Evaluating the Effects of L1 and L2 Regularisation Placement in GAN Architectures

MSc Research Project
Data Analytics

Zuu Zuu Kyaw Shwe
Student ID: x24106585

School of Computing
National College of Ireland

Supervisor: Jaswinder Singh

National College of Ireland
Project Submission Sheet
School of Computing

Student Name:	Zuu Zuu Kyaw Shwe
Student ID:	x24106585
Programme:	Data Analytics
Year:	2025
Module:	MSc Research Project
Supervisor:	Jaswinder Singh
Submission Due Date:	15/09/2025
Project Title:	Evaluating the Effects of L1 and L2 Regularisation Placement in GAN Architectures
Word Count:	5936
Page Count:	20

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature:	*Zuu*
Date:	15th of September

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST:

Attach a completed copy of this sheet to each project (including multiple copies).	☐
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	☐
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	☐

Assignments that are submitted to the Programme Coordinator office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Evaluating the Effects of L1 and L2 Regularisation Placement in GAN Architectures

Zuu Zuu Kyaw Shwe
x24106585

Abstract

Generative Adversarial Networks (GANs) are perceptive to hyperparameter configurations and also unstable during training. Although regularisation is a typical approach to deep learning to help alleviate overfitting and enhance generalisation, its potential in GAN training has not been thoroughly investigated, particularly in regard to placement and type. This paper is a systematic exploration of the impact of L1 and L2 regularisation as positioned in various points of the GAN architecture such as the output of the generator, input of the discriminator and the weights of both networks. Experiments were first conducted on the DCGAN framework using the CIFAR-10 dataset, covering different configurations evaluated with Frechet Inception Distance (FID), Inception Score (IS), loss curve analysis, and visual inspection. The best performing configuration, applying L1 weight regularisation to the early layers of the discriminator (E5_L1_DiscW), achieved the lowest FID among all tested variants, though the improvement margin was modest. This configuration was successfully transferred to a conditional GAN, resulting in modest gains over the CGAN baseline. However, while some configurations improved FID and IS, visual inspection revealed only subtle perceptual differences in image quality. The results indicate that L1 and L2 regularisation may lead to small gains in stability, but it depends significantly on the architecture, the type of penalty and where it is inserted in the GAN. Future work should explore adaptive penalties, larger datasets, and modern GAN variants to realise more substantial gains in generative performance.

1 Introduction

1.1 Background and Motivation

Generative Adversarial Networks (GANs) are a type of deep learning model that learns the distribution of a dataset to produce highly realistic synthetic data. A GAN comprises two neural networks: a generator, which produces synthetic data, and a discriminator, which attempts to distinguish between real and generated samples. These two networks are trained in opposition, creating a dynamic game like scenario where each network attempts to outperform the other (Goodfellow et al.; 2014).

GANs are notoriously hard to train, in spite of their potential. The mode collapse, vanishing gradients, and oscillatory convergence are still common, and they are usually caused by the delicate balance between the generator and the discriminator (Kurach et al.; 2019). There are numerous solutions that have been suggested to tackle these issues, among them are loss function reformulations (such as Wasserstein loss (Arjovsky et al.; 2017)), architectural improvements (such as progressive growing and attention mechanisms). Although such techniques can work, they normally involve major architectural modifications or intensive computational loads, making it difficult to use or reproduce.

A lot of work has been done to understand more regularisation techniques such as the L1 and L2 penalties, in the supervised learning setting, to reduce overfitting and maximise generalisation. They are both computationally lightweight and architecture independent, which makes them appealing to situations in which resources or implementation complexity is a bottleneck. But they have not been studied as much in the context of adversarial training stabilisation, in terms of where along the GAN architecture they should be used, or whether effects generalise across variants of the GAN.

1.2 Problem Statement

Advanced stabilisation methods can enhance the performance of GANs at the expense of simplicity, whereas L1 and L2 regularisation are simple and underutilised in adversarial learning. There is limited evidence in the literature about whether it is desirable to place such penalties on particular parameters, e.g. generator weights or discriminator weights, outputs or inputs, and whether any positive effects in one type of GAN might be transferable to another, e.g. between DCGAN and Conditional GAN (CGAN).

1.3 Research Aim and Objectives

The aim of this research is to determine whether strategically applied L1 and L2 regularisation can improve GAN training stability and performance, and whether such effects are transferable across architectures.

