

Dataset Distillation of 3D Point Clouds via Distribution Matching

Ulsan National Institute of Science and Technology (UNIST)

Jae-Young Yim, Dongwook Kim, and Jae-Young Sim

Problem Definition

Dataset Distillation

- **Concept**
 - A technique that compresses a large dataset (**original dataset**) into a small set of representative samples (**synthetic dataset**).
 - Models trained on the synthetic dataset achieve comparable performance to those trained on the original dataset.
- **Objectives**
 - **Computational efficiency** — reduce training cost and time
 - **Memory efficiency** — minimize data storage requirements

Problem Definition

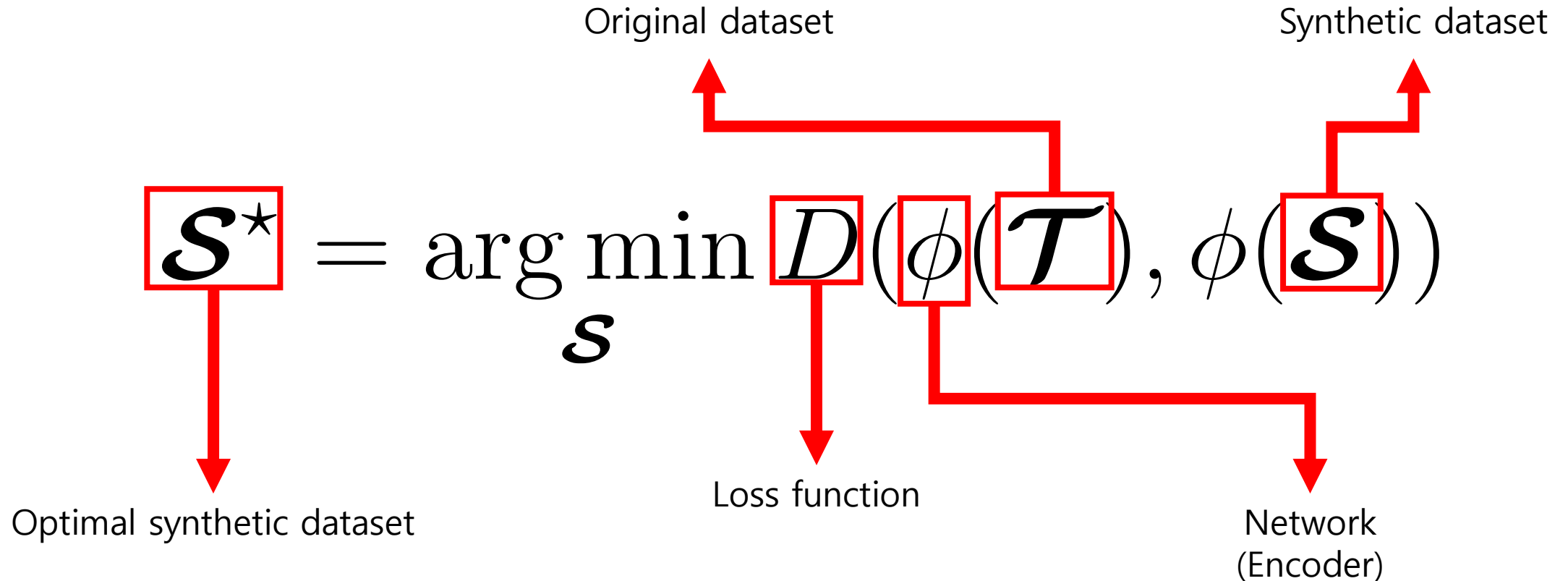
Dataset Distillation

- **Model Parameter Matching**
 - **Gradient Matching** — Align the direction of parameter updates from the synthetic dataset with that from the original dataset.
 - **Trajectory Matching** — Align the optimization trajectory from the synthetic dataset with that of the original dataset.
- **Distribution Matching**
 - Align the data distribution of the synthetic dataset with that of the original dataset in the feature space.

