Study and Analysis of GAN and VAE in continual learning.

Ayush Pande
Roll No.: 20111404
Email id: ayushp20@iitk.ac.in

Narein Rao
Roll No.: 20111414
Email id: nrao20@iitk.ac.in

Pragya Agrawal
Roll No.: 20211402
Email id: pragyaa20@iitk.ac.in

Sri Madhan M
Roll No.: 20111420
Email id: madhan20@iitk.ac.in

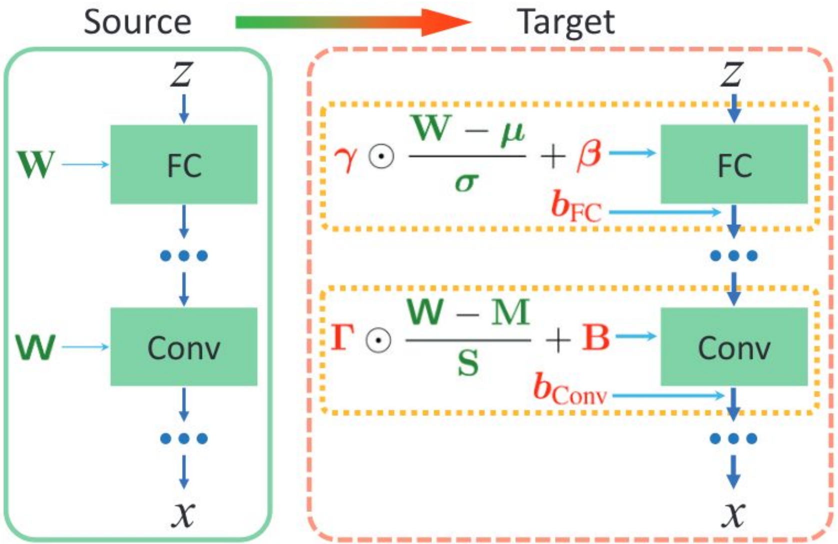
Project synopsis

- To reproduce experimental results and gain a measure of how realistic the generated images are across the two approaches.
- To measure the forgetfulness of the two approaches, find the relationship between number of tasks and forgetfulness and compare the performance of the two approaches in continual learning.
- To generate images over a domain perceptually-distant from the domain the models have been trained over (few-shot learning)

Base Literature

Cong, Yulai et al. "GAN Memory with No Forgetting."

A realistic generative replay framework to alleviate catastrophic forgetting



Src: Cong, Yulai et al. "GAN Memory with No Forgetting."

