Unsupervised 3D Shape Completion through GAN Inversion

Research by Zhang et al.
Presented to CS 348n by David Braun on Feb 14, 2022

ShapeInversion

Very first GAN inversion approach for 3D shape completion with point clouds.

ShapeInversion uses a GAN pre-trained on complete shapes by searching for a latent code that gives a complete shape that best reconstructs the given partial input.

Introduces PatchVariance during GAN pre-training and k-Mask during inversion (inference).

Uses several datasets and benchmarks derived from ShapeNet. Performs well!

Can produce multiple valid complete shapes with good interpolation and generalization.

ShapeInversion takes about 30 seconds per input. The output predicted complete point cloud has 2048 points.

Supervision?

This project is “unsupervised”, not supervised or self-supervised.

Supervised/self-supervised would mean using the same point cloud as the input and output with some data augmentation (randomly deleting sections in the input).

Unsupervised means using unpaired data, e.g., 3D models and point clouds of totally different objects.

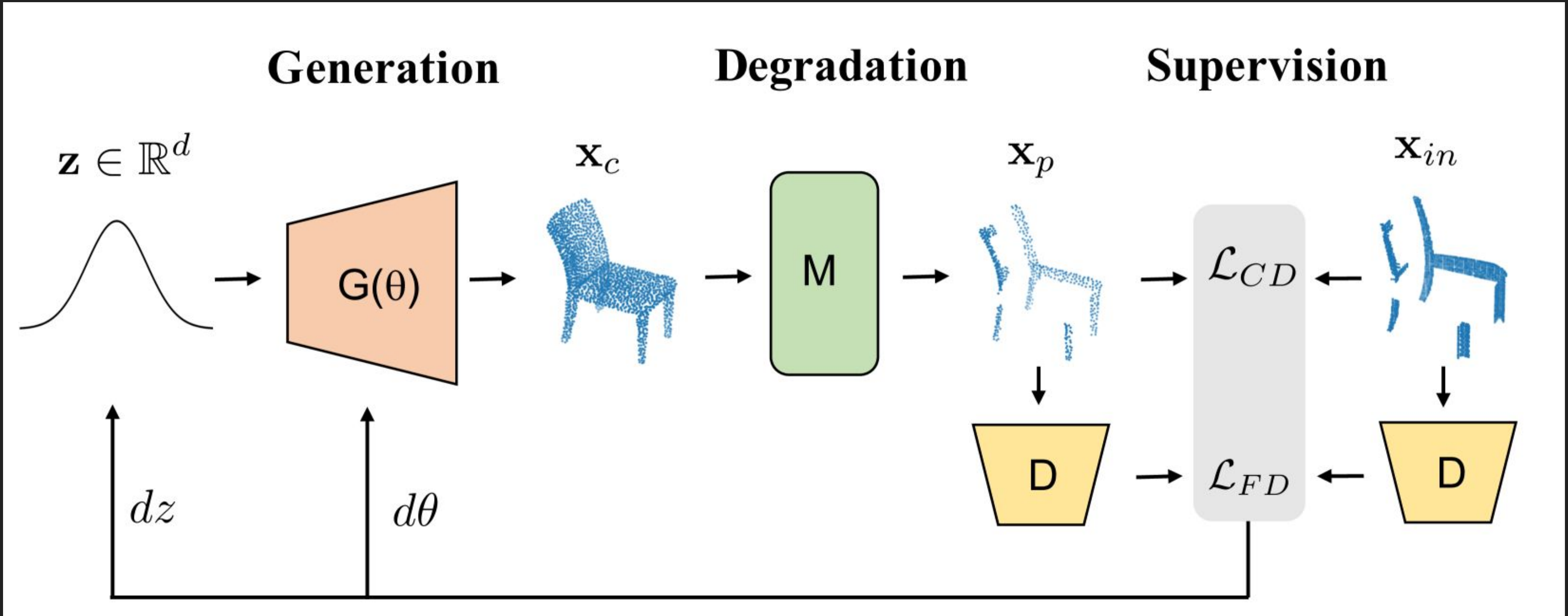
Challenge of applying GAN to 3D

1. Point clouds are unstructured, unlike pixels in a 2D grid.
    a. Previous attempts generated points unevenly distributed over the 3D shape's surface.
2. Point clouds are unordered
    a. For a loss function, making correspondences between two point clouds is tricky.

