ECCV 2018
European Conference on Computer Vision
8 – 14 September 2018 | Munich, Germany

Sub-GAN: An Unsupervised Generative Model via Subspaces

Jie Liang[1], Jufeng Yang[1], Hsin-Ying Lee[2], Kai Wang[1], Ming-Hsuan Yang[2,3]

[1] Nankai University [2] University of California, Merced [3] Google Cloud AI

Homepage: http://liangjie.xyz; http://cv.nankai.edu.cn/ Email: liang27jie@gmail.com

Problems

The ambient space

Generator

Need large-scale training dataset & deep architecture to model the ambient space

Generated (desired)

High-dimensional!

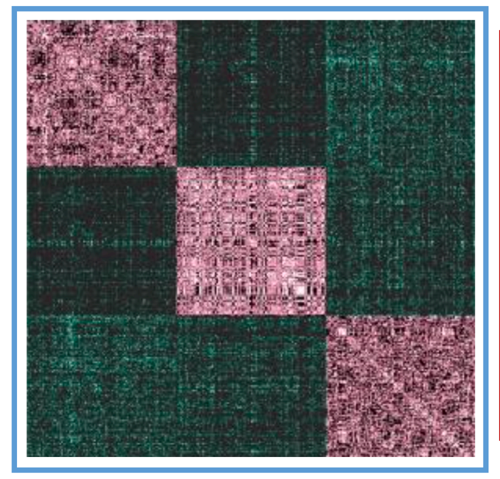
Mode Collapse!

Problems to be solved in this paper:

1. Disentangling the low-dimensional subspaces of the ambient space
2. Generating diverse samples without any supervision

Main Idea

1. To disentangle the low-dimensional subspaces



The Coefficient Matrix in Subspace Clustering is block-diagonal. Each entry reflects the similarity between two samples.

generating ↓↑ clustering

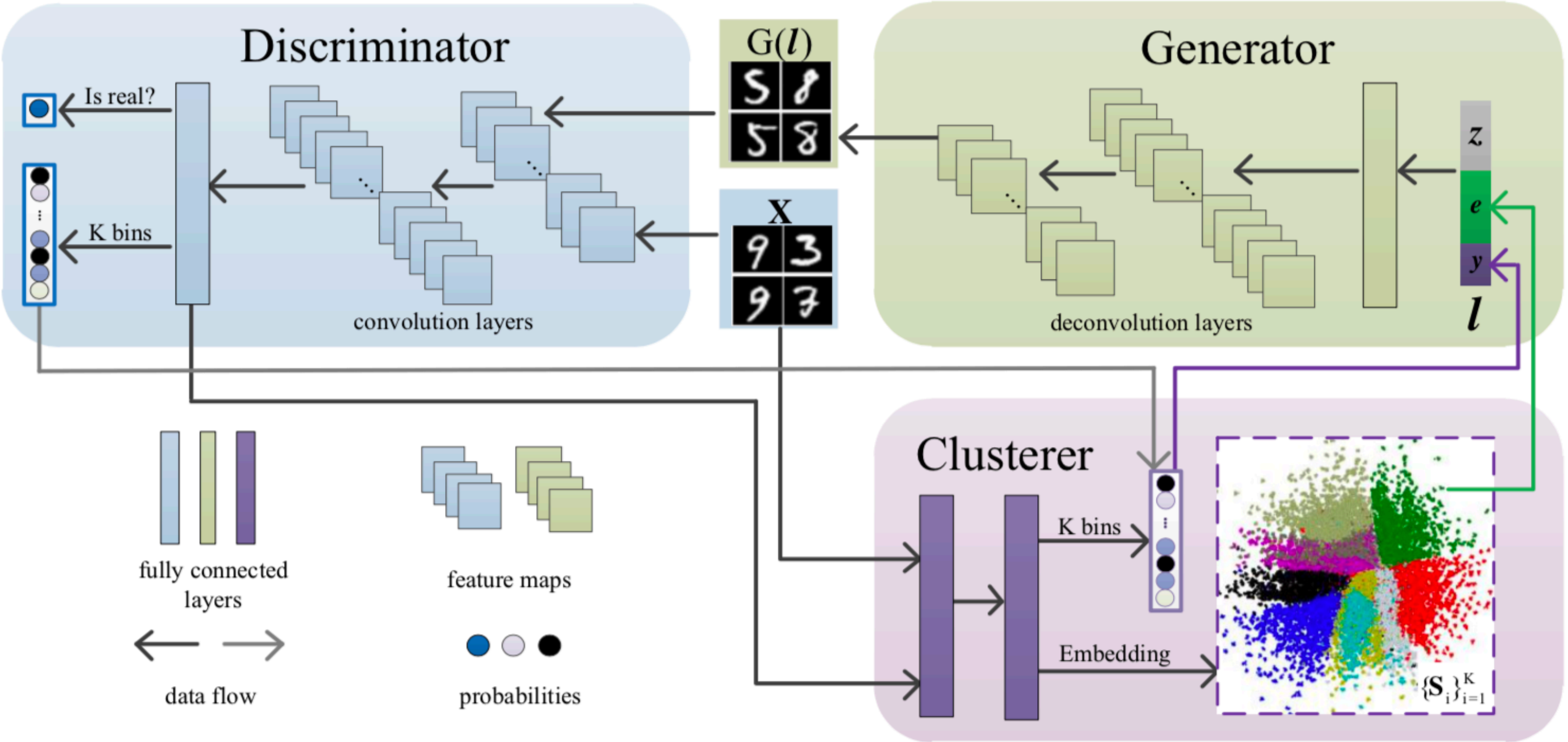
Interaction between G and C. Both modules are improved.