- Each group of style parameters modulates a different generation perspective

  scales $\{\gamma, \Gamma\}$ = textural/structural information

  shifts $\{\beta, B\}$ = control color information

  biases $\{b_{FC}, b_{Conv}\}$ = control the illumination and localized objects

- Style parameters show different strength/focus over the generation

Changes in the overall contrast and illumination of the generation

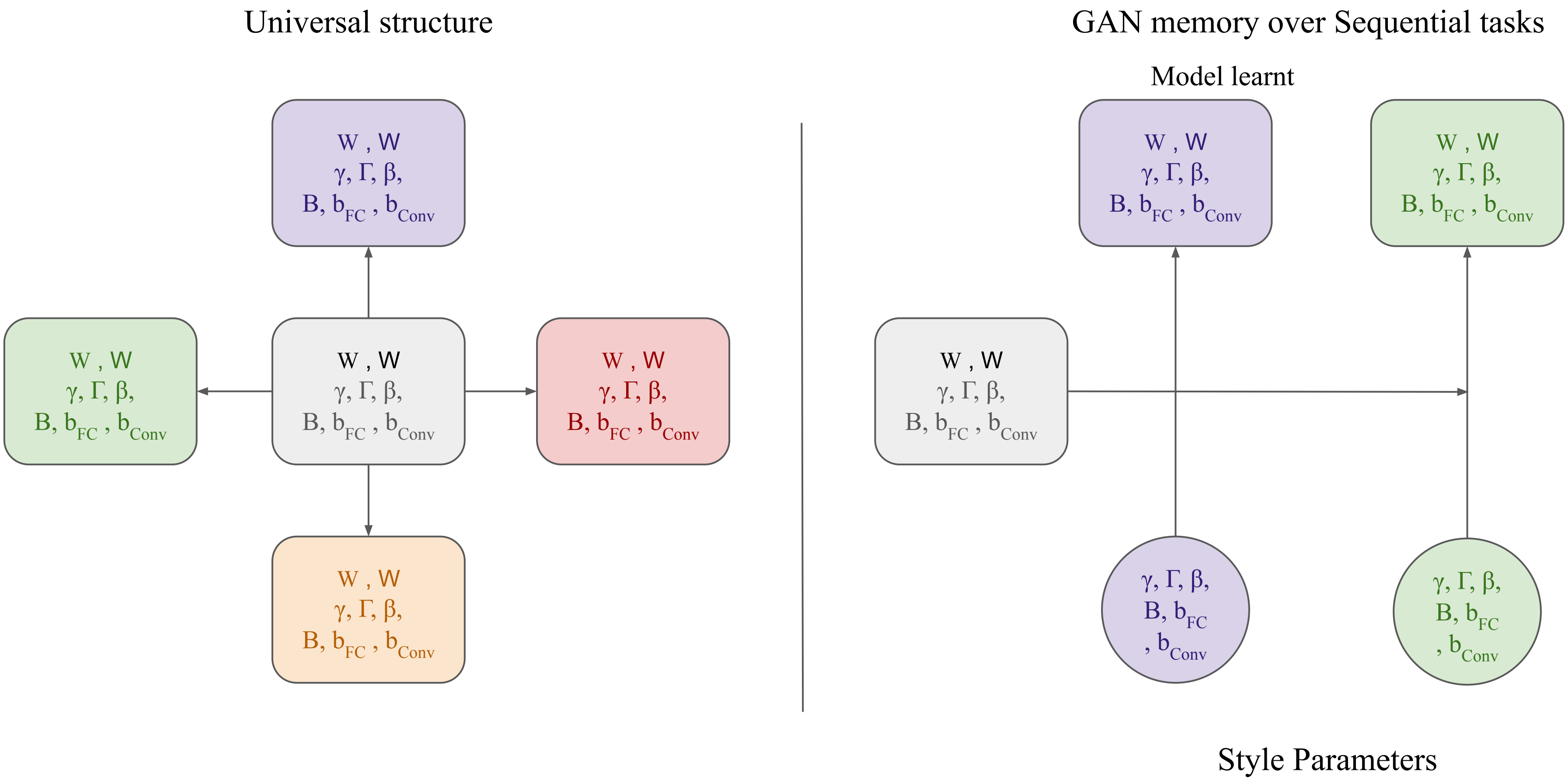
↓

Style parameters effort to modulate

↓

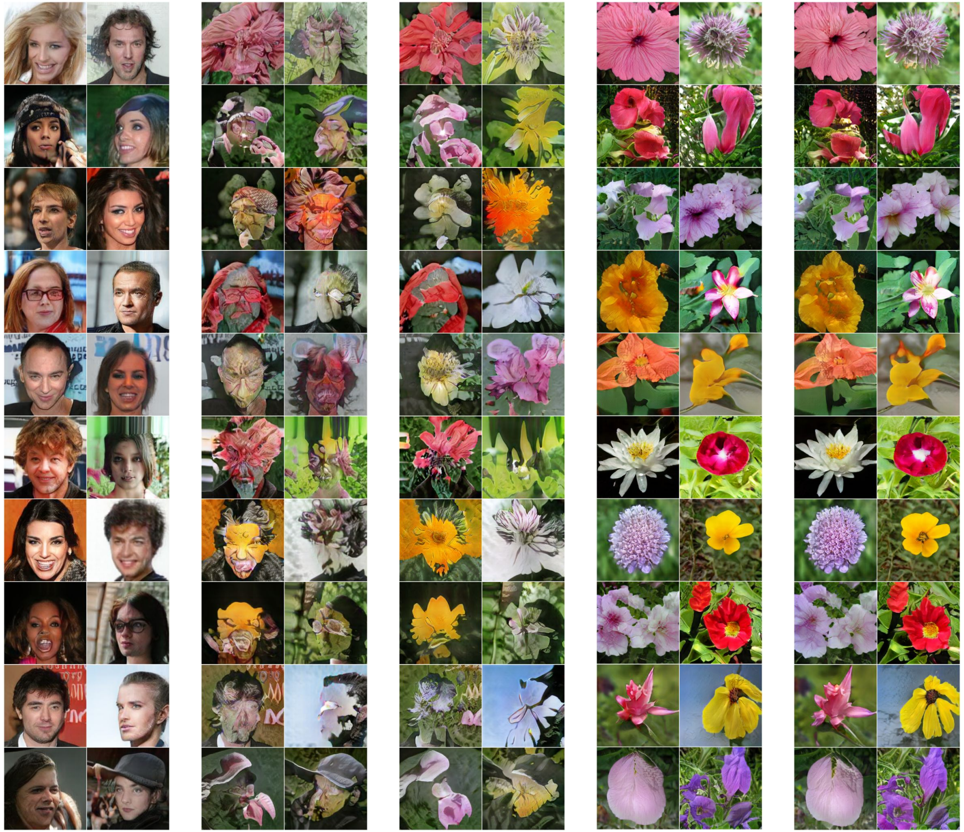
Refining the generation details.