To achieve this aim, the study pursues the following objectives:

- Implement a baseline Deep Convolutional GAN (DCGAN) and systematically apply L1 and L2 regularisation at generator weights, discriminator weights, generator outputs, and discriminator inputs.
- Evaluate each configuration using quantitative metrics (Frechet Inception Distance, Inception Score) and qualitative measures (image samples, loss curves).
- Transfer the best performing regularisation setting from DCGAN to CGAN and assess whether the effects appear to be similar.

1.4 Research Question

How do different placements of L1 and L2 regularisation affect the stability and performance of GAN training, and can these effects generalise across architectures?

1.5 Significance of the Study

This paper explores a lightweight yet replicable method of GAN stabilisation without having to make complicated architectural adjustments. By systematically comparing multiple placements of L1 and L2 penalties under controlled settings, the work seeks to clarify when, where, and how these methods are most effective. Even relatively small increases in stability may be useful to practitioners with computational limits or who want to make interpretable design decisions.

1.6 Scope and Limitations

This research will focus only on the experiments of the CIFAR-10 dataset with DCGAN and CGAN as sample architectures. Fixing the regularisation strengths at $\lambda_{L_1} = 0.0001$ and $\lambda_{L_2} = 0.001$ and training all models to the same number of 100 epochs to be able to compare all the experiments in a consistent manner. Evaluation is based on FID and IS metrics, which may not capture all aspects of generative quality. The interest is in isolating the impacts of regularisation placement, as opposed to maximising absolute performance.

1.7 Structure of the Report

Section 2 reviews existing literature on GAN regularisation and stability. Section 3 outlines the experimental design and dataset. Section 4 describes the model architectures and regularisation configurations. Section 5 details the implementation setup. Section 6 presents the results and analysis, followed by Section 7 contains the key findings of the study and future directions.

2 Related Work

2.1 Training Instability in GANs

Although GANs have achieved success in generating realistic synthetic data, training GANs is notoriously unstable, subject to such issues as mode collapse, vanishing gradients, and oscillatory training. Such difficulties are sometimes due to the adversarial nature of the relationship between the generator and the discriminator that may imply imbalance and failure to converge (Salimans et al.; 2016).

Those difficulties have been attempted to measure through several large scale studies, with (Kurach et al.; 2019) finding that GAN performance is highly sensitive to hyperparameters and architectural design, and (Lucic et al.; 2018) finding that no single GAN variant performs consistently better than others across datasets, pointing to the lack of generalisability in GAN training. (Arora et al.; 2017) also analyzed theoretical limits of GAN generalization and evaluation, revealing that small performance differences can arise purely from sampling variance rather than true model superiority.

The process of evaluation is another issue. Measures like Frechet Inception Distance (FID) and inception score (IS), which are commonly used, are good but sensitive to some artifacts and show variable ranking (Barratt and Sharma; 2018). As a result, diagnostic limitations complicate both training and comparative analysis, motivating the search for stabilisation techniques that are robust yet simple to implement.

2.2 Stabilisation Techniques in GANs

Over time, several mechanisms have been developed to mitigate the instability in the process of training GANs. Among them are reformulations of the loss functions, architectural developments and regularisation heuristics.

One of the most notable advancements was the introduction of the Wasserstein GAN (WGAN) (Arjovsky et al.; 2017), which replaced the Jensen–Shannon divergence with the Earth-Mover (Wasserstein-1) distance, leading to smoother gradients and improved training stability. (Gulrajani et al.; 2017) enhanced this approach with a gradient penalty (WGAN-GP), enforcing Lipschitz continuity to reduce mode collapse and divergence. While highly effective, these methods require additional architecture changes and tuning.

Architectural innovations have also played an important role. DCGAN framework (Radford et al.; 2016) proposed efficient design choices like the use of strided convolutions and batch normalisation and these ideas became fundamental in GAN research. More recent models, such as StyleGAN (Karras et al.; 2021) and BigGAN (Brock et al.; 2019), provided much higher image fidelity through progressive growing and attention, although they are both more complex and require more computation.