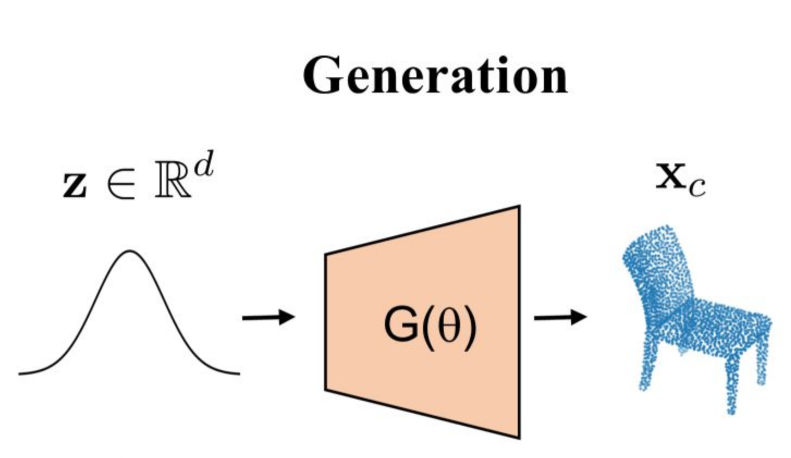
Related Work

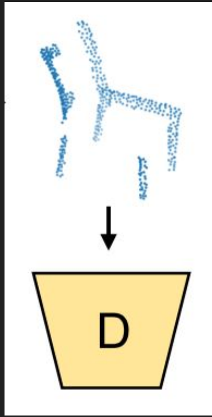
- Morphing Sampling Network (MSN) (Liu, 2019)
  - Known for using “expansion penalty” to try to distribute points uniformly
- Tree-GAN / TreeGCN (Shu, 2019)
  - SOTA Point Cloud Generation
  - But points can be non-uniformly distributed over surface
- Point Completion Network (PCN) (Yuan, 2018)
- Cascaded Refinement Network (CRN) (Wang, 2020)
- Multimodal Shape Completion (MPC) (Wu, 2020)
- pcl2pcl (Chen, 2020)

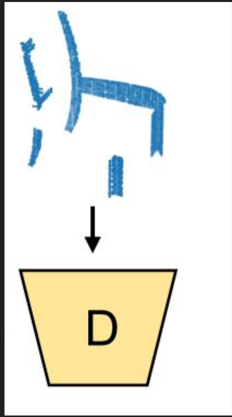
GAN Inversion



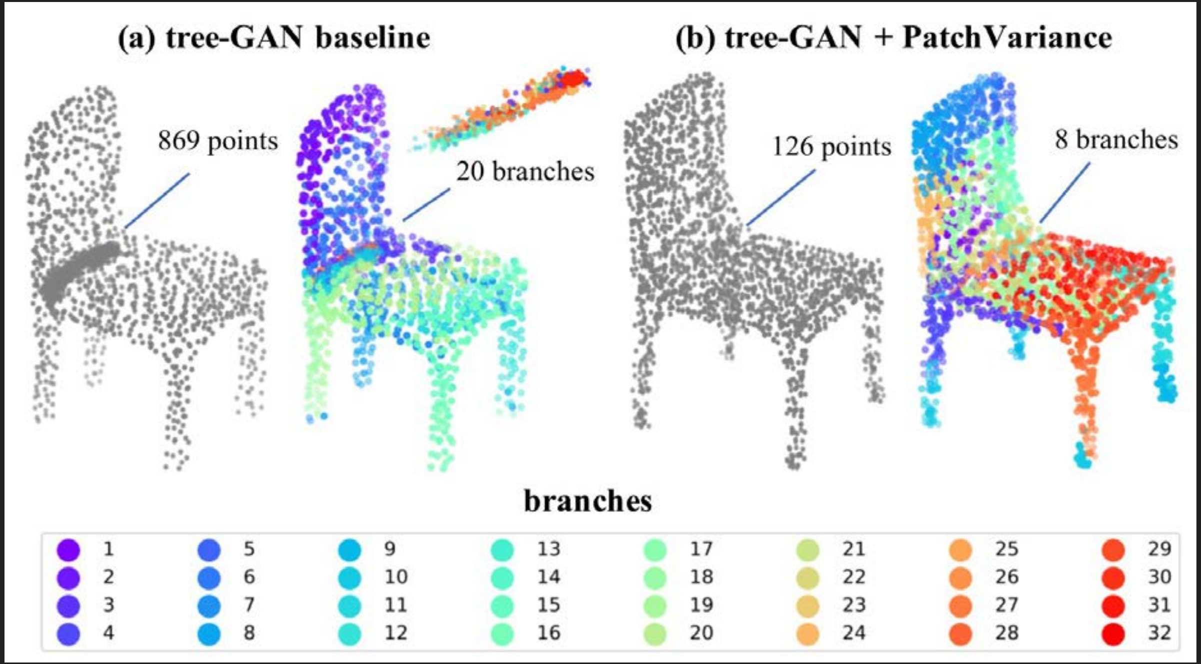
GAN Inversion





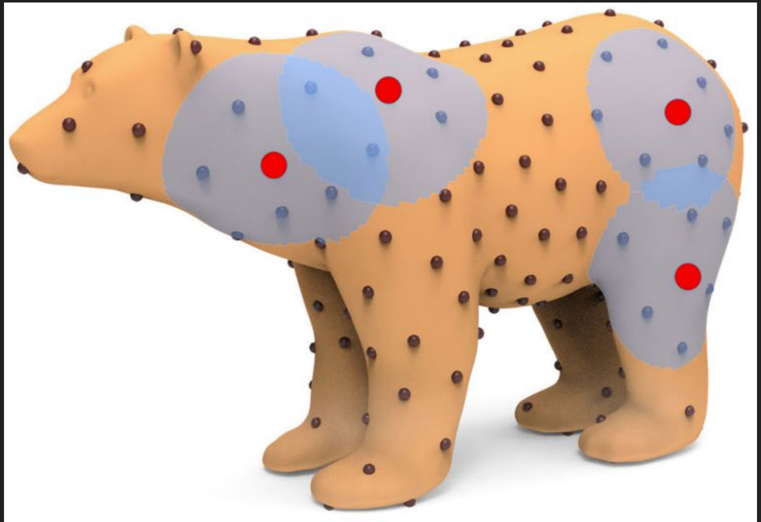


Visualization of uniformity (tree-GAN has a problem)