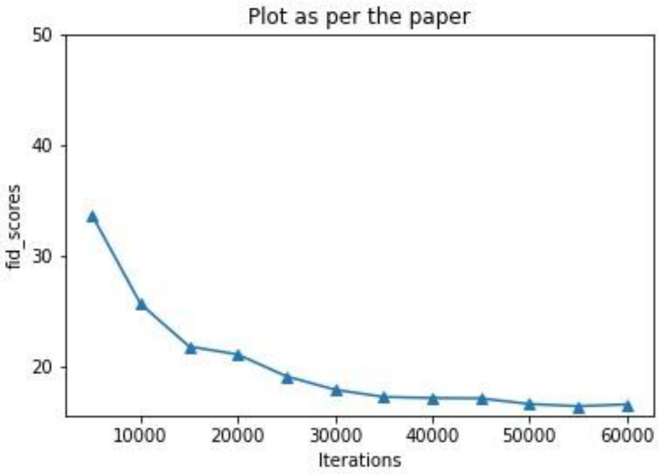
- Normalization contributes significantly to better training efficiency and performance
- GAN memory enables smooth interpolations among generative processes

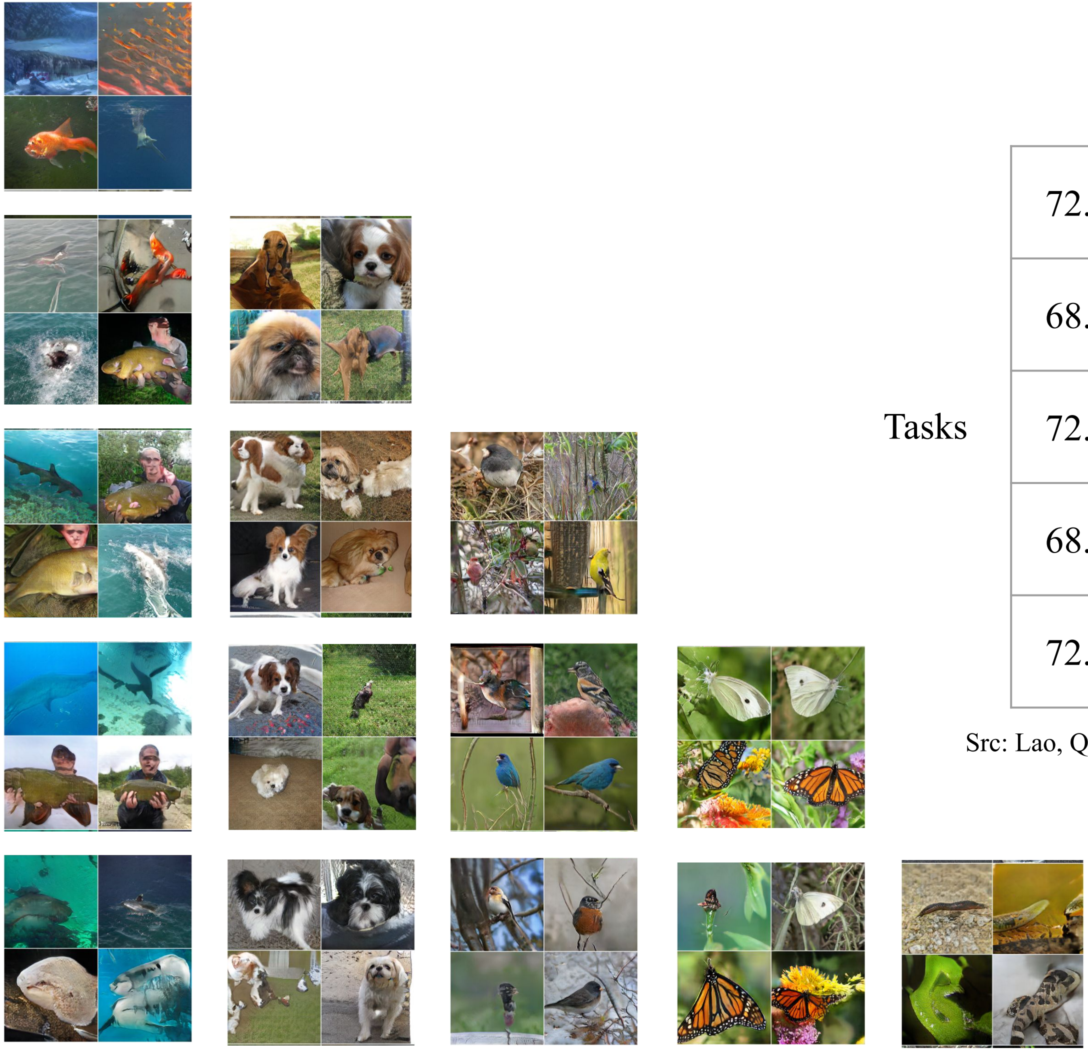

Universal structure
GAN memory over Sequential tasks
Model learnt
W , W
γ, Γ, β,
B, b_{FC} , b_{Conv}
W , W
γ, Γ, β,
B, b_{FC} , b_{Conv}
W , W
γ, Γ, β,
B, b_{FC} , b_{Conv}
W , W
γ, Γ, β,
B, b_{FC} , b_{Conv}
W , W
γ, Γ, β,
B, b_{FC} , b_{Conv}