Problem Definition

Distribution Matching

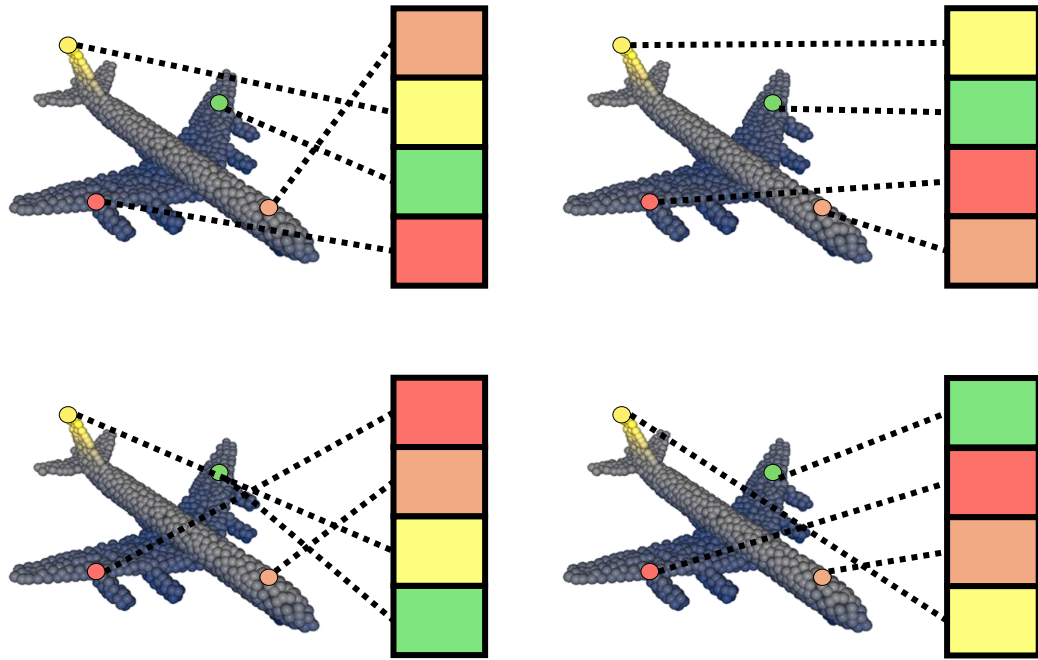
- Given an original dataset and synthetic dataset,



