Latent-space GANs for 3D Point Clouds

P. Achlioptas O. Diamanti
I. Mitliagkas L. Guibas
Stanford University

Motivation

- Object-centered 3D data are becoming ubiquitous
- "Pure" form of input, less clutter than images
- Higher dimensionality, irregular structure, and less data compared to images make for a challenging problem
- Good 3D representations important for learning methods

Why Point Clouds?

Shortcomings of Alternatives

- **CAD models** are highly irregular
- **Voxels** provide canonical frame but are sparse and low resolution [Wu et al. 2016]
- **Image-based** methods live in 2D [Su et al. 2015]

Point Cloud Advantages

- Are homogeneous and compact
- Amenable to geometric transformations (e.g. morphable)
- Plethora of range scanning applications

This Work

Latent Representation for Point Clouds

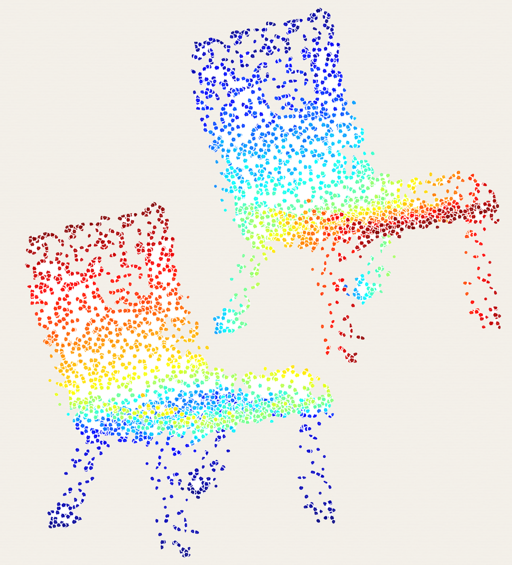
- Powerful Autoencoder-based representation

Generative Models for Point Clouds

- First GAN working in PC space (tedious training & complex architecture)
- Use AE to transform input data and train overly simple GAN in latent space

Structural Losses

Key Challenge



- Point clouds are unordered
- To compare, we need to match them

Two Losses

A. Optimal Assignment

$$d_{EMD}(S_1, S_2) = \min_{\phi: S_1 \to S_2} \sum_{x \in S_1} \|x - \phi(x)\|_2$$

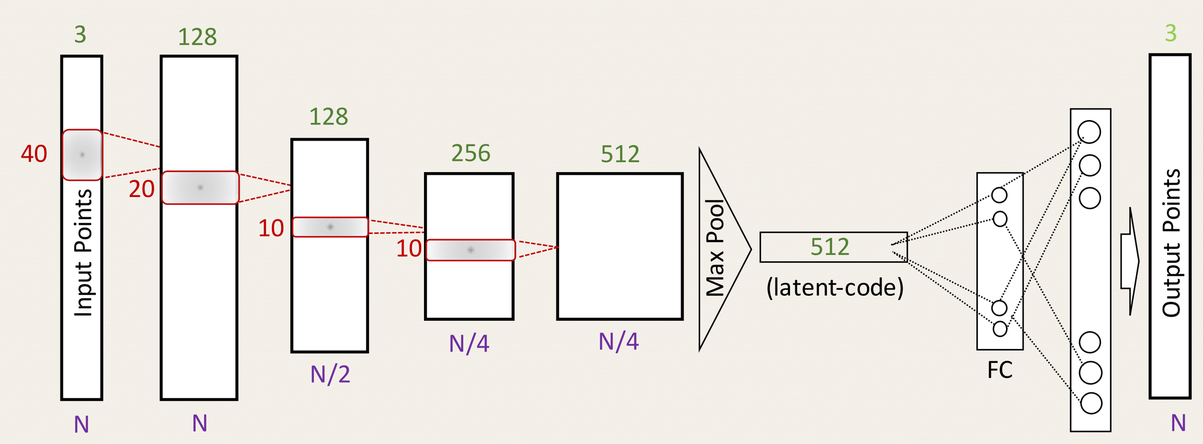
[Fan et al. 2016]

B. Greedy Nearest Neighbor

$$d_{CH}(S_1, S_2) = \sum_{x \in S_1} \min_{y \in S_2} \|x - y\|_2^2 + \sum_{y \in S_2} \min_{x \in S_1} \|x - y\|_2^2$$