W , W
γ, Γ, β,
B, b_{FC} , b_{Conv}
W , W
γ, Γ, β,
B, b_{FC} , b_{Conv}
W , W
γ, Γ, β,
B, b_{FC} , b_{Conv}
γ, Γ, β,
B, b_{FC}
, b_{Conv}
γ, Γ, β,
B, b_{FC}
, b_{Conv}
Style Parameters

Experimental results

GAN







FID Scores

Tasks

72.34				
68.27	64.15			
72.72	63.93	47.85		
68.74	68.13	48.59	46.95	
72.29	63.27	47.65	46.15	83.29

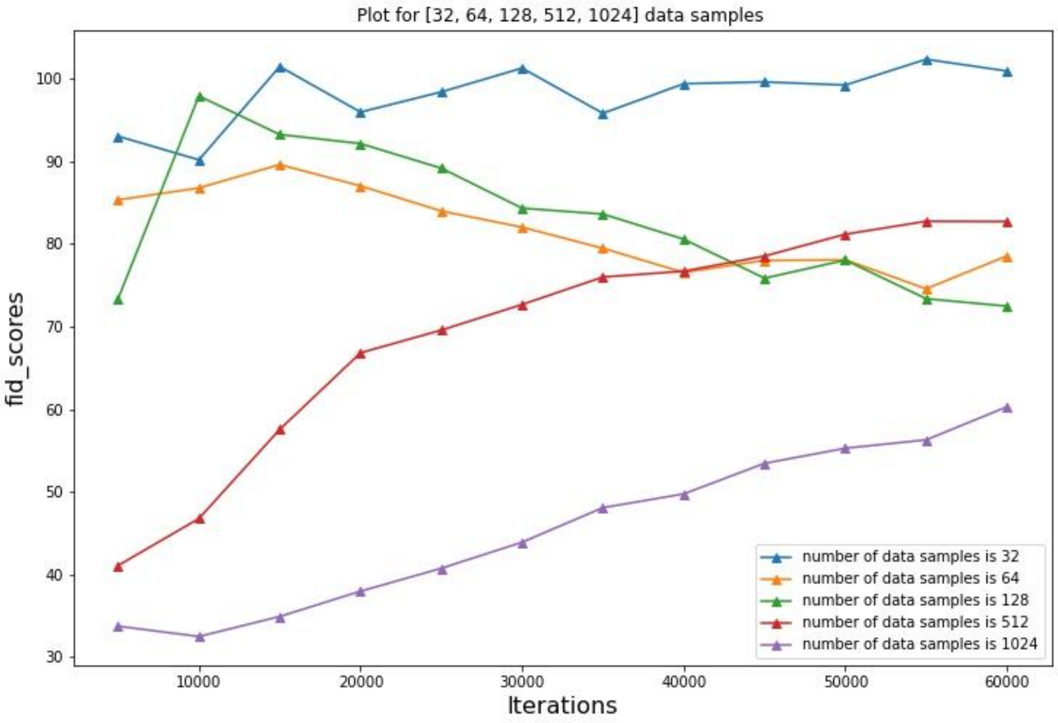
Src: Lao, Qicheng et al. "FoCL: Feature-Oriented Continual Learning for Generative Models."

