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] College of Computer and Control Engineering, Nankai University
[2] School of Engineering, University of California, Merced, USA
[3] Google Cloud

ECCV 2018, Munich

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

NANKAI UNIVERSITY 1919
南開大學
Nankai University

Computer Vision Lab,
College of Computer Science, Nankai University
http://cv.nankai.edu.cn

Outline

- Problem and Related Work
- Motivation
- Main Idea
- The Proposed Sub-GAN method
- Experiments

Problems

In an ideal generative model:

The ambient space

High-dimensional!

Generator

Need large-scale training data and deep architecture to model the ambient space

Generated

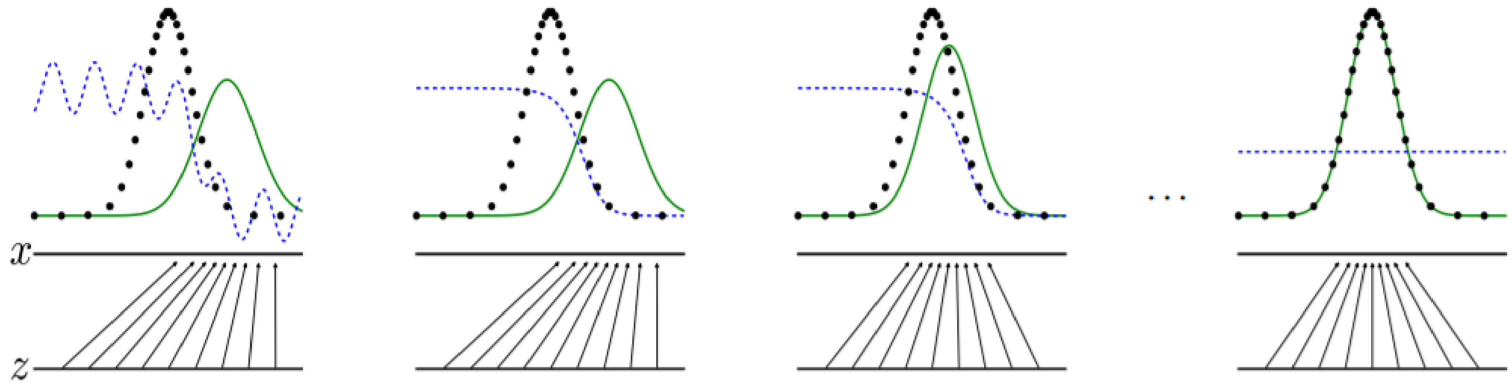
Mode Collapse!

Problems to be solved in this paper:

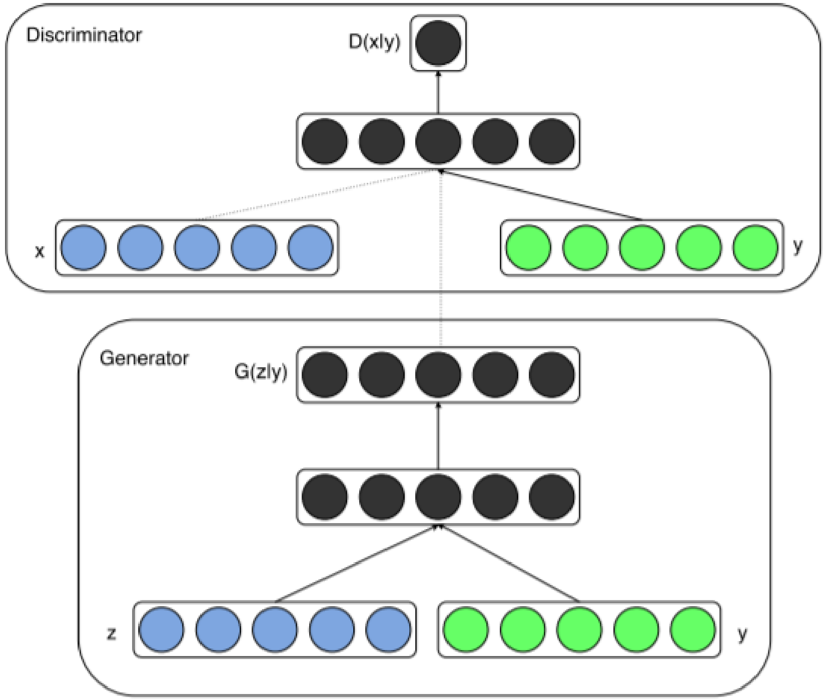
- Disentangling the low-dimensional subspaces of the ambient space
- Generating diverse samples without any supervision

Related Work

Generative Adversarial Networks (GAN):



Training process of the GAN.

