Generation For Adaption: A GAN-Based Approach for Unsupervised Domain Adaption with 3D Point Cloud Data

Junxuan Huang Junsong Yuan Chunming Qiao

State University of New York at Buffalo

{junxuanh, jsyuan ,qiao}@buffalo.edu

Background

Recent deep networks have achieved good performance on a variety of 3d points classification tasks. However, these models often face challenges in "wild tasks" where there are considerable differences between the labeled training/source data collected by one Lidar and unseen test/target data collected by a different Lidar.

Unsupervised domain adaptation (UDA) seeks to overcome such a problem without target domain labels. Instead of aligning features between source data and target data, we propose a method that uses a Generative Adversarial Network (GAN) to generate synthetic data from the source domain so that the output is close to the target domain.

Experiments show that our approach performs better than state-of-the-art UDA methods in three popular 3D object/scene datasets

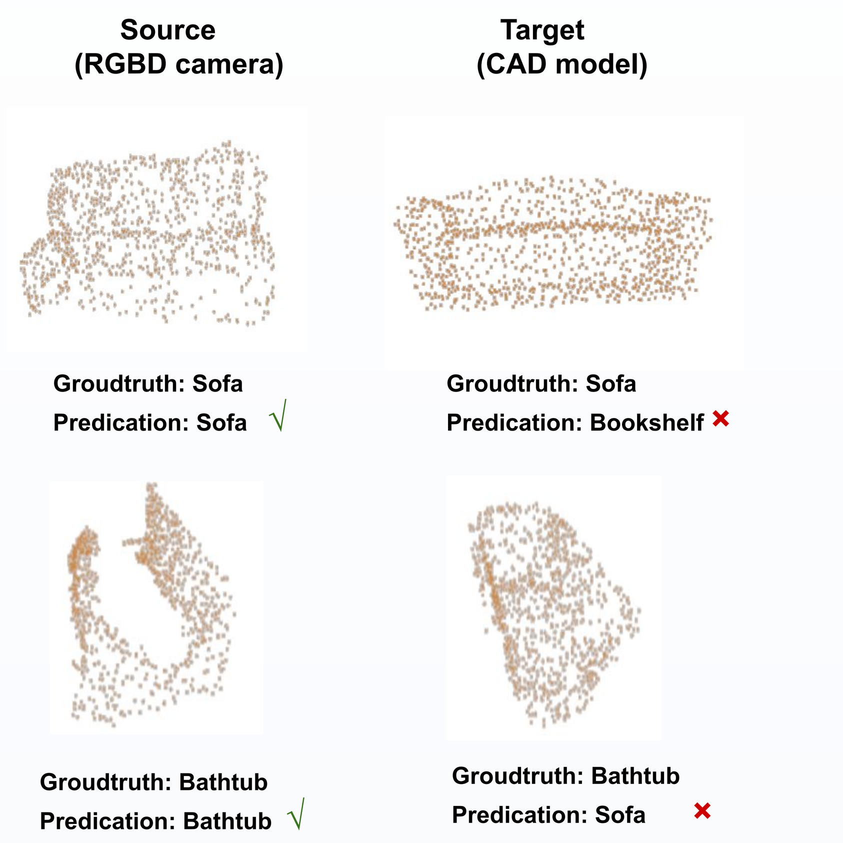
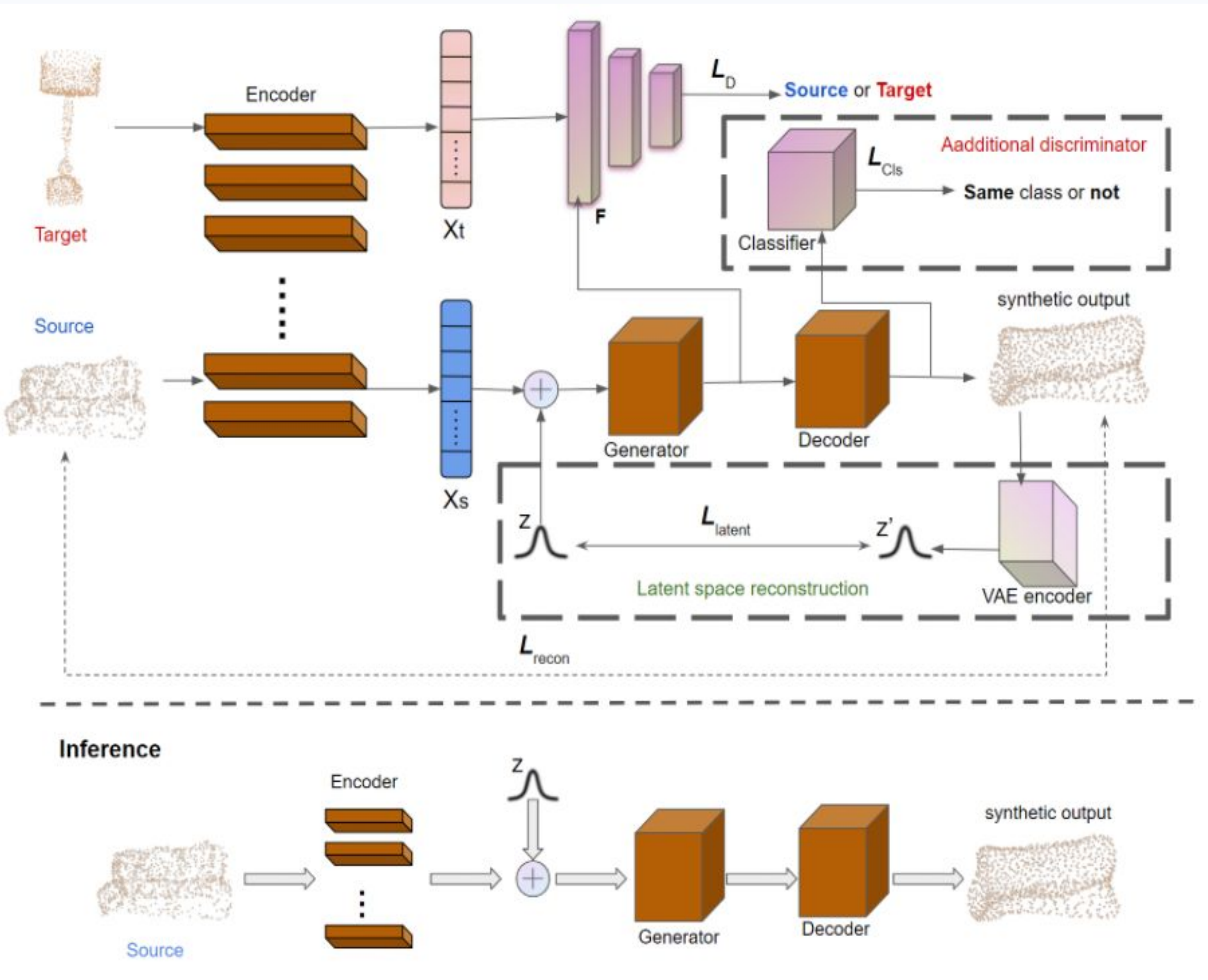