2. To control the diversity of generated samples without supervision

Latent code of the generator: *Random vector* | *Subspace embedding* | *Predicted label* → Generator

Eigenvectors of the Laplacian Matrix

Algorithm



Clusterer: 1. Self-Representation Optimization:

$$\min_{C} \|\boxed{X} - XC\|_2^2 + \lambda \|\boxed{C}\|_1, \quad s.t.\ \mathrm{diag}(C) = \mathbf{0}$$

Data samples — Coefficient Matrix

2. Eigenvectors of the Laplacian matrix:

$$[e_1, e_2, \cdots, e_K] = \mathrm{eig}(M)$$

Each of the K eigenvectors reflects the embedding of a subspace.

Generator: The latent code is composed of three vectors:

$$l = z \oplus e \oplus \hat{y}.$$

Discriminator: 1) Distinguish between real and fake samples; 2) Classify the inputs into subspaces.

Minimax objective:

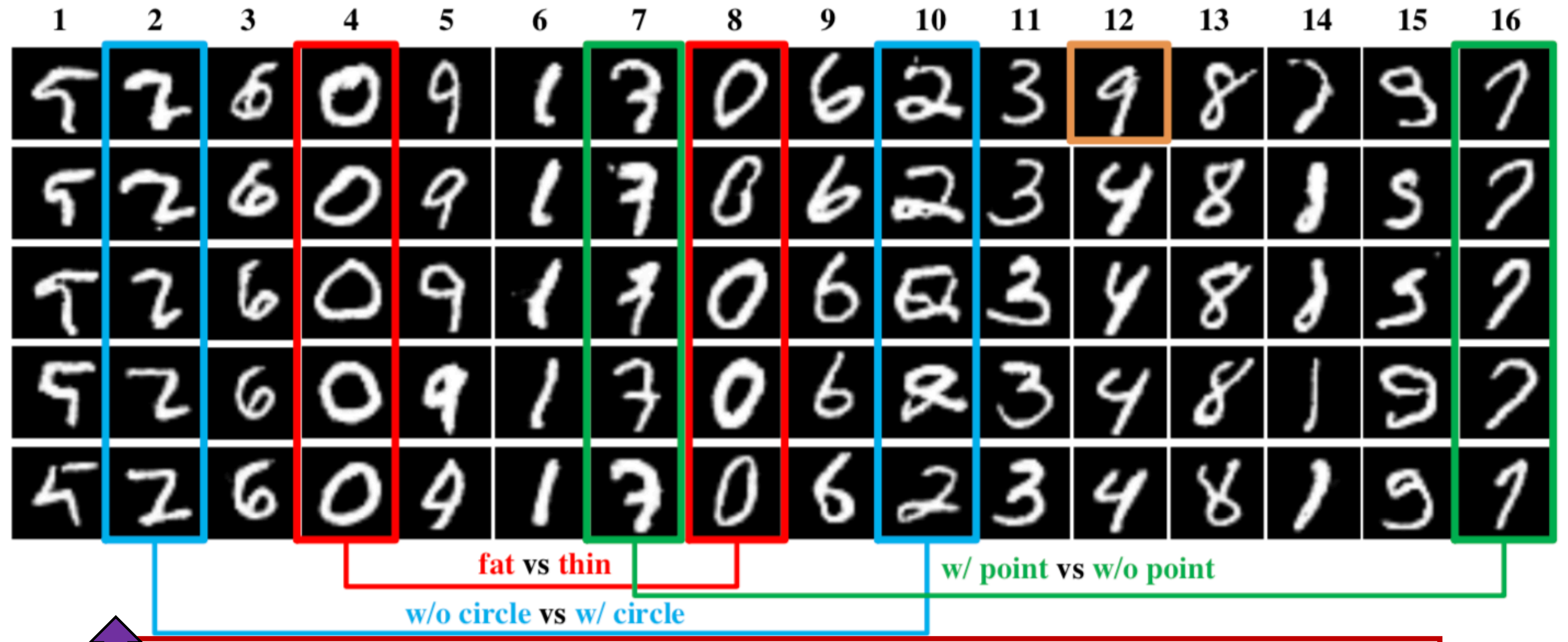
$$\min_{G,C} \max_{D} \mathcal{L}(D,G,C),$$