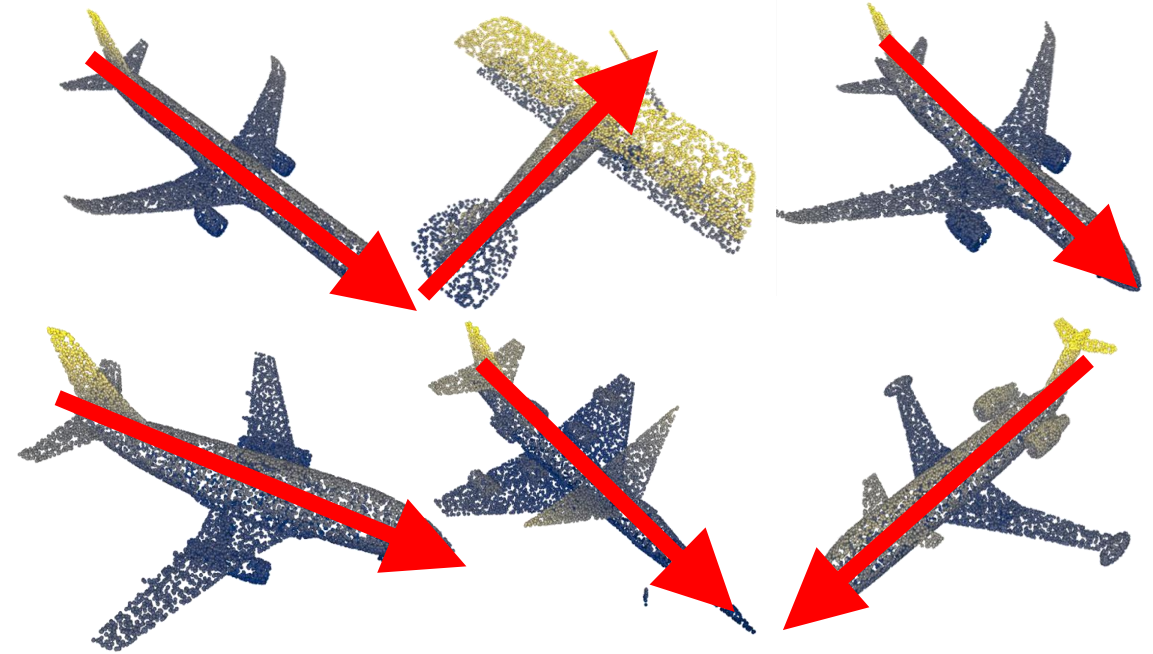
Motivation

5

Characteristics of Point Clouds

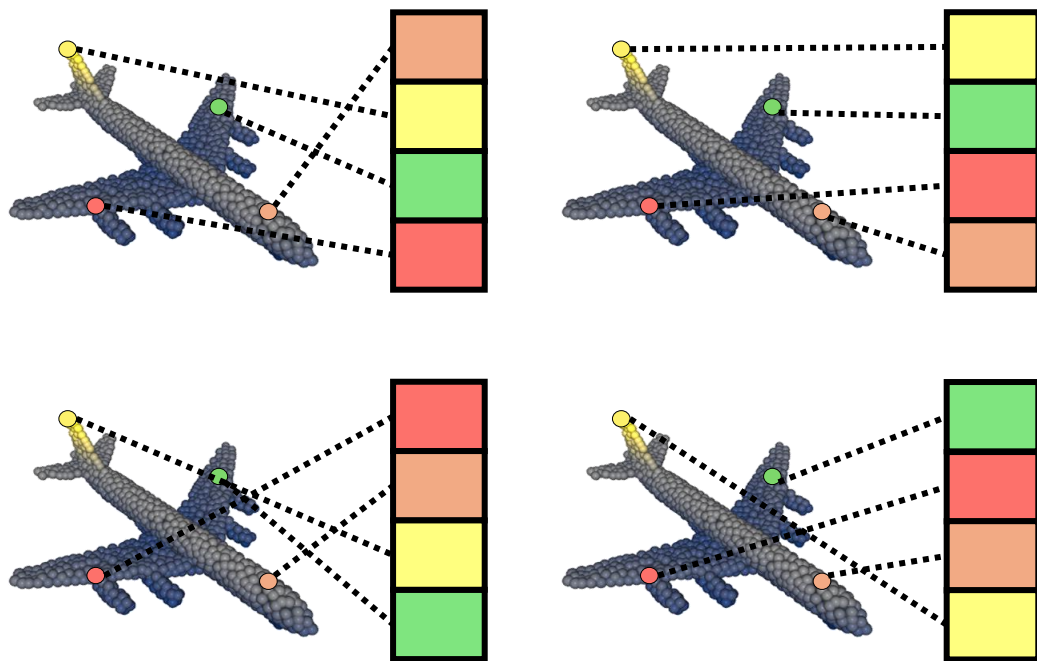


Indexed in different orders

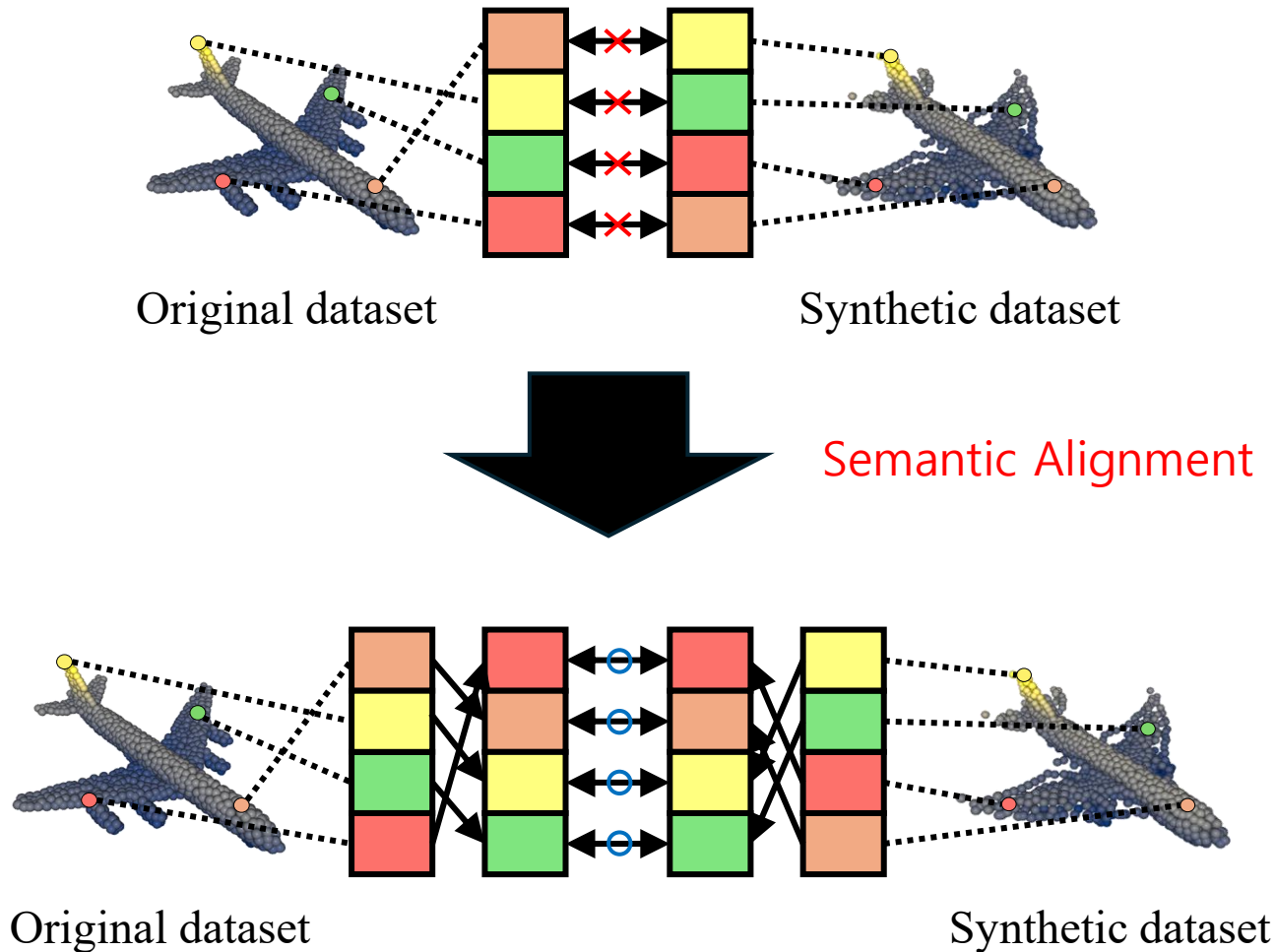


Arbitrary orientation

Semantically Aligned Distribution Matching



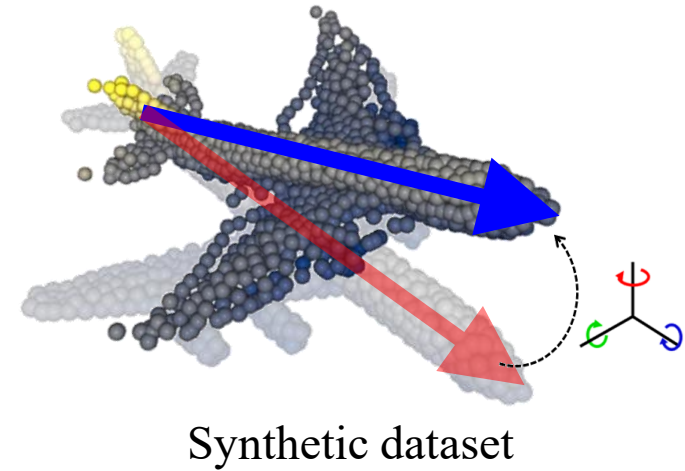
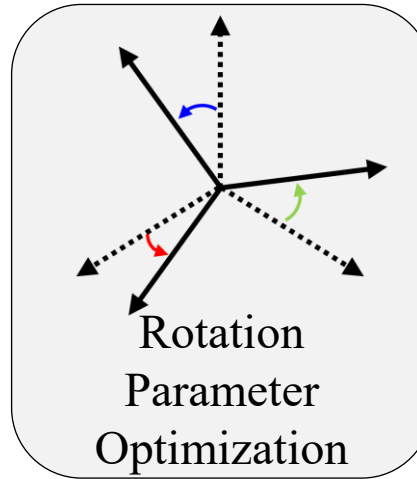