PatchVariance

Image from PU-GAN (Li, 2019)

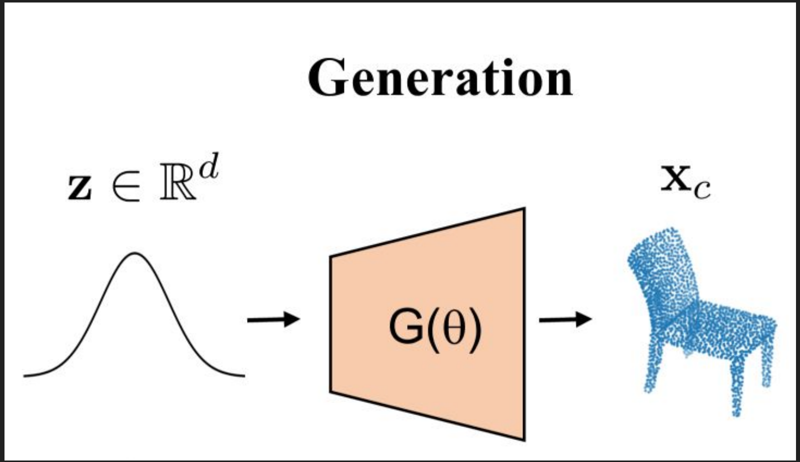


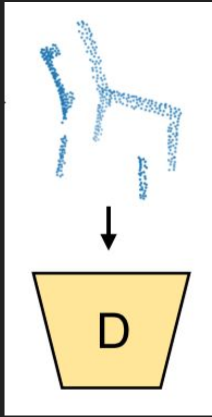
PatchVariance - a new loss which samples small patches

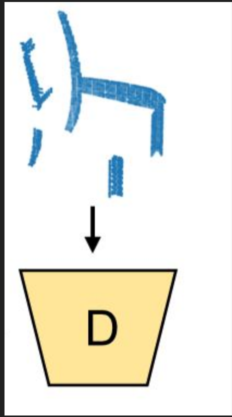
- Penalize the variance of average distances between patch centers and their nearest neighbors
- Regularize the density variance among randomly sampled local patches

$$\mathcal{L}_{patch} = Var(\{\rho_j\}_{j=1}^{n}), \quad \rho_j = \frac{1}{k}\Sigma_{i=1}^{k} dist_{ij}^2$$