Autoencoder

Lex-sort for aligned 3D data enables **1D Convolutions**



1D-Conv: increasing stride and # filters
decreasing filter size

Object Classification Success

Method	MN40	MN10
SPH (Kazhdan et al., 2003)	68.2%	79.8%
LFD (Chen et al., 2003)]	75.5%	79.9%
T-L Network (Girdhar et al., 2016)	74.4%	-
VConv-DAE (Sharma et al., 2016)	75.5%	80.5%
3D-Gan (Wu et al., 2016)	83.3%	91.0%
ours, EMD loss	84.4%	**95.3%**
ours, CD loss	**85.7%**	95.0%

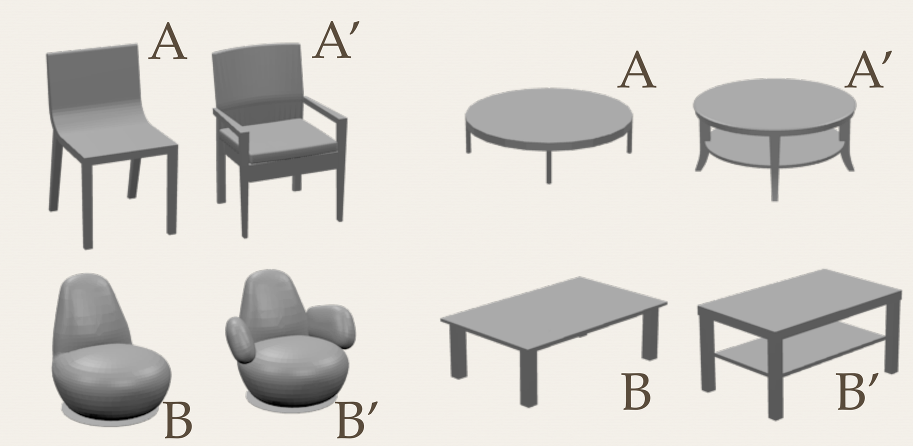
Improve SoA in unsupervised feature learning

