# Comparison and Analysis of GAN Variants on CIFAR-10



| Sr.No | Name | BITS ID |
|---|---|---|
| 1. | Hari Shankar Jaiswal | 2023aa05106 |
| 2. | Warun Kumar | 2023aa05244 |
| 3. | Gandhi Disha | 2023aa05388 |



## 1. Introduction

This report presents a detailed analysis of different Generative Adversarial Networks (GANs) trained on the CIFAR-10 dataset. The following GAN architectures were implemented and evaluated:

- **Task 1:** Conditional Wasserstein GAN (WGAN)
- **Task 2:** Spectrally Normalized GAN (SNGAN)
- **Task 3:** Self-Attention GAN (SAGAN) without spectral normalization and TTUB
- **Task 4:** Complete Self-Attention GAN (SAGAN)

Each model was evaluated based on the **Inception Score (IS)** and **Fréchet Inception Distance (FID)**, and their generated images were qualitatively assessed.

## 2. Comparative Table of GAN Variants

| Model | Stability | Mode Collapse | Training Speed | IS Score | FID Score | Image Quality |
|---|---|---|---|---|---|---|
| **Conditional WGAN** | Moderate | Low | Slow | 1.1573 ± 0.0065 | 63.8121 | Good |
| **SNGAN** | High | Very Low | Moderate | 1.1435 ± 0.0046 | 35.1275 | Very Good |
| **SAGAN (No Spectral Norm)** | Low | High | Fast | 1.1440 ± 0.0058 | 92.0718 | Poor |
| **Complete SAGAN** | Very High | Very Low | Moderate | 1.1035 ± 0.0039 | 85.2918 | Good |

## 3. Detailed Analysis of Each Model

### 3.1 Conditional WGAN (Task 1)

- **Architecture:** Uses a Wasserstein loss function with gradient penalty and conditions the generator and discriminator on class labels.
- **Observations:**
  - More stable than vanilla GANs but requires careful tuning.
  - Image quality is **good** but not as sharp as in later models.
  - Mode collapse is **low** due to the conditional setting.
  - Training is relatively **slow** due to the gradient penalty computation.
- **Metrics:** IS = 1.1573 ± 0.0065, FID = 63.8121
- **Conclusion:** Performs well but is outperformed by SNGAN and SAGAN in terms of image quality and FID score.

### 3.2 SNGAN (Task 2)

- **Architecture:** Uses **spectral normalization** to stabilize training and improve discriminator regularization.
- **Observations:**
  - Provides **significantly better stability** than WGAN.
  - Mode collapse is almost **eliminated**.
  - Generates **sharper** and **more diverse** images.
  - Training is **faster** than WGAN.
- **Metrics:** IS = 1.1435 ± 0.0046, FID = 35.1275
- **Conclusion:** SNGAN **outperforms WGAN** in stability, training efficiency, and image quality. Image quality in this SNGAN is really good.

### 3.3 SAGAN without Spectral Normalization and TTUB (Task 3)

- **Architecture:** SAGAN but without **spectral normalization** and **TTUB (Truncated Truncated Upper Bound)**.
- **Observations:**
  - **Unstable** training, leading to frequent mode collapse.
  - Poor quality and **noisy artifacts** in generated images.
  - Fast training but at the cost of low-quality outputs.
- **Metrics:** IS = 1.1440 ± 0.0058, FID = 92.0718
- **Conclusion:** Removing spectral normalization negatively impacts performance, making training unstable.

### 3.4 Complete SAGAN (Task 4)

- **Architecture:** Incorporates **self-attention** along with **spectral normalization** to improve feature representation.
- **Observations:**
  - **Good image quality** with **sharp, high-resolution features**.
  - **Most stable training** with **minimal mode collapse**.
  - Generates **realistic images** with **high diversity**.
  - Also observed problem of Overfitting due to model collapse.
  - Training is **moderately fast** but provides superior results.

- **Metrics:** IS = 1.1035 ± 0.0039, FID = 85.2918
- **Conclusion:** The complete SAGAN is the **best-performing** model in terms of both **quantitative scores** and **visual quality**.

# 4. Final Observations and Takeaways

## 4.1 Performance Summary

- **SAGAN (complete)** is the best model overall,but model collapses resulting into overfitting..
- **SNGAN** is a great balance between stability and efficiency.
- **Conditional WGAN** performs well but is outclassed by newer architectures.
- **SAGAN without spectral normalization** suffers from severe instability and poor results.

## 4.2 Lessons Learned

- **Spectral normalization** is crucial for stabilizing training.
- **Self-attention mechanisms** improve image quality by capturing long-range dependencies.
- **Gradient penalty (WGAN)** improves training but slows down convergence.

## 4.3 Recommendations for Future Improvements

- Use a **larger dataset** to improve generalization.
- Experiment with **hybrid models**, combining elements from different architectures.
- Apply **advanced regularization techniques** to enhance training stability.

# 5. Conclusion

This study demonstrates the **progression of GAN architectures** from traditional WGANs to modern self-attention GANs. The findings highlight the importance of **spectral normalization, attention mechanisms, and architectural design choices** in improving GAN performance. **SAGAN with self-attention and spectral normalization achieves the best balance of stability, diversity, and realism.**