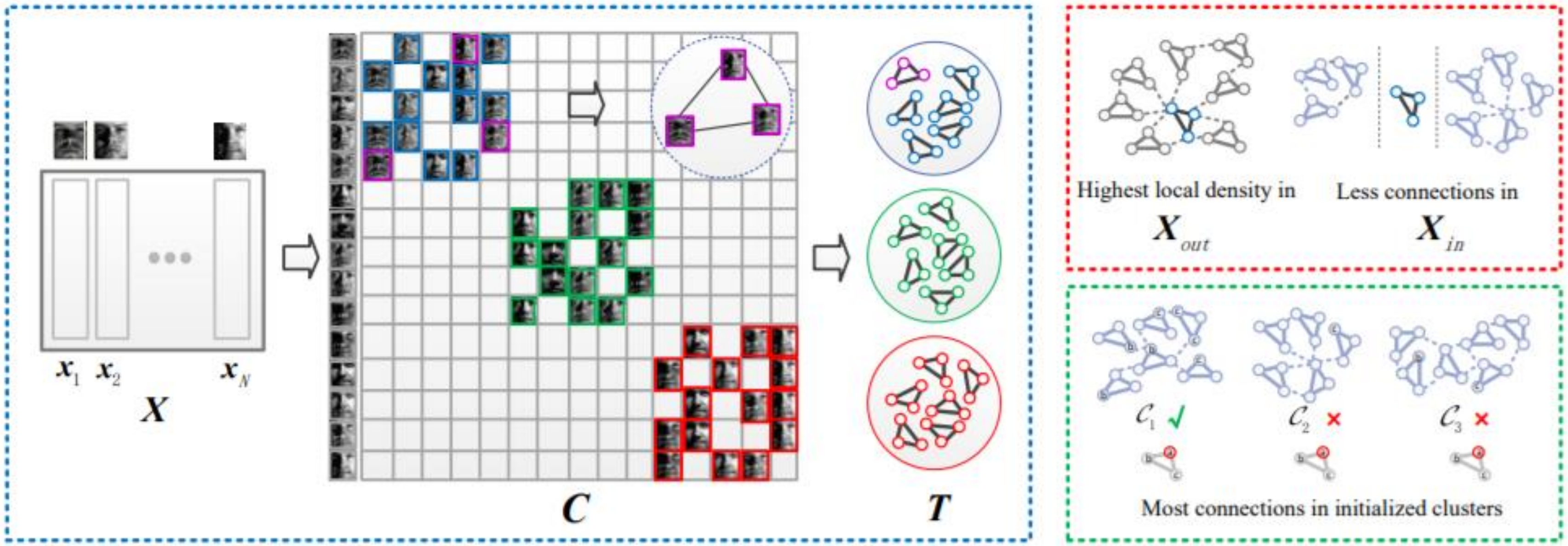


Framework of the conditional GAN.

Goodfellow et. al, Generative Adversarial Nets, NIPS 2014.
Mirza et. al, Conditional Generative Adversarial Nets.

Related Work

Subspace Clustering:



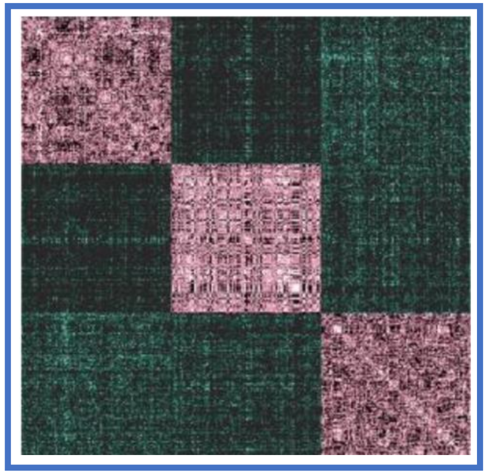
$$\min_{C} \|\boxed{X} - XC\|_2^2 + \lambda \|\boxed{C}\|_1, \quad s.t.\ \mathrm{diag}(C) = \mathbf{0}$$

Data samples Coefficient Matrix

Elhamifar et. al, Sparse Subspace Clustering, CVPR 2009
Yang et. al, Automatic Model Selection in Subspace Clustering via Triplet Relationships, AAAI 2018

Motivation

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 generating module and clustering module. Finally, Both modules are improved.