Forgetfulness score: 0.473

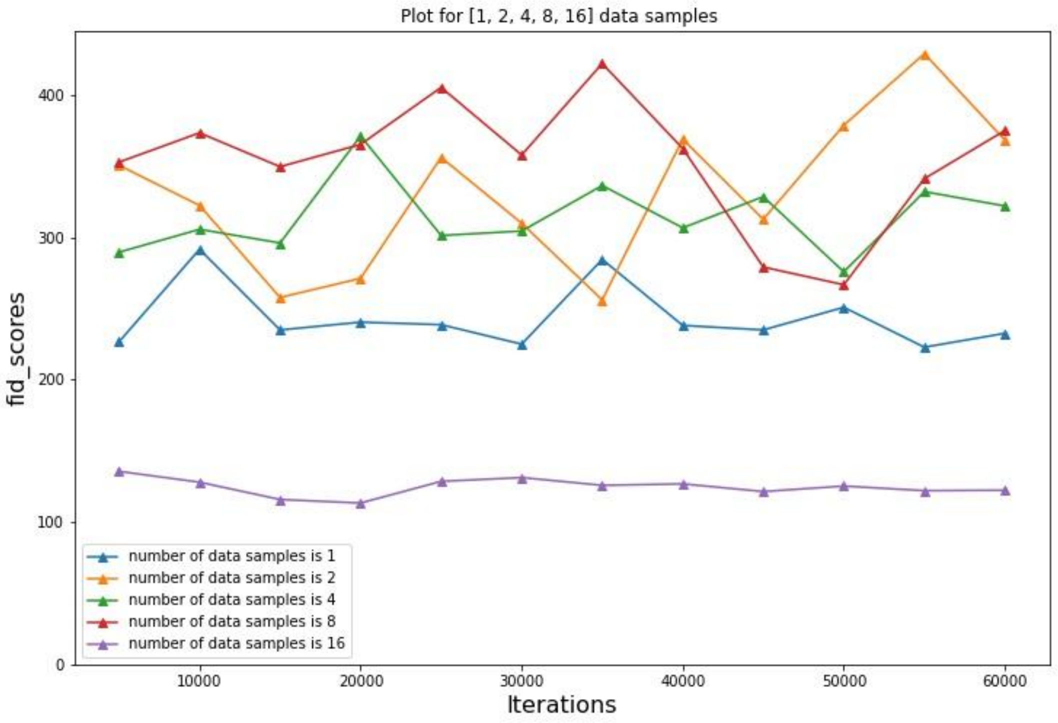
$$FS_t = \frac{1}{t-1}\sum_{i=1}^{t-1}(d_t^{(i)} - d_i^{(i)}).$$

$$FS = \frac{2}{T*(T-1)}\sum_{t=2}^{T}(t-1)FS_t.$$

- Mode collapse seen higher for smaller sample sizes
- Mode collapse and poor generation quality issues



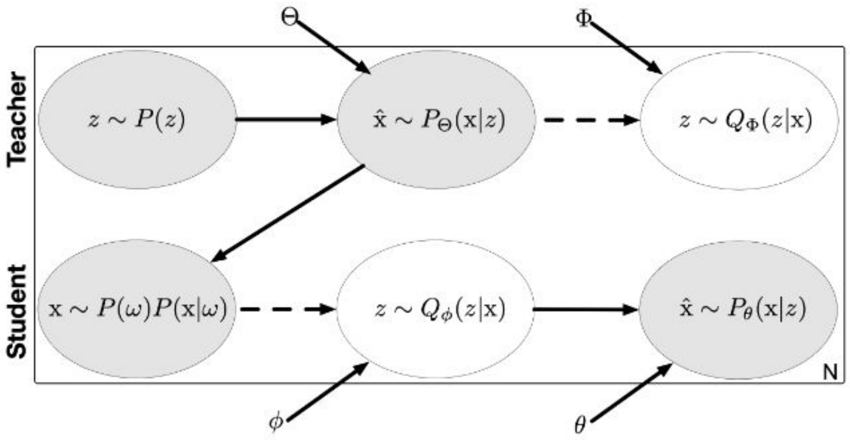
- Fid score trend for varying sample sizes
- Low quality generation despite low FID score



Generation capabilities with varying sample size (1,2,4,8,16,32,128,1024)

Base Literature

Ramapuram, Jason et al. "Lifelong Generative Modeling."