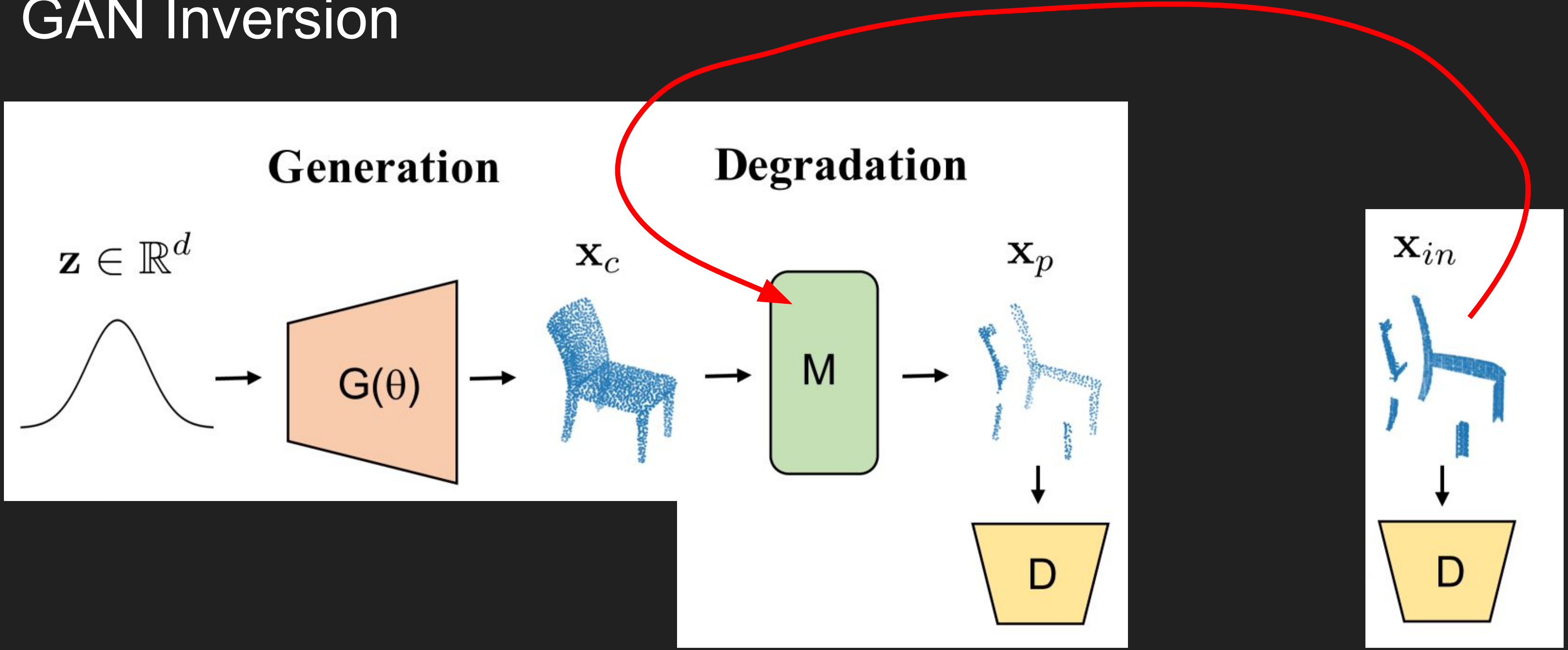
GAN Inversion







GAN Inversion



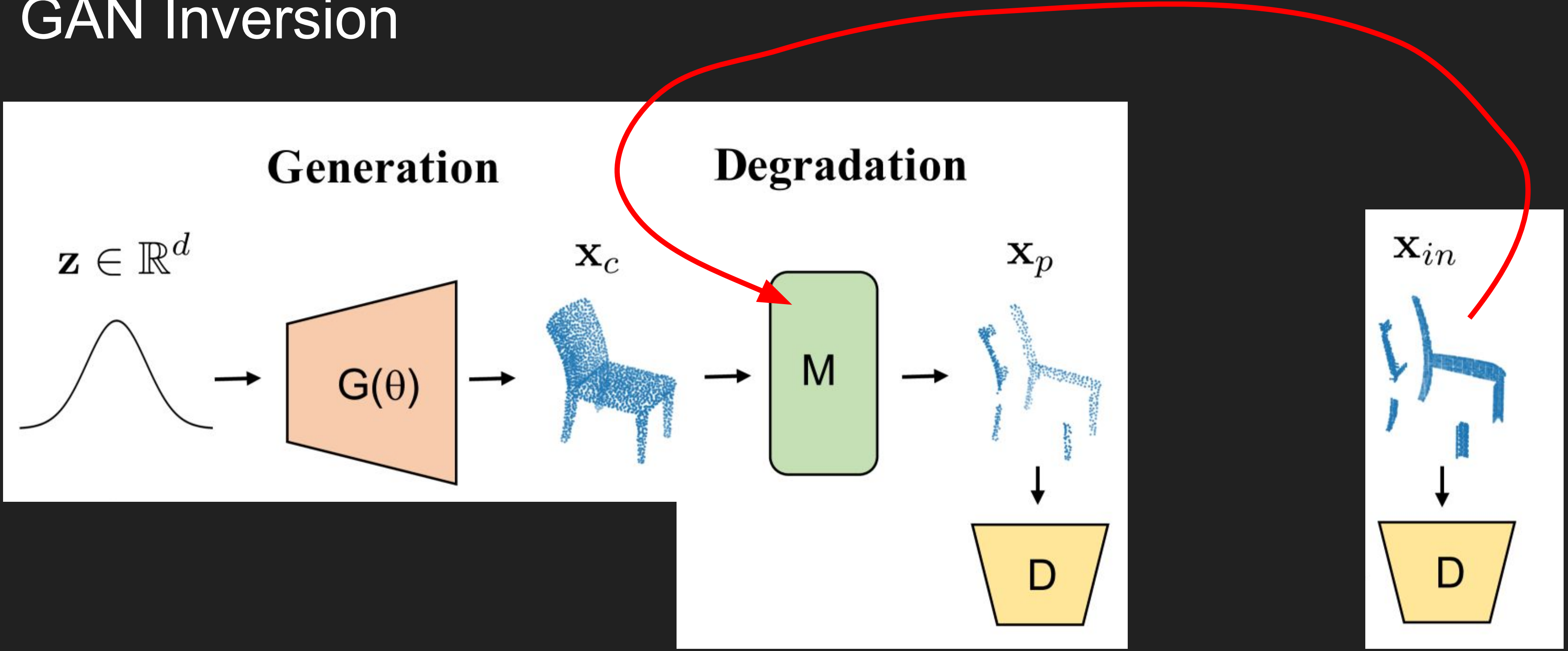
k-Mask Degradation

Estimate the point correspondences between the partial input and predicted shape. For each point in the partial input, we look for its k-nearest neighbors from the complete shape. Then the mask is the union of these k-nearest neighbors.