2. To control the diversity of the generated samples without supervision

Latent code of the generator

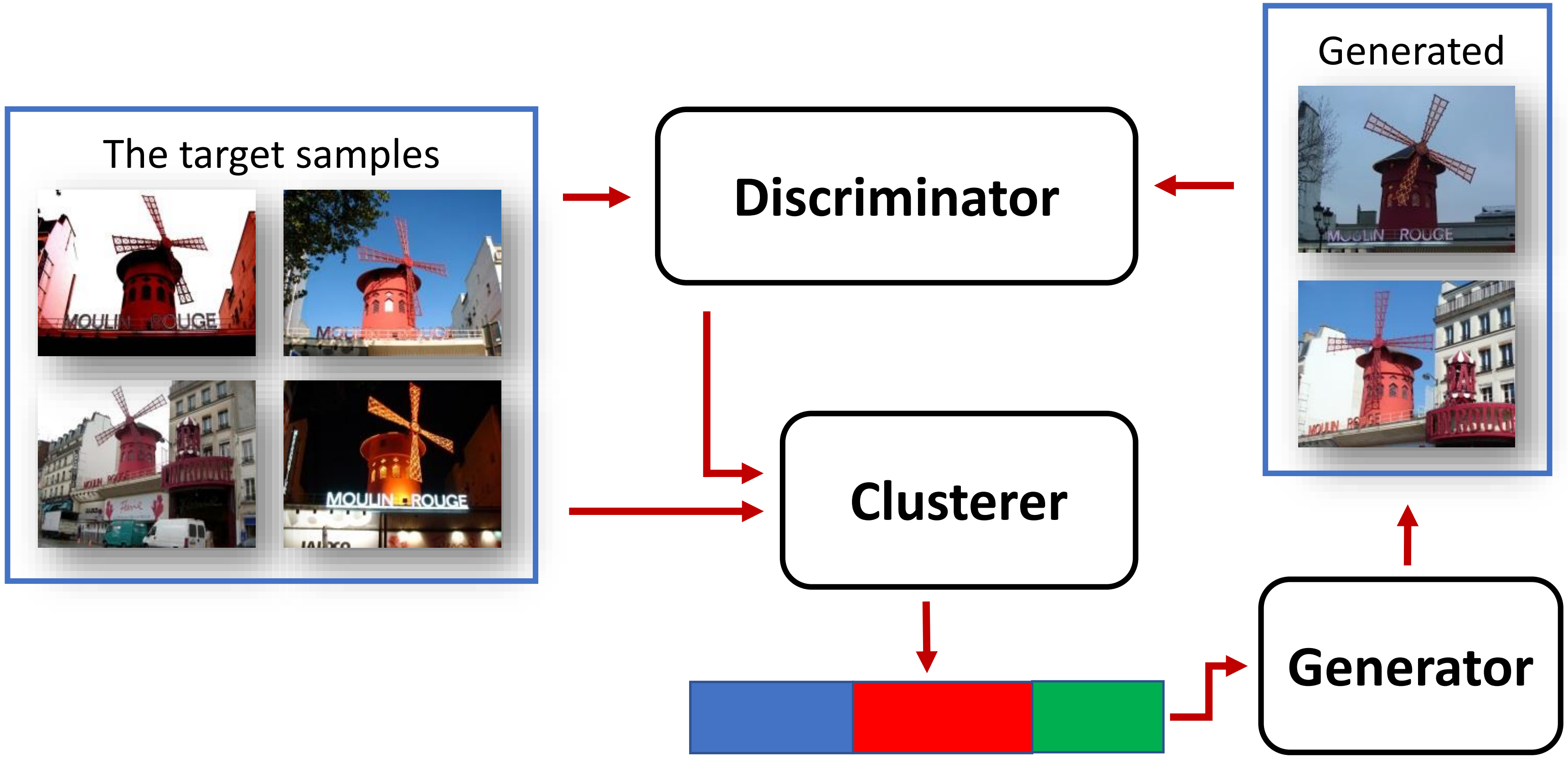
Random vector	*Subspace embedding*	*Predicted label*

Eigenvectors of the Laplacian Matrix

Generator

The subspace embedding and the predicted label are iteratively updated.

Main Idea



The three modules are mutually optimized during training.