Qualitative Results

Latent interpolations between objects

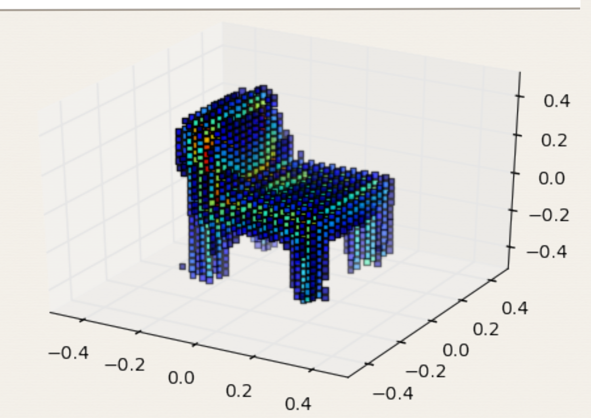
Structural editing: e.g. "add" armrests

Shape analogies



Evaluation Metrics

Distributional Divergence



Jensen-Shannon Divergence in 3D space

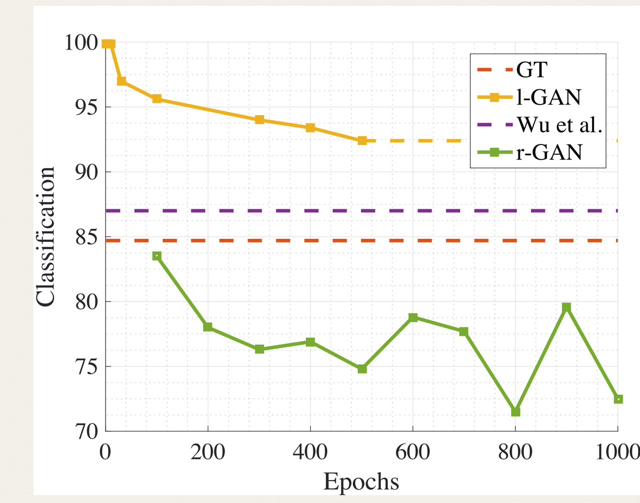
Minimum Matching Distance

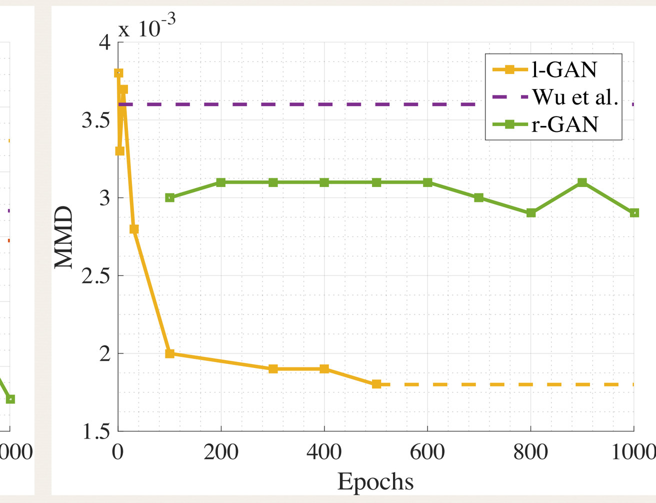
Matching distance of **every** training object and any synthetic one (indicative of mode collapse)