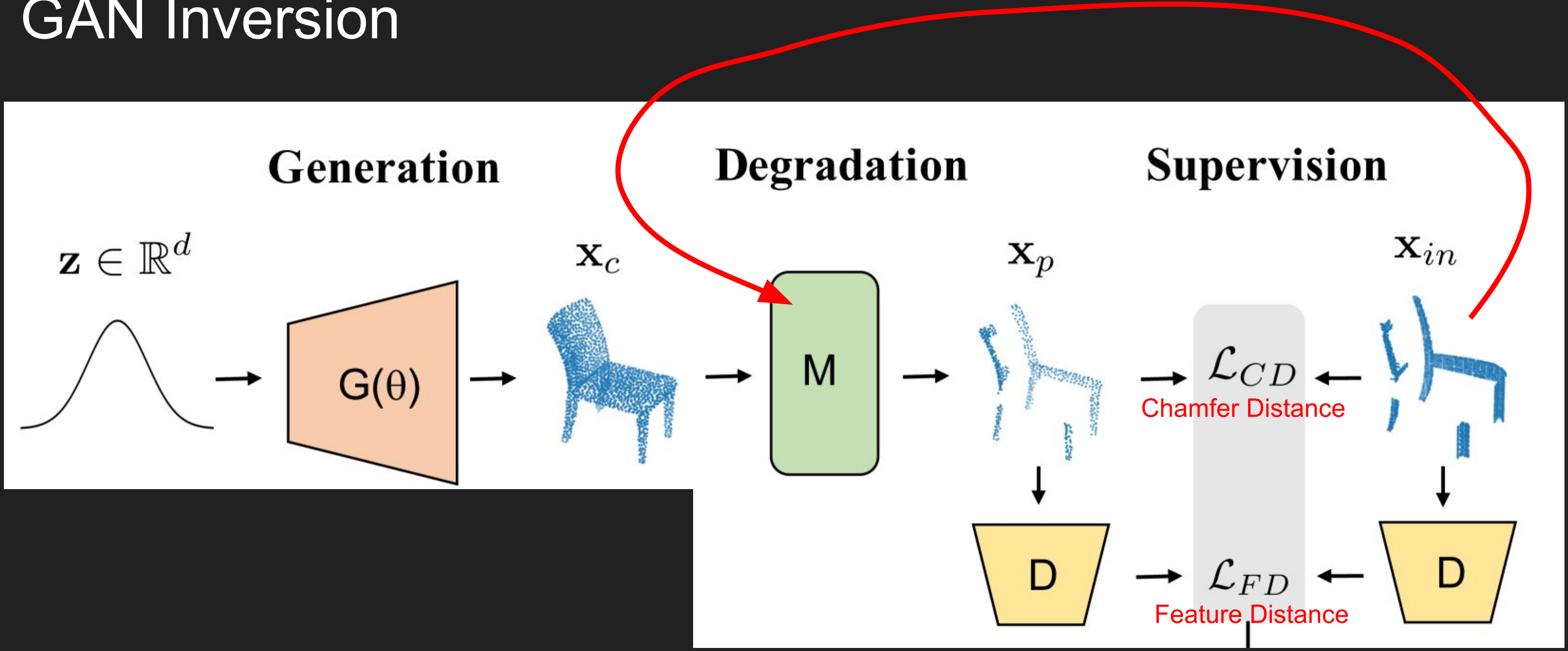
k-Mask is better than voxel-Mask and tau-Mask.

$$\mathbf{x}_p = \bigcup_{i=1}^{n} \{ q_j \in N_k^{\mathbf{x}_c}(p_i) \quad | \quad p_i \in \mathbf{x}_{in} \}$$