Algorithm 1 Data Flow

Teacher:
Sample Prior: $\boldsymbol{z}_j \sim P(\boldsymbol{z})$
Decode: $\hat{\mathbf{x}}_j \sim P_{\boldsymbol{\Theta}}(\mathbf{x}|\boldsymbol{z})$

Student:
Sample : $\mathbf{x}_j \sim P(\boldsymbol{\omega})P(\mathbf{x}|\boldsymbol{\omega})$
Encode : $\boldsymbol{z}_j \sim Q_{\boldsymbol{\phi}}(\boldsymbol{z}|\mathbf{x})$
Decode: $\hat{\mathbf{x}}_j \sim P_{\boldsymbol{\theta}}(\mathbf{x}|\boldsymbol{z})$

Src: Ramapuram, Jason et al. "Lifelong Generative Modeling."

$$L_{\theta,\phi}(x) = \underbrace{E_{\theta_\phi(z_c,z_d|x)}[\log P_\theta(x \mid z_c, z_d)] - KL[Q_\phi(z_c z_d \mid x) \mid\mid P(z_c, z_d)]}_{(a)} + \underbrace{(1-w)KL[Q_\phi(z_c, z_d \mid x) \mid\mid Q_\Phi(z_c, z_d \mid x)]}_{(b)} - \underbrace{\lambda I(\hat{X}, z_c)}_{(c)}$$

$$\mathbf{x} \sim P(\boldsymbol{\omega})P(\mathbf{x}|\boldsymbol{\omega}),\ \boldsymbol{\omega} \sim \mathrm{Ber}(\pi)$$

$$P(\boldsymbol{\omega})P(\mathbf{x}|\boldsymbol{\omega}) = \begin{cases} P_{\boldsymbol{\Theta}}(\mathbf{x}|\boldsymbol{z}) & \boldsymbol{\omega} = 0 \\ P_i(\mathbf{x}) & \boldsymbol{\omega} = 1 \end{cases}$$

Aim to maximize the above equation

(a) ELBO equation of a standard VAE

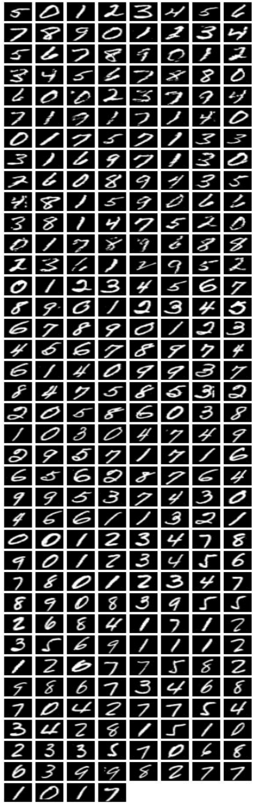
(b) Posterior regularization - can change posterior given new information

(c) Incorporates mutual information - enforces discriminate information about each distribution in model