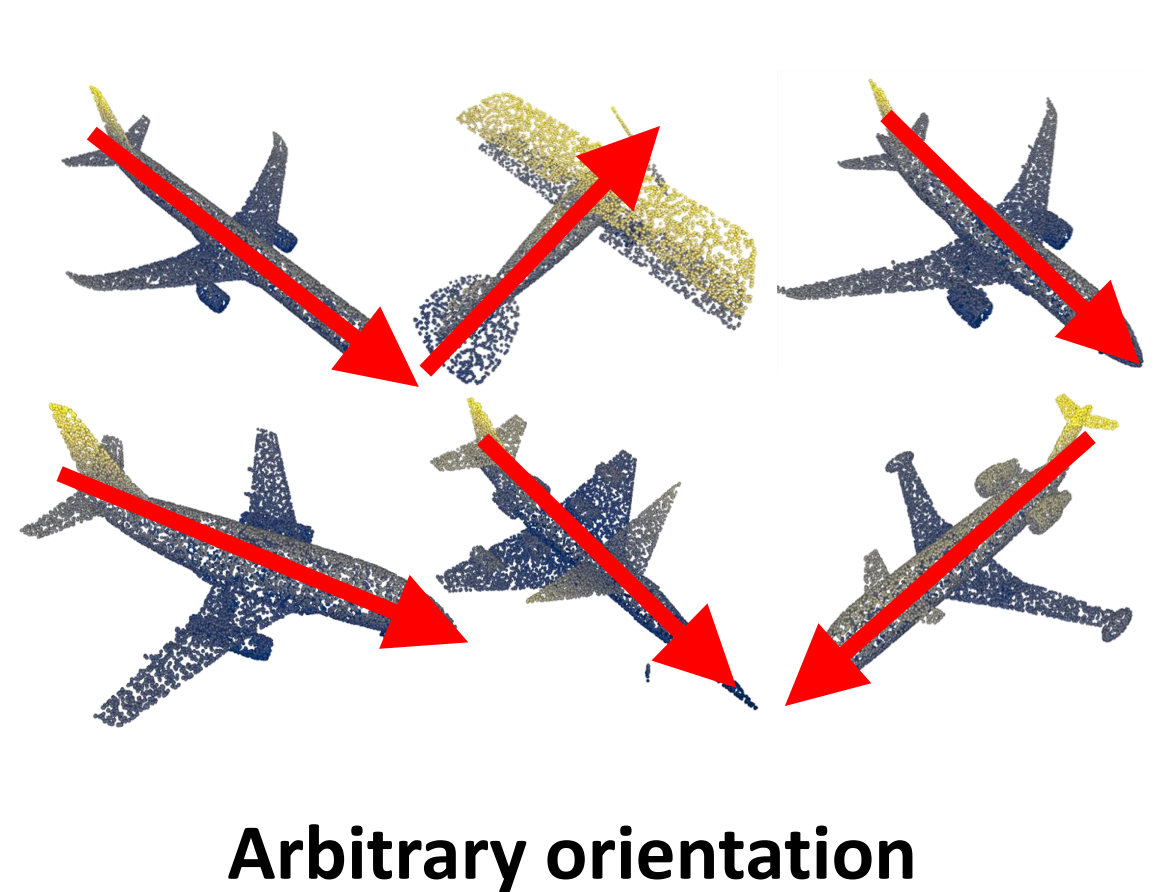
Indexed in different orders



Methods

7

Estimation of Optimal Rotations



Quantitative Results

Table 1

Datasets	ModelNet10			ModelNet40			ShapeNet			ScanObjectNN		
PPC	1	3	10	1	3	10	1	3	10	1	3	10
Ratio	0.25	0.75	2.5	0.4	1.2	4.0	0.15	0.45	1.5	0.15	0.45	1.5
Whole	91.41			87.84			82.49			63.84		
Random	35.5	75.2	85.3	34.6	60.0	74.1	34.0	54.8	63.1	13.9	20.4	34.8
Herding	40.1	78.0	86.9	54.4	68.5	78.8	49.5	59.8	66.9	15.7	27.7	38.7
K-center	40.1	77.6	83.2	54.4	63.0	65.3	49.5	51.4	47.8	15.7	19.8	24.0
DM	31.9	77.6	86.1	32.3	62.0	75.2	27.8	56.1	64.3	14.9	23.3	35.6
DC	43.9	75.0	86.1	52.0	66.6	75.6	49.8	60.1	64.6	18.6	24.3	35.3
Ours	44.7	84.4	87.8	55.8	72.1	80.1	50.2	63.7	68.4	17.4	31.6	43.9