Fig. 1. Point clouds acquired from different sensors and being misclassified

Proposed GAN-based DA method



Our model architecture consists of four parts: **generator**, **latent reconstruction module**, **discriminator**, and **feature encoder/decoder**.

In the training, source domain object and target domain object will go through a shared encoder. The encoded features from source domain will be sent to the generator and a discriminator tries to distinguish features from generator or target domain. For adding multimodal information to the model, we also have Gaussian samples *z* for latent condition input to the generator. To force the generator to use the Gaussian samples *z* ,we introduce a VAE encoder to recover z from the synthetic output. In addition, in order to enhance the quality of output object from *G*, we have an additional discriminator, a classifier *C* in training the model.

To reconstruct the shape of point clouds object we choose Earth Mover's Distance (EMD) to measure the distance between reconstructed object and input object. So that we can restrict the synthetic outputs and make it close to the input's shape. But we do not want the synthetic object to having the exact shape of input. Because we are building a synthetic dataset which means the variety is also significant. So we bring a random sampled variable z into our model and a Variational Autoencoder (VAE) is trained to encode synthetic objects to recover latent input vector, encouraging the use of conditional mode input z.

Experiments results

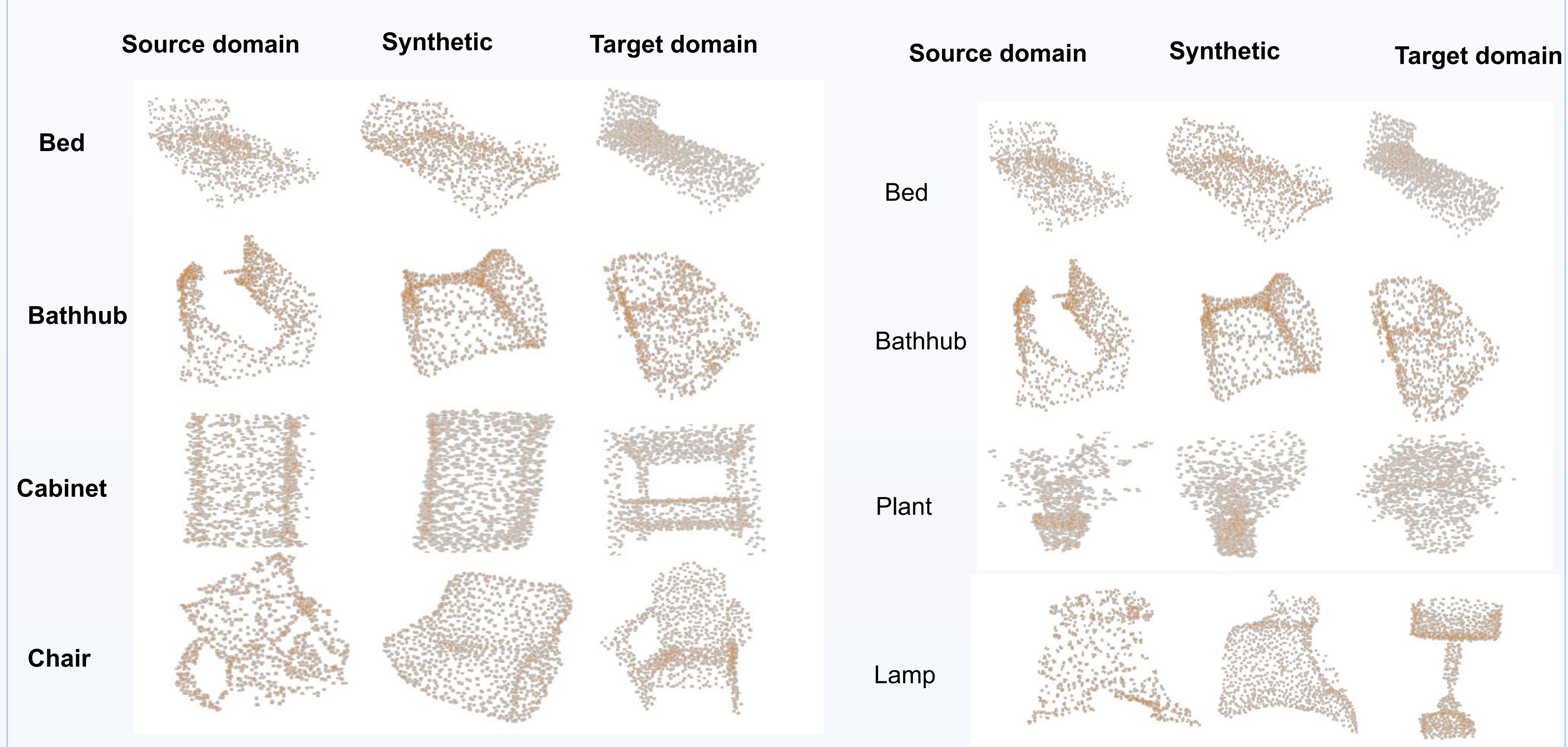
Quantitative classification results (%) on PointDA-10 Dataset

	M→S	M→S*	S→M	S→S*	S*→M	S*→S	Avg
w/o Adapt	42.5	22.3	39.9	23.5	34.2	46.9	34.9
MMD[19]	57.5	27.9	40.7	26.7	47.3	54.8	42.5
DANN[4]	58.7	29.4	**42.3**	30.5	48.1	56.7	44.2
ADDA[5]	61.0	30.5	40.4	29.3	48.9	51.1	43.5
MCD[6]	62.0	31.0	41.4	31.3	46.8	59.3	45.3
PointDAN[7]	62.5	31.2	41.5	31.5	46.9	59.3	45.5
Ours	**62.8**	**36.5**	41.9	**31.6**	**50.4**	**65.7**	**48.1**
Supervised	90.5	53.2	86.2	53.2	86.2	90.5	76.6

M means ModelNet and S denotes ShapeNet while S* represents ScanNet.

We consider six types of adaptation scenarios which are **M→S, M → S*, S →M, S → S*, S*→M and S*→ S,** where **M, S** and **S*** represent subset of Modelnet, Shapenet and Scannet respectively.