Algorithm

Clusterer:

Optimizing a Self-Representation problem:

$$\min_{\boldsymbol{C}} \|\boxed{\boldsymbol{X}} - \boldsymbol{X}\boldsymbol{C}\|_2^2 + \lambda \|\boxed{\boldsymbol{C}}\|_1, \quad s.t. \ \mathrm{diag}(\boldsymbol{C}) = \boldsymbol{0}, \qquad (1)$$

Data samples Coefficient Matrix

Calculating the eigenvectors of the Laplacian matrix:

$$[\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_K] = \mathrm{eig}(\boxed{\boldsymbol{M}}), \qquad (3)$$

Laplacian Matrix

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

Algorithm

Generator:

The latent code of G is composed of three vectors:

Random

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

Predicted label

Subspace embedding

Discriminator:

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

Minimax objective:

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

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

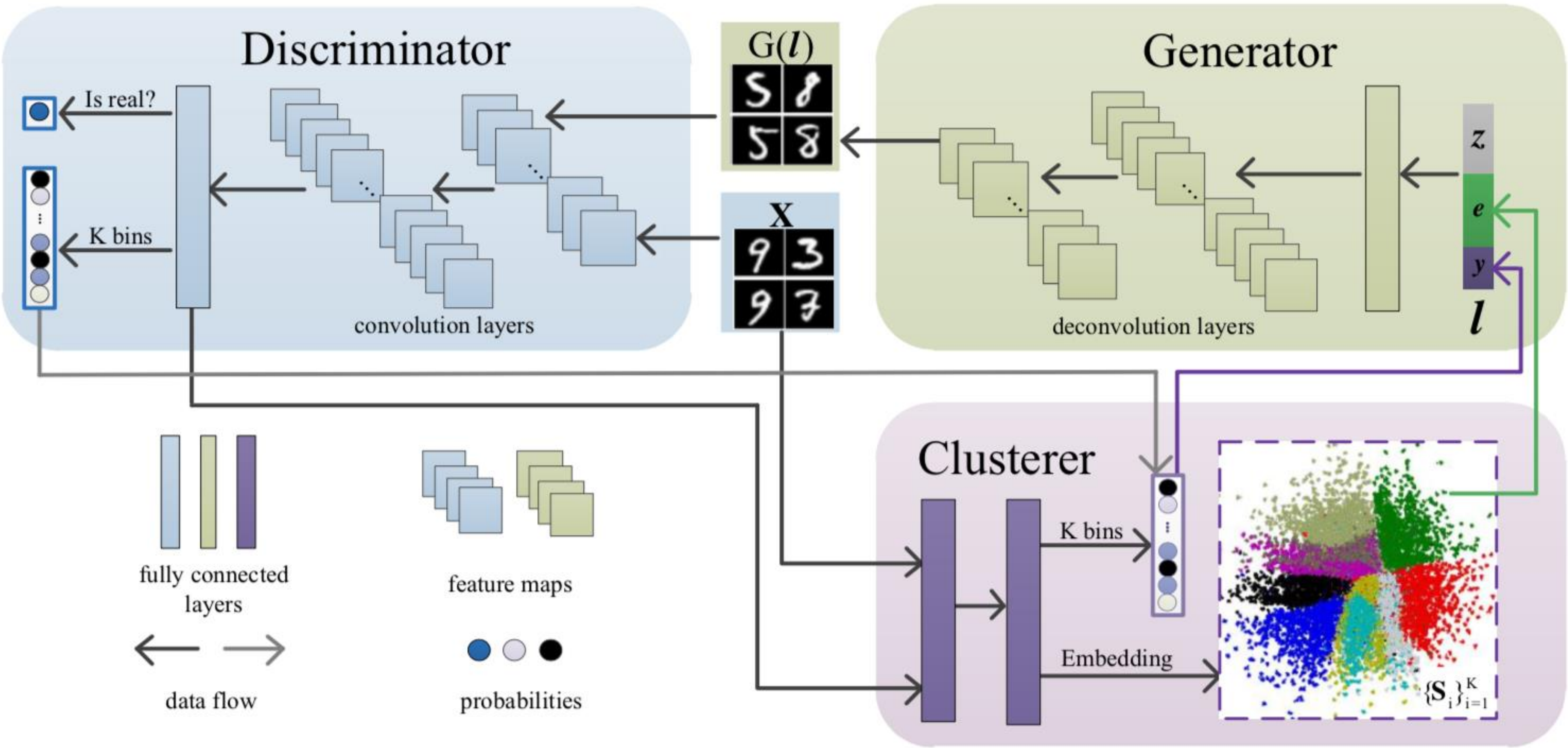
Loss of D

Loss of G

Loss of C

Algorithm

Pipeline of the proposed Sub-GAN:



Experiment

Setup

Datasets: MNIST and CIFAR-10

Contrastive Methods:

Generation: CGAN, Improved GAN, Improved WGAN, DCGAN and InfoGAN.
Clustering: K-means, SSC, LSR, SMR, NSN, SSC-OMP, ORGEN, iPursuit, DEC, CatGAN and InfoGAN.