$$\mathcal{L}(D,G,C) = \mathbb{E}_{x \sim p_S(x)}[\log D(x)] + \mathbb{E}_{l \sim p_L(l)}[\log(1 - D(G(l)))] + \mathrm{KL}(Q_S \| P_S).$$

Experiments

Methods	MNIST			CIFAR		
	$K=10$	$K=16$	$K=20$	$K=10$	$K=16$	$K=20$
K-means	53.49	60.36	62.55	42.62	46.81	51.02
SSC [4]	62.71	66.82	70.19	50.31	52.77	53.98
LSR [46]	66.85	70.21	73.83	53.97	55.80	59.24
SMR [47]	73.39	81.27	83.63	56.24	59.02	62.73
NSN [48]	68.75	71.04	73.67	52.29	56.55	59.03
SSC-OMP [35]	76.33	79.25	82.52	51.21	53.02	57.84
ORGEN [31]	71.04	74.07	78.65	52.29	55.61	58.08
iPursuit [49]	61.35	64.28	68.84	59.21	62.53	65.66
DEC [39]	84.30	83.28	83.02	61.03	65.29	67.31
CatGAN [50]	80.21	84.92	90.30	67.42	67.85	68.76
InfoGAN [29]	70.63	73.77	78.69	71.02	73.64	74.07
Sub-GAN	**85.32**	**90.36**	**90.81**	**78.95**	**81.35**	**82.44**

Unsupervised clustering performance (adjusted clustering accuracy) with different K's.



The Sub-GAN can discover multiple hidden attributes of the data.
The diversity of generation can be controlled by the given number of clusters.

Datasets	Real	CGAN [2]	IGAN [43]	IWGAN [44]	DCGAN [45]	InfoGAN [29]	Sub-GAN
MNIST	2.96	0.92	1.81	1.78	1.63	2.11	**2.36**
CIFAR	3.21	1.02	2.20	2.03	1.95	2.48	**2.72**

Diversity scores with K = 10.

Refinement in D	1^{st} Epoch	20^{th} Epoch	40^{th} Epoch	Last Epoch
W/o	75.23	82.96	83.11	83.87
W/	**77.12**	**83.45**	**84.24**	**85.32**

Clustering accuracy.

Some wrongly grouped using global similarity. ---> Refine the assignment in D based on the similarity of samples in local batches.

Conclusion

1. We present a joint unsupervised framework to simultaneously learn the subspaces of the ambient space and generating instances accordingly, where both tasks are mutually optimized;
2. We address the mode collapse problem by specifying the number of distinct subspaces, from which we generate meaningful and diverse images with informative visual attributes.