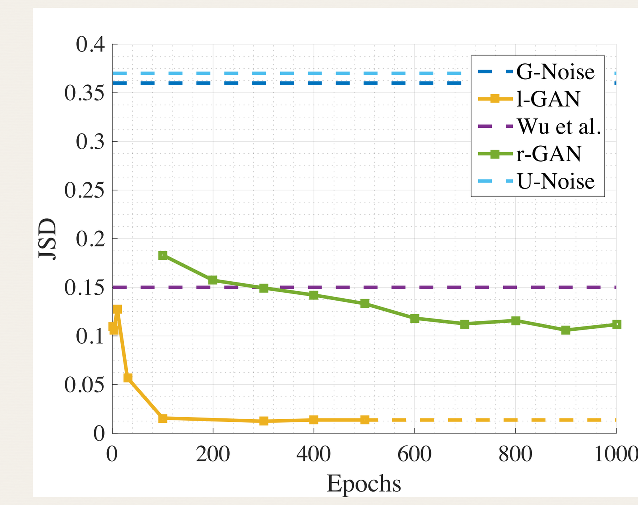
Realism

Average classification probability with SoA point cloud classifier

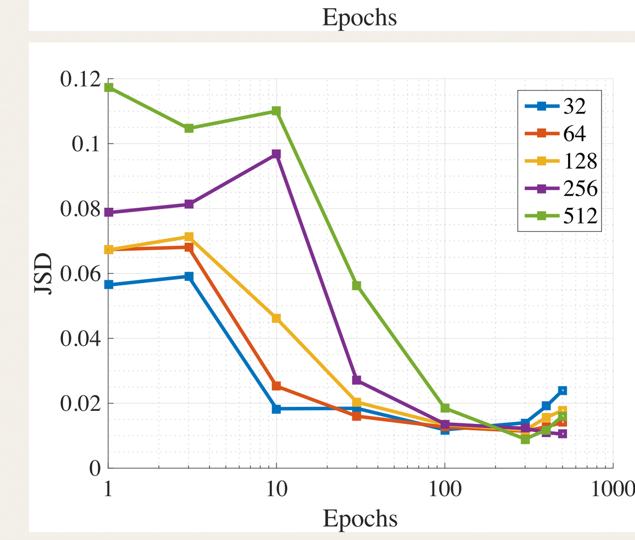
Quantitative Results

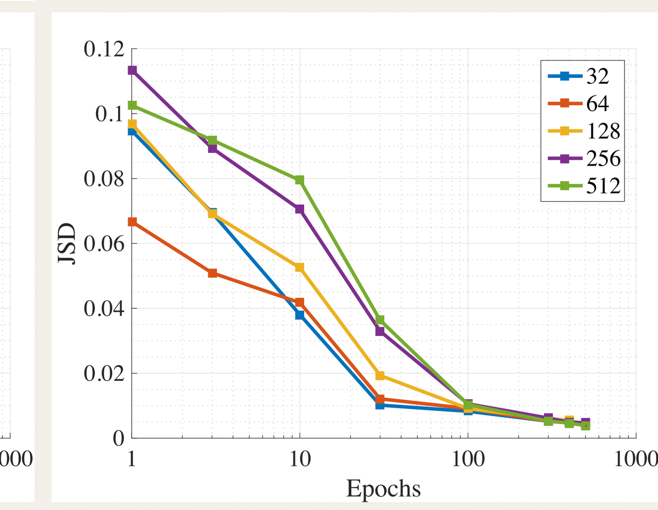






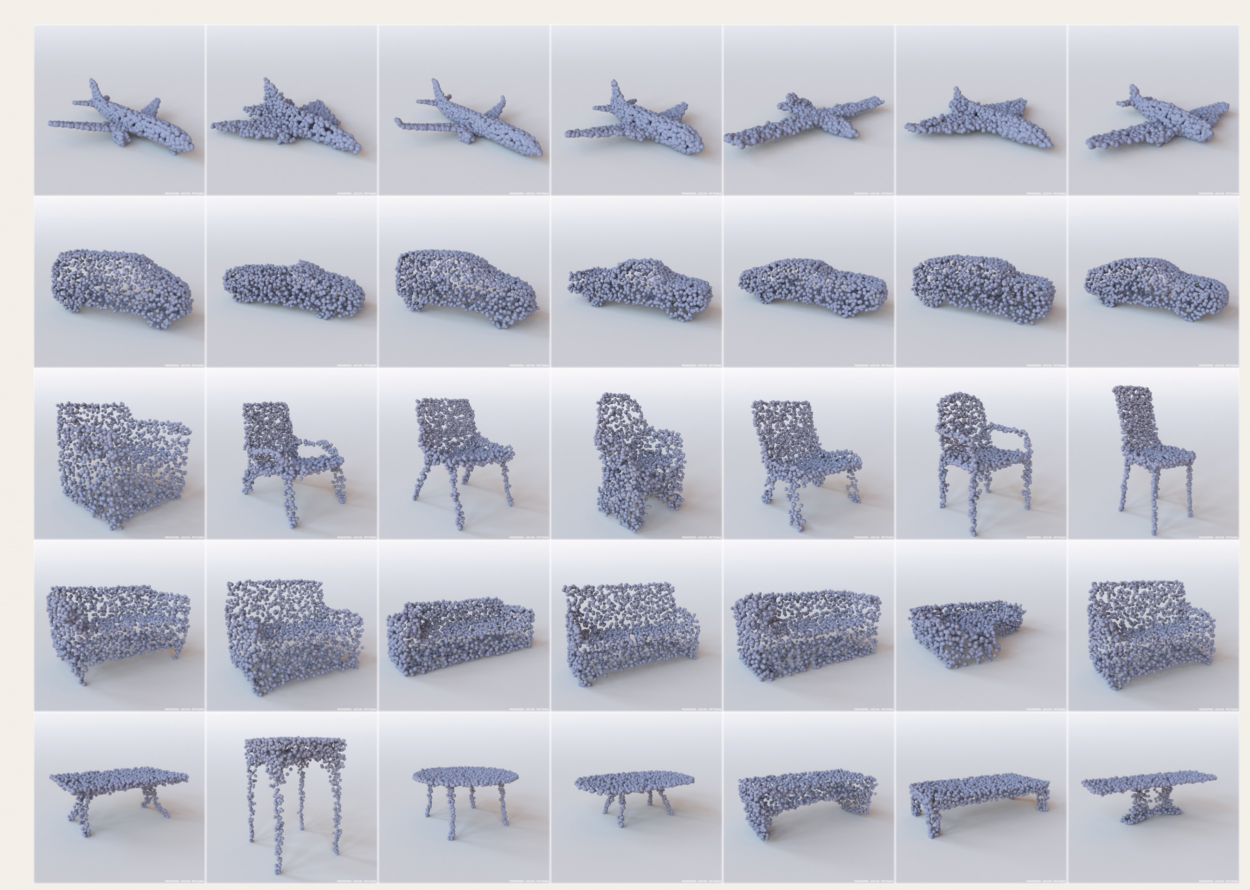
Latent-GANs create **more realistic** point clouds, and **better match** the ground truth distribution





Smaller embeddings (32 vs. 512) trained **faster without hurting fidelity** (GAN & iWGAN)

Generative Results



Project Page: bit.ly/pcgan