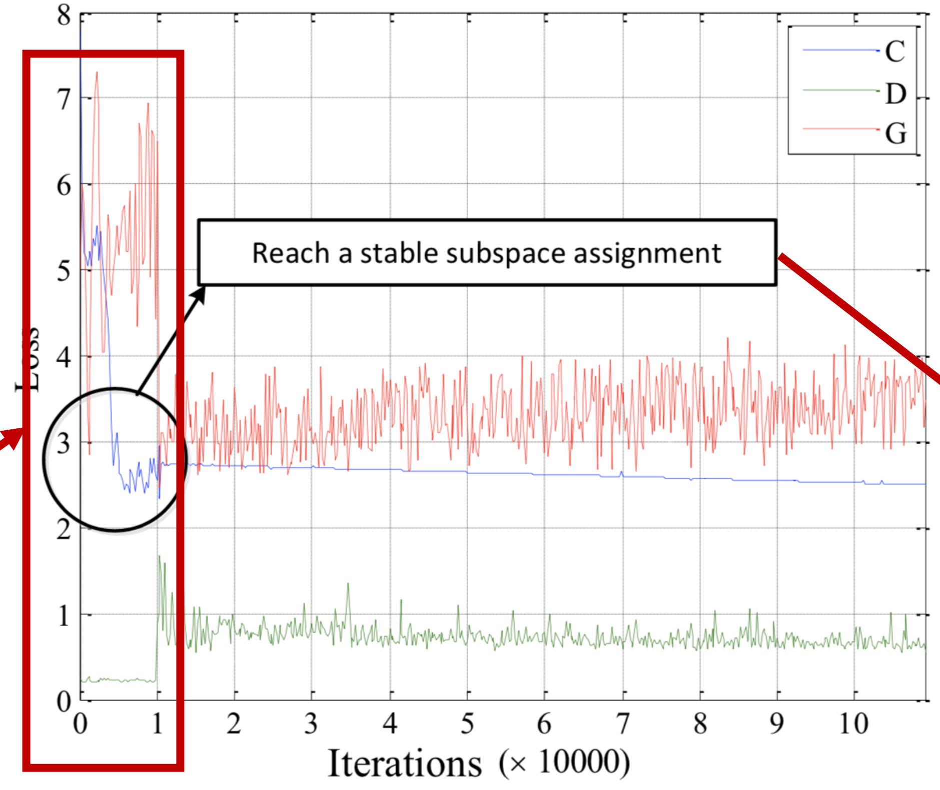
Evaluation Metrics:

Inception Score
Diversity Score
Adjusted Accuracy for Clustering

Experiment

Training Process:

The loss of the Clusterer demonstrates a downward trend before around 10, 000-th iteration. Sequentially, the training of the Generator and Discriminator is unstable, e.g., D can easily discriminate the fake images from the real one so that the loss of discriminator is low.



Optimization losses of three modules.

After C reaches a stable subspace assignment, the framework begins a normal adversarial training of the three modules G, D and C.