GAN Inversion



GAN Inversion



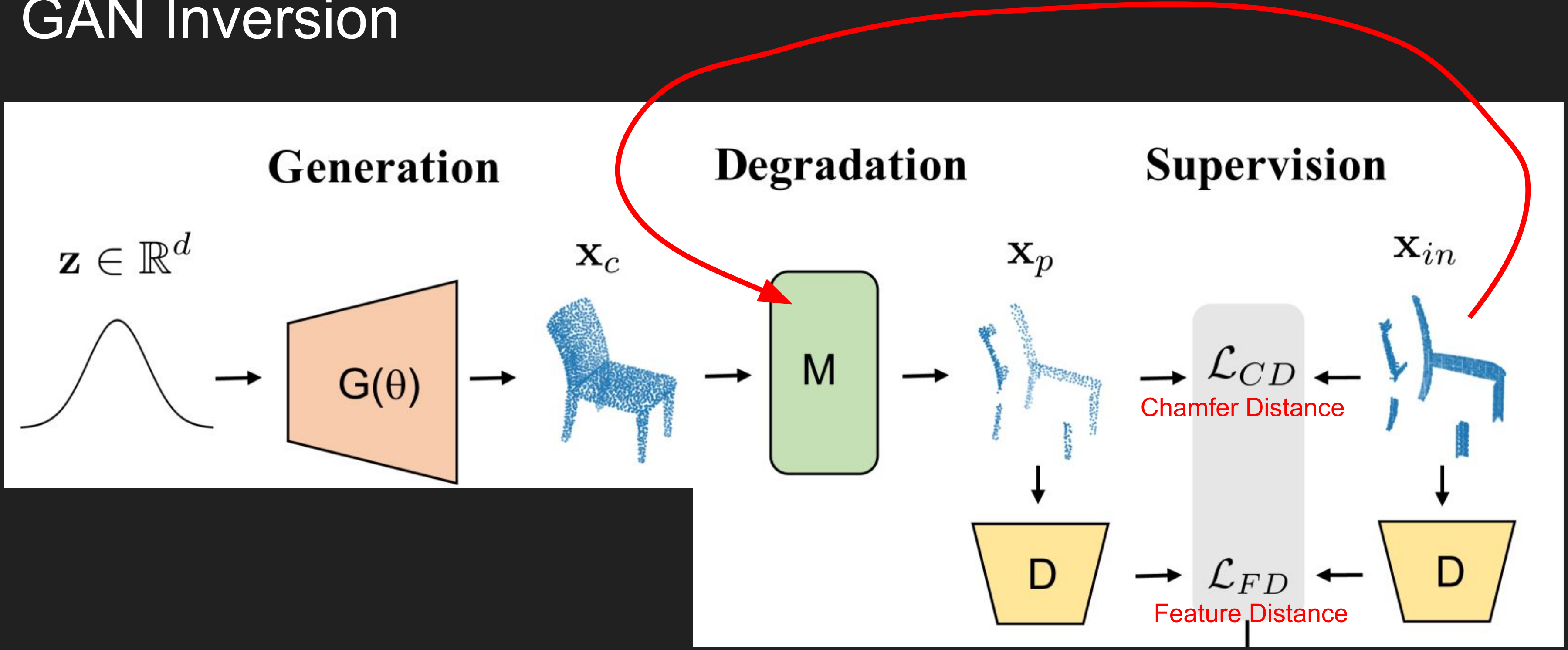
Loss Function for Inversion

$$\mathscr{L}_{CD}(\mathbf{x}_p, \mathbf{x}_{in}) = \frac{1}{|\mathbf{x}_p|} \sum_{p \in \mathbf{x}_p} \min_{q \in \mathbf{x}_{in}} ||p - q||_2^2 + \frac{1}{|\mathbf{x}_{in}|} \sum_{q \in \mathbf{x}_{in}} \min_{p \in \mathbf{x}_p} ||p - q||_2^2$$

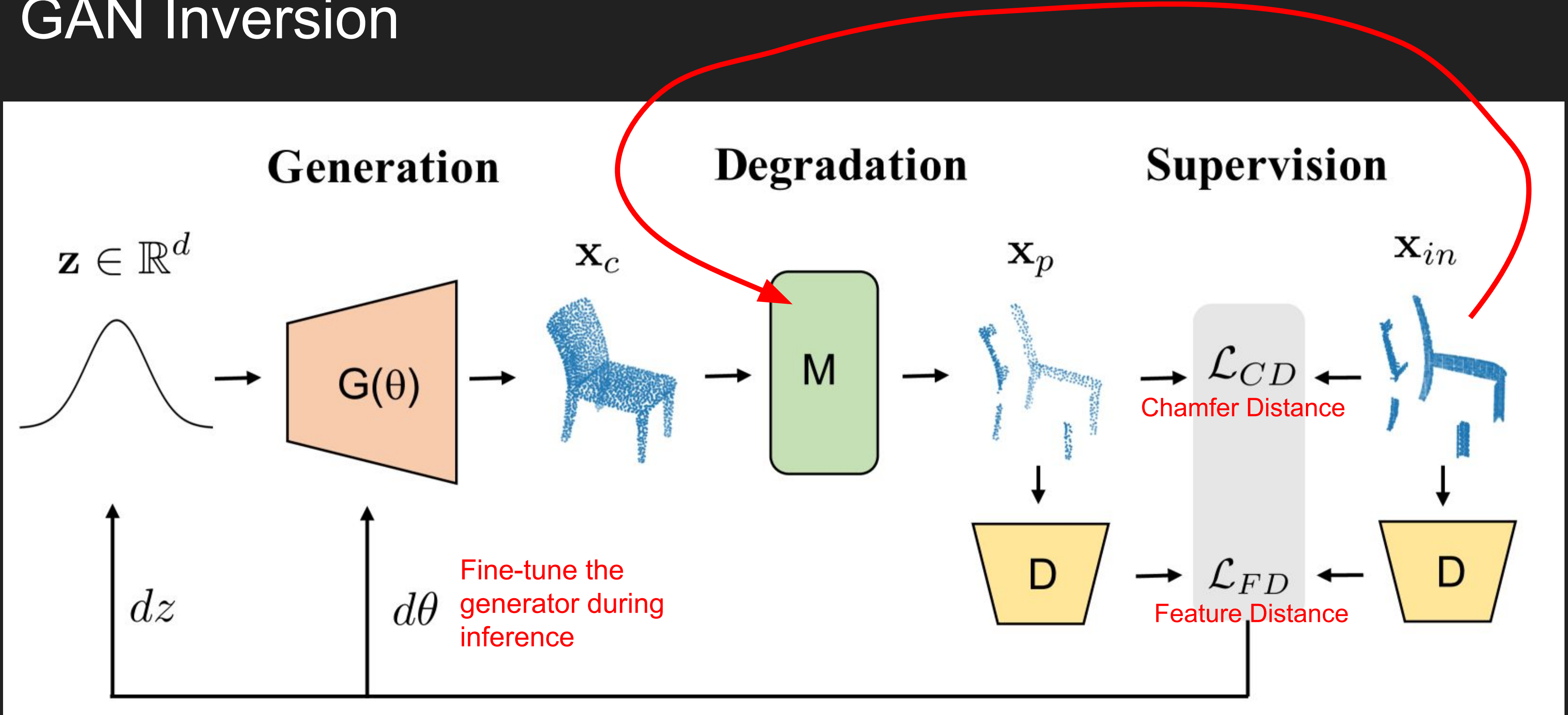
$$\mathscr{L}_{FD} = ||D(\mathbf{x}_p) - D(\mathbf{x}_{in})||_1$$

$$\mathscr{L} = \mathscr{L}_{CD}(\mathbf{x}_p, \mathbf{x}_{in}) + \mathscr{L}_{FD}(\mathbf{x}_p, \mathbf{x}_{in})$$

GAN Inversion



GAN Inversion



Ablation Study

Improvements shown with:

- PatchVariance during pre-training (compared to expansion penalty and repulsion loss)
- k-Mask in inversion stage (compared to voxel-mask and tau-Mask)
- Feature distance in inversion stage (compared to just Chamfer distance)