Visualization



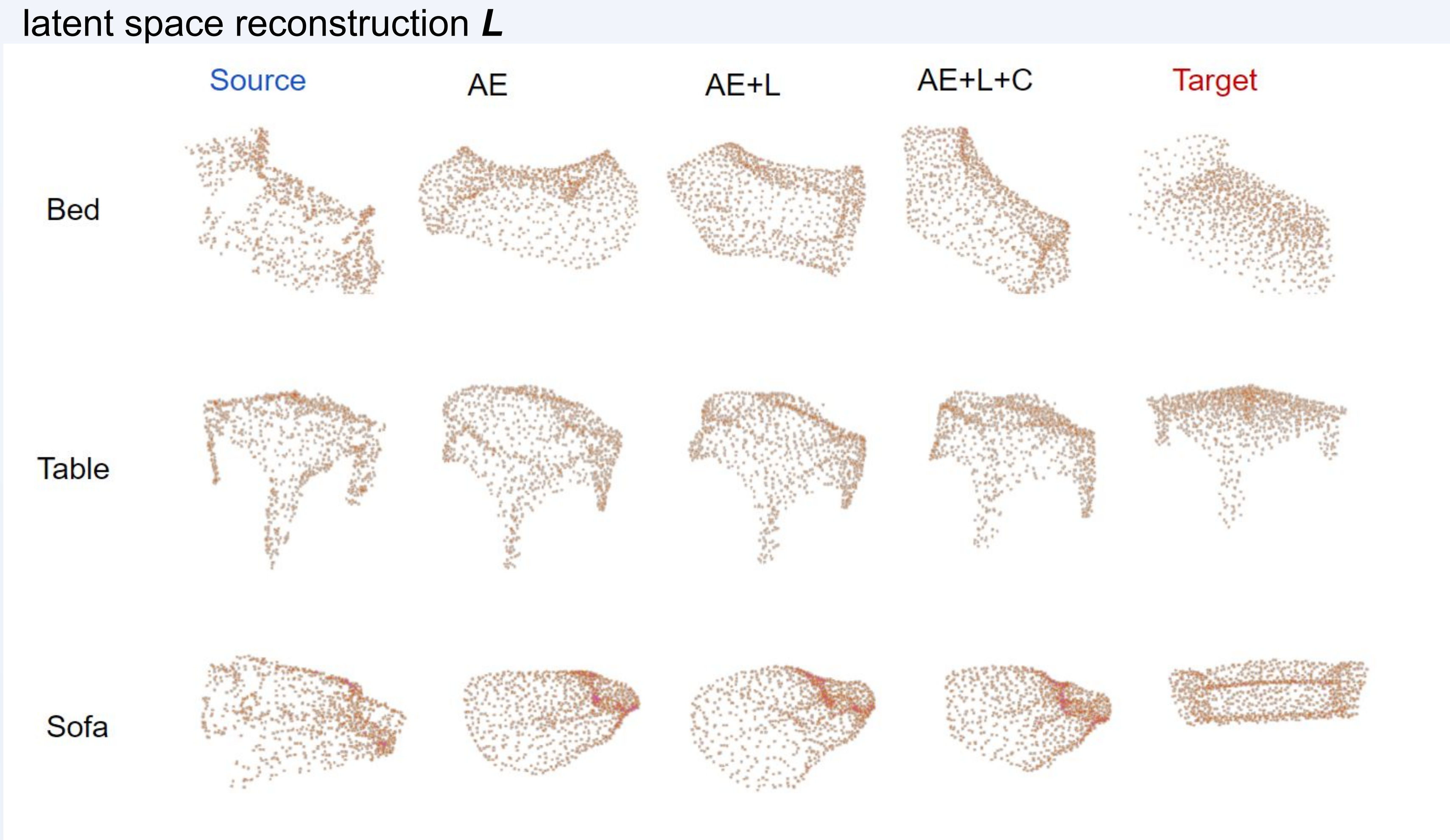
Ablation Study

Table 2. Ablation analysis

	AE	L	C	M→S	M→S*	S→M	S→S*	S*→M	S*→S	Avg
w/o Adapt				42.5	22.3	39.9	23.5	34.2	46.9	34.9
only AE	√			59.5	33.5	34.2	16.1	43.3	55.4	40.3
AE+L	√	√		62.6	34.1	40.4	29.1	49.6	64.3	46.7
GFA	√	√	√	**62.8**	**36.5**	**41.9**	**31.6**	**50.4**	**65.7**	**48.1**
Supervised				90.5	53.2	86.2	53.2	86.2	90.5	76.6

AE means use autoencoder with reconstruction loss in model ,**L** denotes latent space reconstruction with VAE , **C** represents the additional discriminator, a classifier.

From Table2 , we could see the latent space reconstruction play a important role, the classifier's performance significantly increases in all six scenarios after adding the latent space reconstruction ***L***



Conclusion

We have proposed a novel generative approach to unsupervised domain adaptation in the 3D classification task. The basic idea is to transfer source training data into the style of target domain rather than selecting domain invariant feature or implementing feature alignment. Furthermore, we implemented latent reconstruction module and an addition discriminator for enhancing the performance