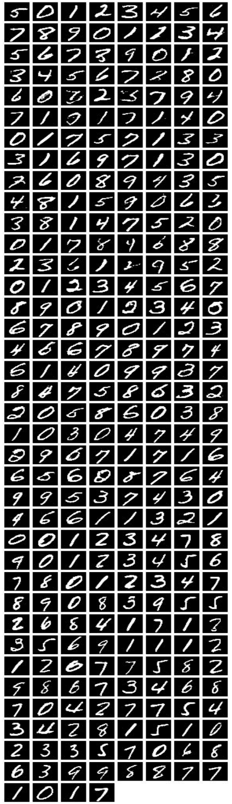
Experimental results

VAE

Original

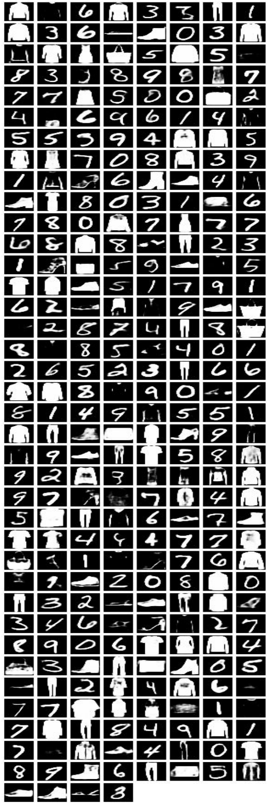


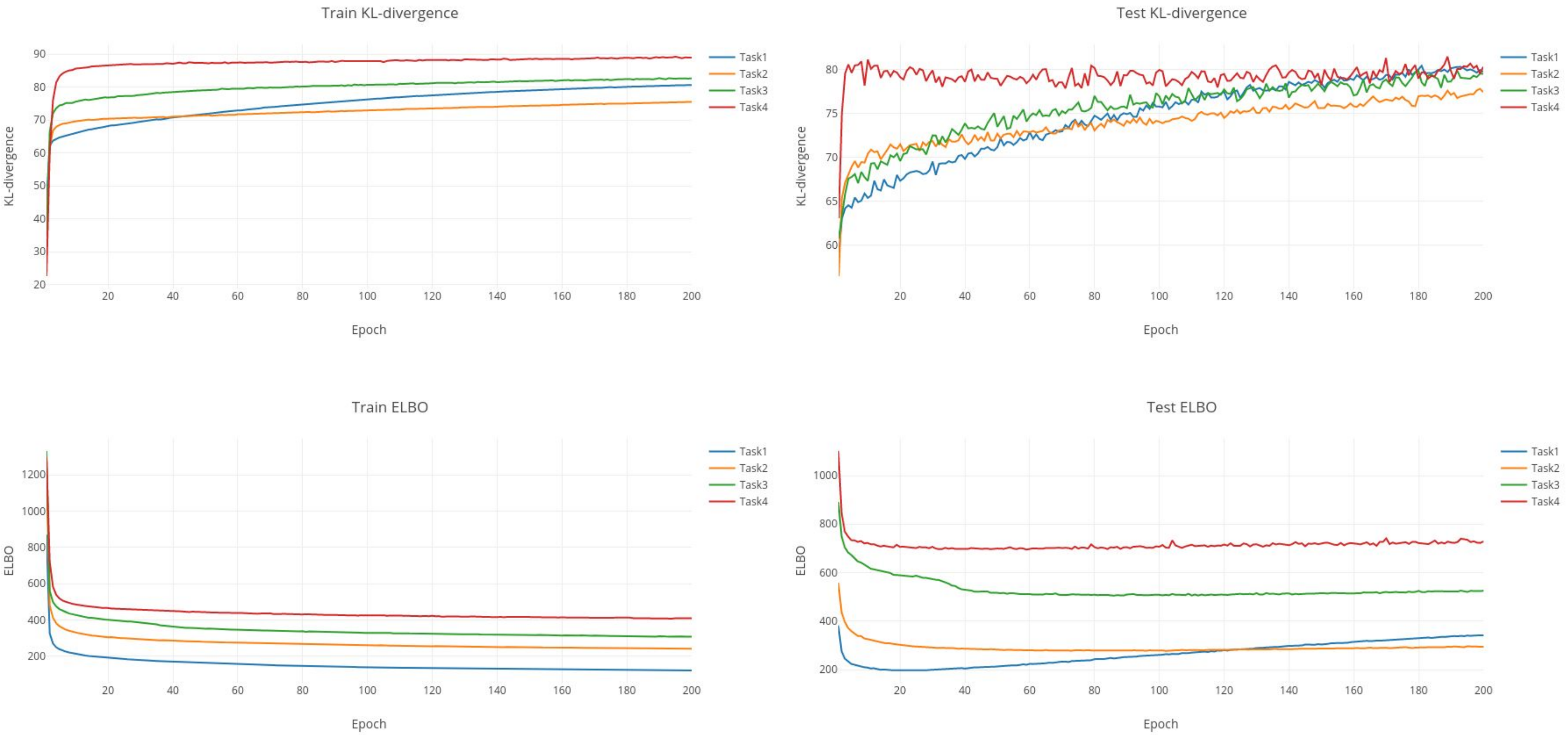
Reconstructed



Original

Reconstructed




Train KL-divergence
KL-divergence
Epoch
Task1
Task2
Task3
Task4
Test KL-divergence
KL-divergence
Epoch
Task1
Task2
Task3
Task4
Train ELBO
ELBO
Epoch
Task1
Task2
Task3
Task4
Test ELBO
ELBO
Epoch
Task1
Task2
Task3
Task4

Task 1 | Task 2 | Task 3 | Task 4

FID Scores

Tasks

172.61			
178.33	149.74		
175.87	172.72	188.46	
185.81	204.12	214.34	220.11

Src: Lao, Qicheng et al. “FoCL: Feature-Oriented Continual Learning for Generative Models.”