Effectiveness of k-Mask and Feature Distance

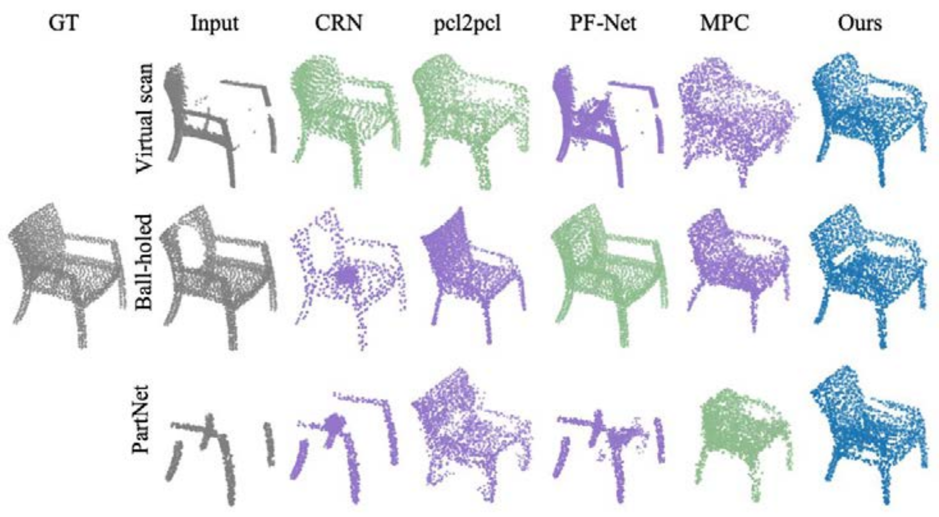
Methods	CD ($\times 10^4$) $\downarrow$	acc. $\uparrow$	comp. $\uparrow$	F1 $\uparrow$
Ours w/ voxel-Mask	19.3	84.7	79.7	81.5
Ours w/ τ-Mask	18.9	82.9	81.6	81.6
Ours w/o $\mathscr{L}_{FD}$	16.3	83.6	81.7	81.9
Ours	**14.9**	**85.0**	**84.0**	**83.9**

Shape Completion Results on ShapeNet

	Methods	Plane	Cabinet	Car	Chair	Lamp	Sofa	Table	Boat	Average
sup.	**PCN [36]**	3.5/96.5	11.3/86.4	6.4/94.0	11.0/86.0	11.6/84.6	11.5/85.2	10.4/89.4	7.4/91.7	9.1/89.2
	TopNet [28]	4.1/96.0	12.9/84.1	7.8/91.3	13.4/82.3	14.8/79.4	16.0/80.8	12.9/85.7	8.9/89.3	11.4/86.1
	MSN [21]	2.9/97.4	12.5/85.5	7.1/92.3	10.6/86.8	9.3/88.6	12.0/83.3	9.6/91.3	6.5/93.1	8.8/89.8
	CRN [30]	2.3/98.3	11.4/86.2	6.2/93.8	8.8/89.7	8.5/90.2	11.3/85.1	9.3/92.9	6.1/94.2	8.0/91.3
unsup.	**pcl2pcl [7]**	9.8/89.1	27.1/68.4	15.8/80.0	26.9/70.4	25.7/70.4	34.1/58.4	23.6/79.0	15.7/77.8	22.4/74.2
	Ours	**5.6/94.3**	**16.1/77.2**	**13.0/85.8**	**15.4/81.2**	**18.0/81.7**	**24.6/78.4**	**16.2/85.5**	**10.1/87.0**	**14.9/83.9**

Robustness to Varying Partial Forms from ShapeNet

The metric is Chamfer Distance

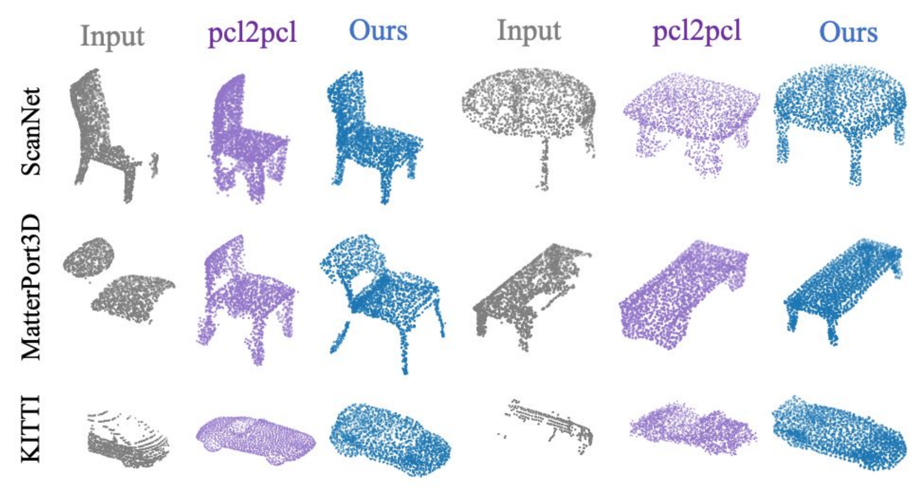


Target	Methods	Source	Chair	Table	Lamp
Virtual scan	CRN	Virtual scan	8.8	9.3	8.5
	MPC [33]	PartNet	45.9	88.9	63.0
	Ours	-	**15.4**	**16.2**	**18.0**
Ball-holed	PF-Net [15]	Ball-holed	11.9	9.9	23.1
	MSN	Virtual scan	79.6	46.6	55.4
	CRN	Virtual scan	44.7	52.9	52.1
	pcl2pcl	Virtual scan	18.6	18.5	21.2
	MPC	PartNet	44.7	28.9	69.5
	Ours	-	**10.1**	**16.0**	**17.3**
PartNet	MPC	PartNet	40.0	51.0	82.0
	MSN	Virtual scan	198.0	143.2	229.9
	CRN	Virtual scan	177.4	140.6	185.9
	pcl2pcl	Virtual scan	51.0	**76.6**	111.2
	Ours	-	**36.8**	77.8	**100.8**