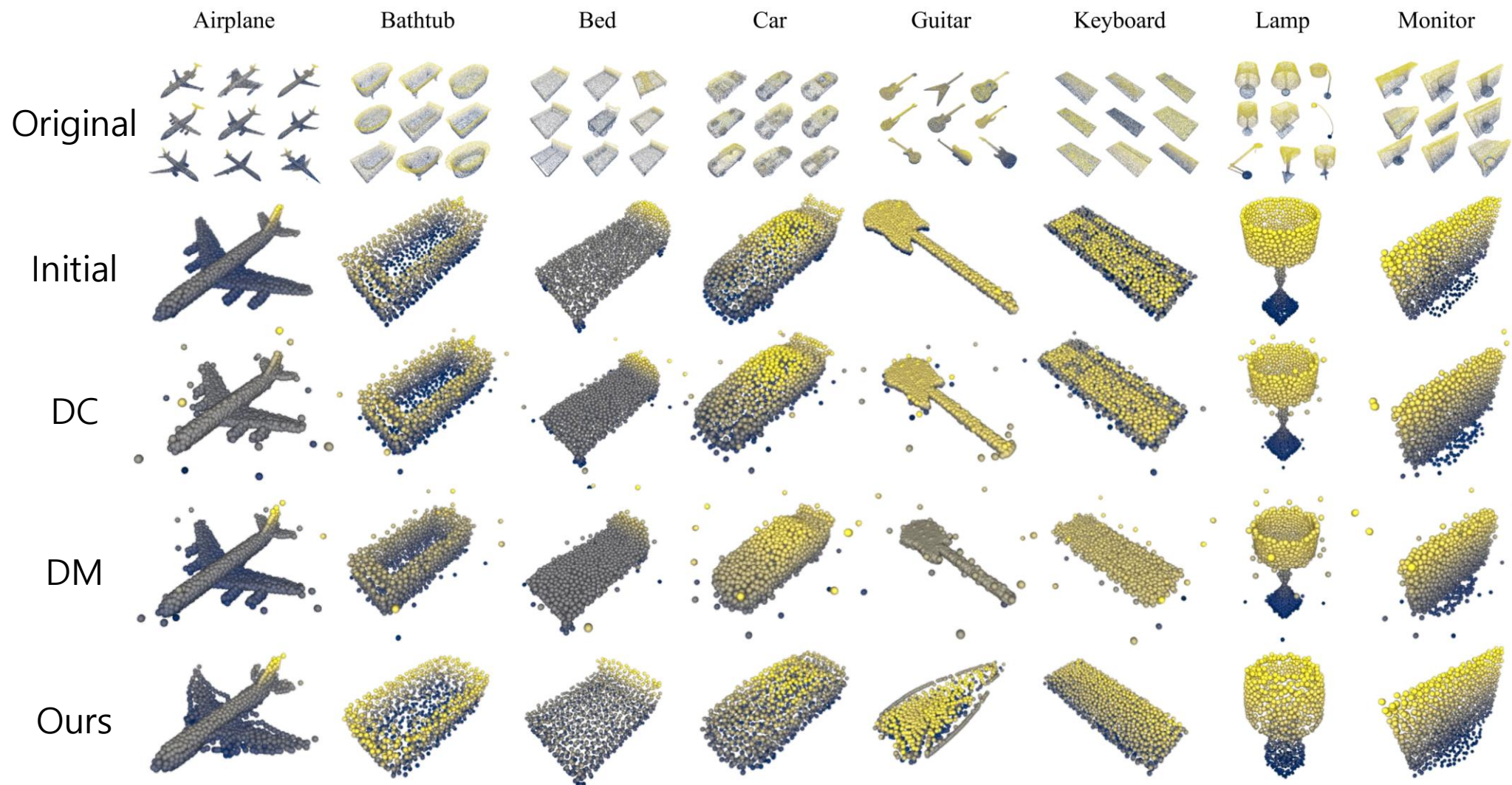
Table 2

Datasets	ModelNet10			ModelNet40			ShapeNet			ScanObjectNN		
Method	DM	DC	Ours	DM	DC	Ours	DM	DC	Ours	DM	DC	Ours
PointNet++	35.0	51.5	82.9	14.1	46.9	68.5	20.3	34.3	53.2	15.5	11.6	22.0
DGCNN	56.6	62.6	79.0	34.8	52.3	66.9	13.8	38.0	52.8	17.7	12.5	19.0
PointConv	33.4	40.0	56.6	20.0	37.9	51.3	17.4	34.0	47.4	15.2	14.2	16.8
PT	48.5	61.8	77.6	26.7	47.5	61.6	34.2	42.6	55.5	17.1	14.9	21.9

Results

9

Qualitative Results



Thank you!