$$FS_t = \frac{1}{t-1}\sum_{i=1}^{t-1}(d_t^{(i)} - d_i^{(i)}).$$

$$FS = \frac{2}{T*(T-1)} \sum_{t=2}^{T}(t-1)FS_t.$$

Forgetfulness score: 8.59

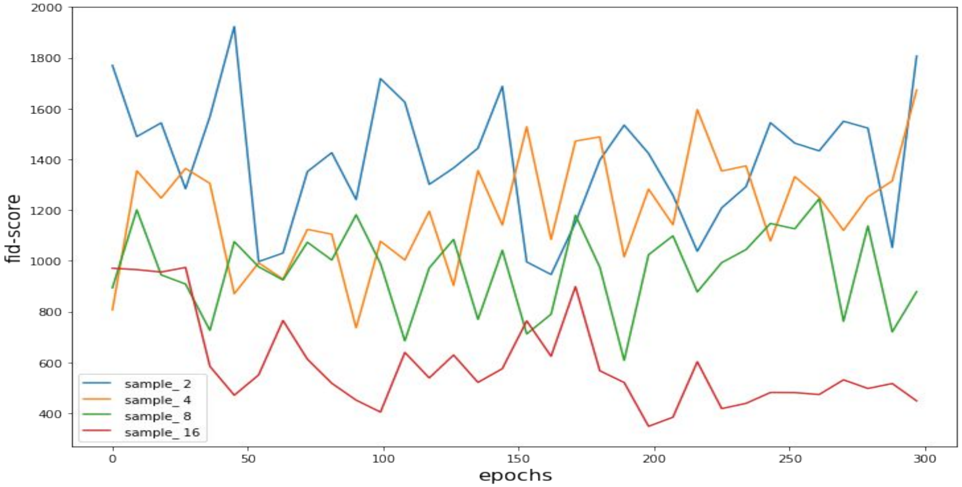
Task1 - MNIST
Task2 - FashionMNIST + regenerated Task1
Task3 - SVHN + regenerated Task2
Task4 - CIFAR10 + regenerated Task3

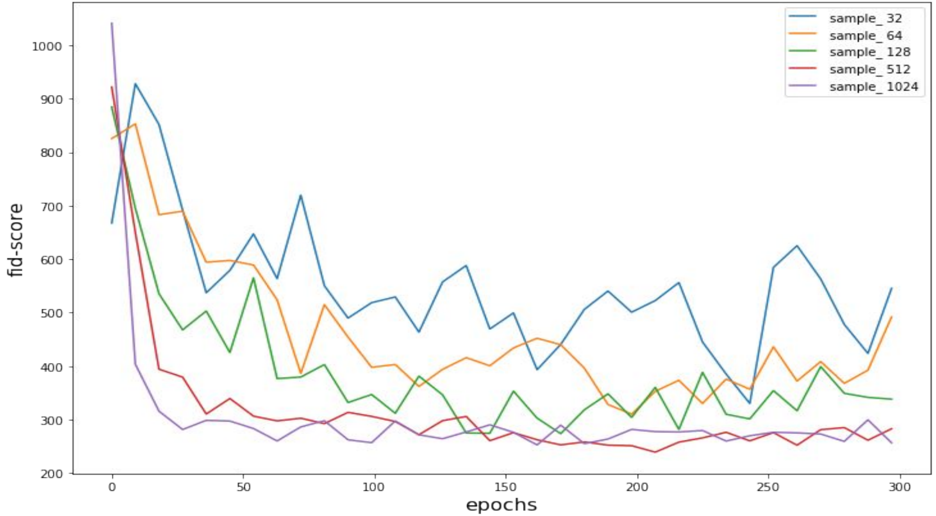
Analysis

VAE

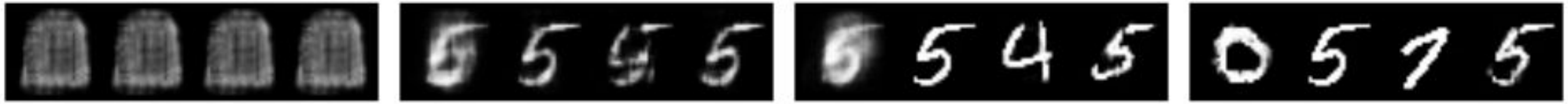
- As new task comes, the generated images of previously trained dataset degrades
- FID scores increases for all the datasets, as new dataset is added
- Poor regeneration of colored images
- Most of the regenerated images are blurry and few images are even blacked-out

Few-shot on VAE: (Sample size: 2, 4, 8, 16, 32, 64, 128, 512, 1024)





Sample size 8



(a) num-epochs:1 (b) num-epochs:100 (c) num-epochs:200 (d) num-epochs:300

Sample size 16

(e) num-epochs:1 (f) num-epochs:100 (g) num-epochs:200 (h) num-epochs:300

Sample size 32

(i) num-epochs:1 (j) num-epochs:100 (k) num-epochs:200 (l) num-epochs:300

Sample size 1024

(u) num-epochs:1 (v) num-epochs:100 (w) num-epochs:200 (x) num-epochs:300

Possible future work

Could perform the following to gain an improvement of performance of VAE

- Save latent vectors (VAE)
- Compression using SVD (reference: Cong, Yulai et al. "GAN Memory with No Forgetting.")
- A metric to quantify the mode collapse problem of GAN.