Robustness on Real World Scans

UCD = Unidirectional Chamfer Distance

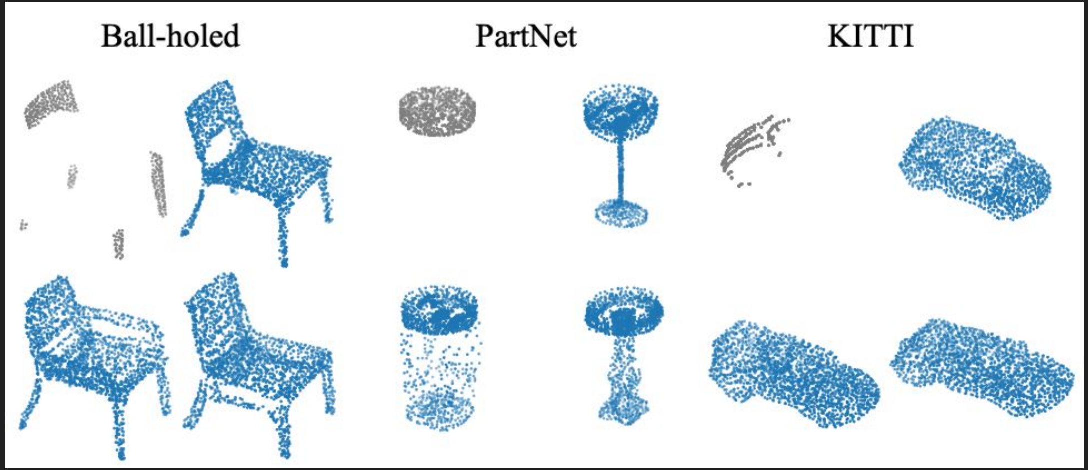
UHD = Unidirectional Hausdorff Distance



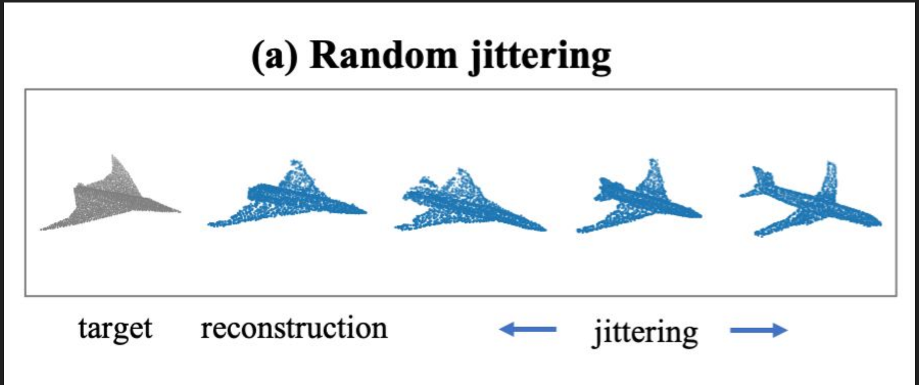
Methods	ScanNet		MatterPort3D		KITTI
	Chair	Table	Chair	Table	Car
pcl2pcl	17.3/10.1	9.1/11.8	15.9/10.5	6.0/11.8	9.2/14.1
Ours	**3.2**/10.1	**3.3**/11.9	**3.6**/10.0	**3.1**/11.8	**2.9**/13.8
Ours+UHD	4.0/**9.3**	6.6/**11.0**	4.5/**9.5**	5.7/**10.7**	5.3/**12.5**

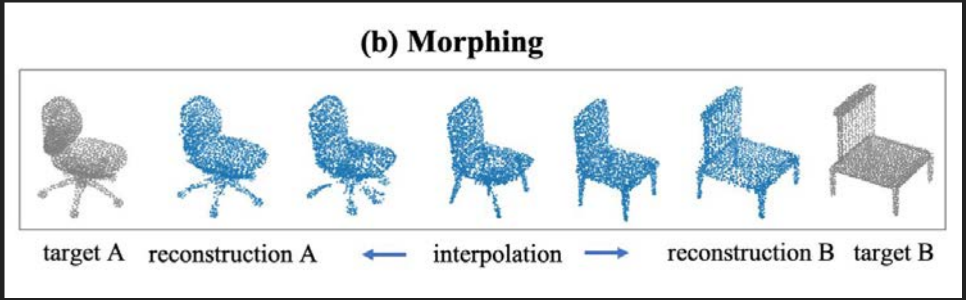
Multiple Valid Outputs under Ambiguity

In inversion initialization step, simply randomly sample more latent vectors to find more which have a desirably low initial starting loss.



Manipulation of Complete Shapes





Future Work

Improve the fidelity of multi-class models, which could provide more possibilities such as cross-category shape completion via a conditional GAN.