On the Evaluation of Generative Adversarial Networks By Discriminative Models

Amirsina Torfi, Mohammadreza Beyki, Edward A. Fox

Introduction

- Use Siamese Neural Networks (SNN) to evaluate GANs
- Domain Agnostic
- Comparable to FID Score and Inception Score

Targets

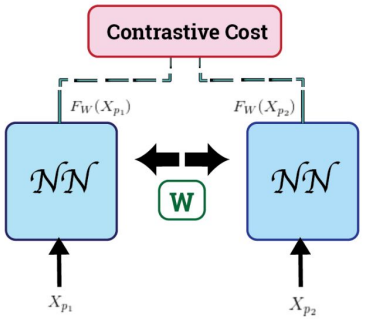
- Visual Quality
- Mode Collapse
- Mode Drop
- Mode Invention

Approach

- Siamese Architecture
- Train on Real Dataset
- Loss should be smaller in Genuine vs. Impostor Pairs

Contrastive Cost

- **Genuine Pairs** are from the same category
- **Impostor Pairs** are from different categories
- Calculate cost for all pairs using the distance function



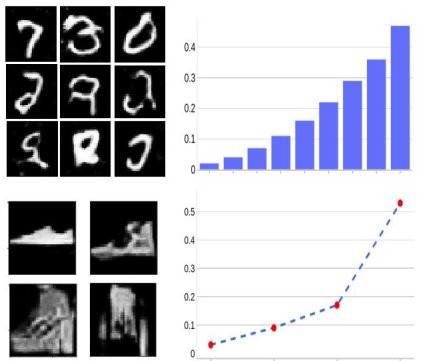
Evaluation

This is SDS on MNIST and FashionMNIST datasets

SDS score increases as the image quality lowers

SDS score is aligned with human evaluation

Graphs represent images from left to right and then from top to bottom



Loss Functions

Distance Function

$$D_W(X_{p_1}, X_{p_1}) = ||F_W(X_{p_1}) - F_W(X_{p_2})||_2.$$

Loss for Genuine and Impostor Pairs

M is the margin obtained through Cross-Validation

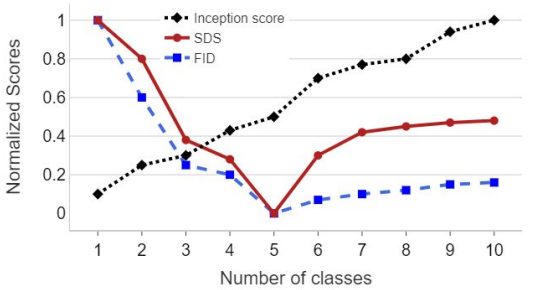
$$\begin{cases} \mathcal{L}_{gen}(D_W) = \frac{1}{2}(D_W)^2 \\ \mathcal{L}_{imp}(D_W) = \frac{1}{2}(max\{0, M - D_W\})^2 \end{cases}$$

Loss

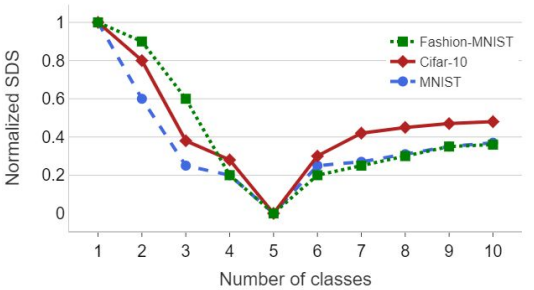
$$\mathcal{L}_W(Y_i, (X_{p_1}, X_{p_2})_i) = Y * \mathcal{L}_{gen}(D_W(X_{p_1}, X_{p_2})_i) + (1 - Y) * \mathcal{L}_{imp}(D_W(X_{p_1}, X_{p_2})_i$$

Procedure

- Feed real and fake samples to the network and calculate their features
- Find all real samples that closely resemble fake samples
- Classify fake samples
- Calculate SDS Score for all samples



SDS Score is sensitive to mode drop and mode invention compared to FID and Inception score
Discriminator is Trained on 5 classes



SDS Score is sensitive to mode drop and mode invention on MNIST, FashionMNIST and Cifar-10
Discriminator is Trained on 5 classes

Conclusion

- Novel approach to evaluate GANs
- **Domain agnostic**
- Can be used on other generative models
- Sensitive to common GAN failures
- Captures visual imperfections