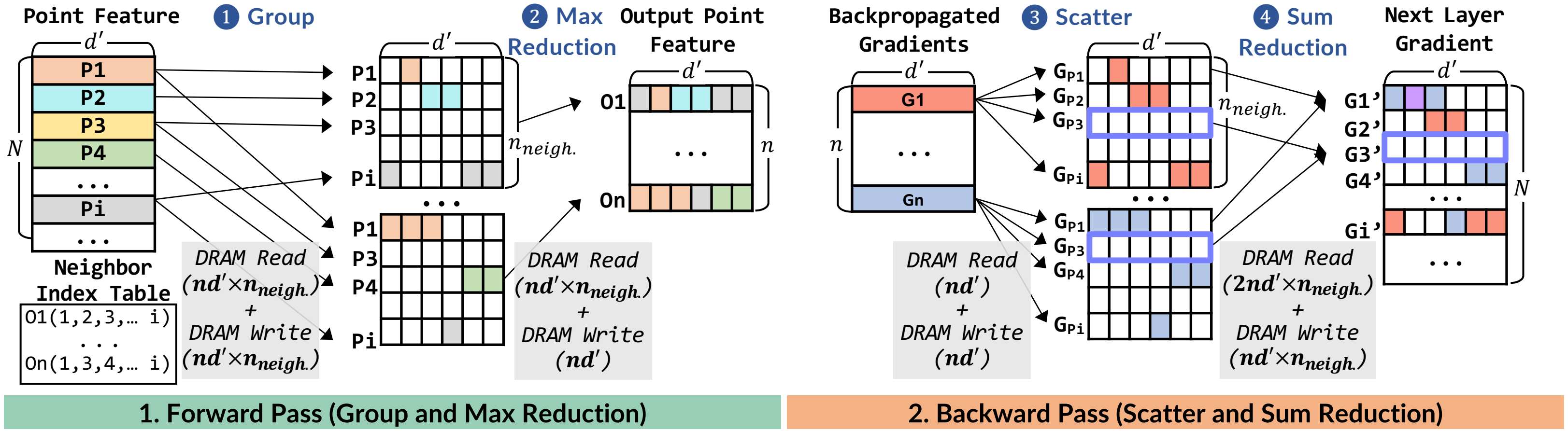


[Forward] "$3nd' \times n_{neigh.} + nd'$" memory access

[Backward] "$3nd' \times n_{neigh.} + 2nd'$" memory access

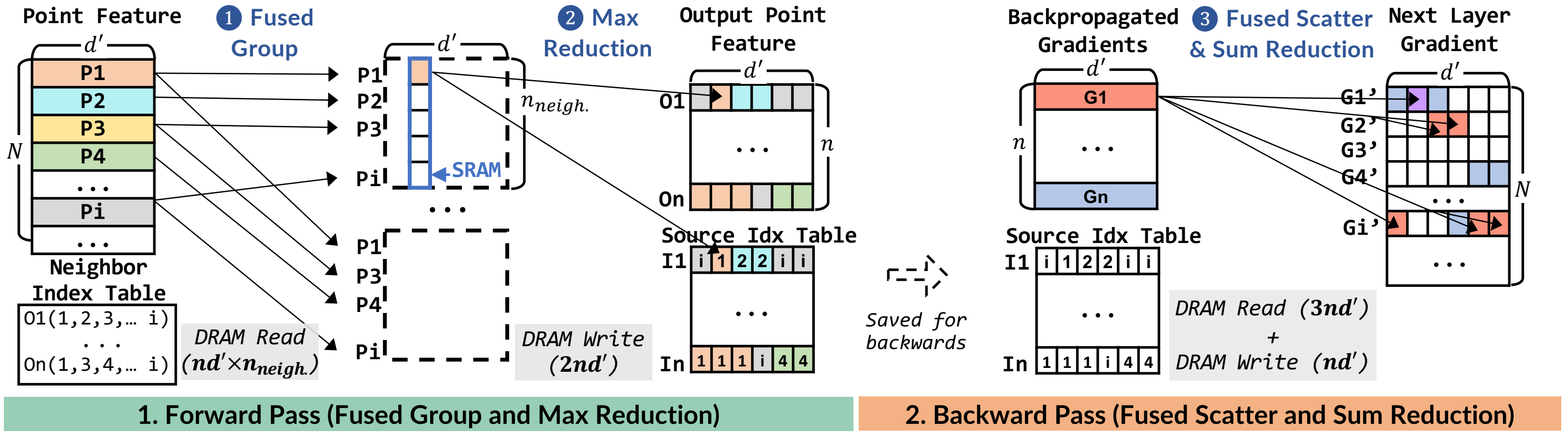
Aggregation - Observations

Observation #2. Ineffectual computations are performed in the backward pass.



Technique #2. Fused Aggregation

- We propose ***Fused Aggregation***, which significantly reduces redundant memory accesses.

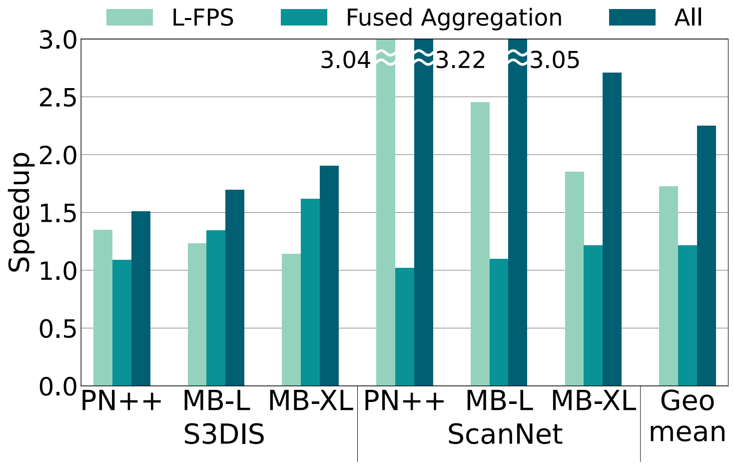


[Forward] Memory access reduced from "$3nd' \times n_{neigh.} + nd'$" to "$nd' \times n_{neigh.} + 2nd'$"

[Backward] Memory access reduced from "$3nd' \times n_{neigh.} + 2nd'$" to "$4nd'$"

Evaluation

Dataset	Model	Accuracy (Stdev.)	
		Baseline	L-FPS
S3DIS	PN++	63.19 (0.54)	63.39 (0.34)
	MB-L	69.82 (0.40)	69.76 (0.40)
	MB-XL	70.67 (0.37)	70.74 (0.43)
ScanNet	PN++	59.42 (0.26)	59.54 (0.57)
	MB-L	70.52 (0.27)	70.54 (0.31)
	MB-XL	71.78 (0.28)	71.74 (0.44)



NVIDIA RTX 3090

Accuracy Max **0.06** mIoU loss, potential mIoU gain of **0.2**

Throughput 2.25x end-to-end speedup

Please contact to the author or refer to the full paper for more details
http://arc.snu.ac.kr/pubs/eccv24_pointcloud.pdf

Frugal 3D Point Cloud Model Training via Progressive Near Point Filtering and Fused Aggregation

Donghyun Lee
eudh1206@snu.ac.kr

Yejin Lee
yejinlee@snu.ac.kr

Jae W. Lee
jaewlee@snu.ac.kr

Hongil Yoon
hongilyoon@google.com

Sourced code available at https://github.com/SNU-ARC/Frugal_PN_Training

Department of CSE
Seoul National University

European Conference on Computer Vision (ECCV 2024), Milan, Italy