Experiment

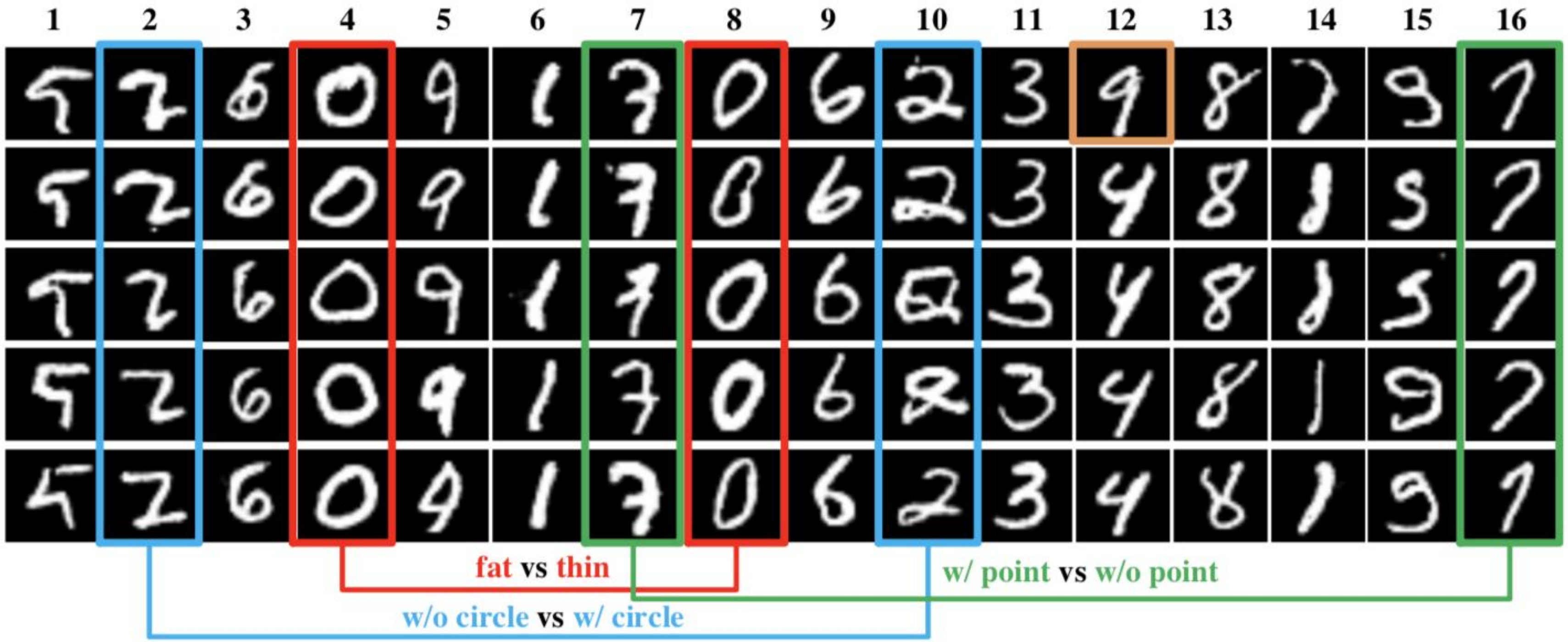
Generation:

Image Quality	Modeling Subspaces?
Not Satisfied	Yes
Satisfied	No
favorable	Yes

Generation Performance of the contrastive methods.

Experiment

Generation:



Samples generated from joint unsupervised training on the MNIST dataset using the proposed Sub-GAN by setting K=16. The Sub-GAN can discover multiple hidden attributes of the data. The diversity of generation can be controlled by the given number of clusters.

Experiment

Generation:

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**

Comparison of the diversity scores on both MNIST and CIFAR datasets with K = 10. The proposed Sub-GAN achieves best performance against contrastive methods, which alleviates the mode collapse problem in training GANs.

Experiment

Generation:

Samples

Models	CGAN [2]	IGAN [43]	IWGAN [44]	DCGAN [45]	InfoGAN [29]	Sub-GAN
Inception Score	4.28 ±0.08	8.09±0.07	7.86±0.07	6.16 ±0.07	7.26±0.05	7.95 ±0.04

Inception scores for samples derived from various generative models on the CIFAR-10 dataset.

Experiment

Clustering:

Some samples might be wrongly grouped based on the global similarity to all others.

➡

Refine the assignment in D based on the similarity of samples in local batches.

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 (%) on the MNIST dataset under K = 10 with/without the refinement operation in the discriminator.

Experiment

Clustering Performance:

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) of the contrasted methods on the MNIST and CIFAR datasets with different K's.

Conclusion

The proposed unsupervised Sub-GAN model:

- Jointly learning the latent subspaces of the ambient space and generating instances correspondingly
- **Clusterer**: aims to discover distinctive subspaces of high-dimensional data in an unsupervised fashion, which is updated on each epoch based on the feedback from the discriminator
- **Generator**: produces samples conditioned on a one-hot vector indicating the belonged cluster and a base vector of subspace derived from the clusterer
- **Discriminator**: not only needs to distinguish between real and fake samples, but also requires to classify them to belonged subspaces. It also provides distinctive representations of data samples for updating the clusterer

Thank you