Strategies that involve normalization and regularisation such as spectral normalization (Miyato et al.; 2018) and self-supervised training strategies (Chen et al.; 2019), have enhanced model robustness. Gradient based penalty functions like R1/R2 gradient penalties to stabilize the discriminator proposed by (Mescheder et al.; 2018) to penalise steep gradients. Such advanced methods generally come with heavy implementation overhead, and lack experimentation with L1/L2 style placements of regularisation.

2.3 L1/L2 Regularisation in Deep Learning and GANs

L1 and L2 regularisation are well known methods in supervised learning; they can be applied in an effort to improve generalisation and prevent overfitting: L1 regularisation imparts sparsity, and L2 regularisation imparts smoothness in the distribution of the weights involved (Ng; 2004). However, their effect in adversarial training settings is less studied.

While many GAN regularisation methods exist, few directly apply L1/L2 penalties. (Roth et al.; 2017) introduced a simple discriminator output regulariser; (Tseng et al.; 2021) applied a limited data regulariser mechanism based on moving average predictions. These methods, while relevant, differ from direct weight or input/output penalisation.

A more general context, including regularisation and normalisation of GAN training, is discussed by (Li et al.; 2023) who, however, do not systematically investigate the application of placement specific L1/L2 in this respect. None of these works investigate whether regularisation applied at different model components yields consistent improvements across GAN architectures.

2.4 Limitations of Advanced Stabilisation Approaches

Most modern state-of-the-art stabilisation methods (e.g., WGAN-GP, StyleGAN2, DiffAugment) produce encouraging results but tend to be harder to operate because of their overhead and, in many cases, sensitivity to detail, making them inaccessible to practitioners. For instance, gradient penalties in WGAN-GP introduce additional backward passes that increase training time and require tuning of penalty coefficients (Gulrajani et al.; 2017). Similarly, StyleGAN has an additional complexity of the progressive growing and mapping networks which tends to demand expensive GPUs and long training hours (Karras et al.; 2021).

This is restrictive to use in low resource context or learning. In addition, it is hard to reproduce a result when weights are not pretrained or fine tuned hyperparameters are not presented. The focus on these issues has prompted some researchers to reconsider established methods, including weight decay, dropout, and others, which are easier to apply and less demanding in terms of resources, even if less powerful in isolation (Brock et al.; 2019).

2.5 Regularisation as Inductive Bias in GANs

Another angle to consider is the role of regularisation as a means of introducing inductive bias into the training process. In conventional supervised learning, L1 and L2 regularisation help encode assumptions about sparsity or smoothness, effectively constraining the hypothesis space. In GANs, however, where training involves a min–max game between two networks, such inductive biases may have more subtle or even conflicting effects (Arora et al.; 2017).

For example, regularising the generator too strongly may limit its capacity to capture the target distribution, while excessive constraints on the discriminator might weaken its adversarial signal. The complexity of this implies that assumptions about regularisation improving generalisation do not map well to GANs.

2.6 Impact of Configuration Choices on GAN Behaviour

Several studies have highlighted the wide range of possible configurations in GANs, including variations in architectures, loss functions, training routines, and regularisation methods. (Lucic et al.; 2018) systematically compared dozens of GAN variants under controlled settings. They revealed that tweaking architecture or training schedule could make major performance swings sometimes more than that accomplished by switching to new loss functions or models.

Such findings underscore that evaluation must control for confounding factors. When analysing regularisation effects like in this project, architecture, optimiser and data distribution shall be fixed, isolating only the factor under study. Our work contributes to this literature by focusing exclusively on the placement of L1 and L2 regularisation techniques under a consistent training regime.

2.7 Alternative Regularisation Mechanisms in GANs

There are two popular deep learning regularation techniques, Dropout and Batch Normalisation (BatchNorm), controversy still exists as to how they are applied within GANs. The technique of dropout does help in supervised learning, but it was demonstrated that this method interferes with the training dynamics of GANs, particularly the discriminator, where noise is introduced to the adversarial feedback loop (Radford et al.; 2016). BatchNorm has since become a standard component in DCGAN like architectures, helping to facilitate convergence and reduce internal covariate shift.

Notably, neither dropout nor BatchNorm addresses overfitting in the same way as L1/L2 penalties. Dropout acts as implicit model averaging, while weight decay explicitly limits parameter magnitude. In adversarial settings, their comparative effects are still not well understood, and few studies directly contrast them. Our decision to isolate L1 and L2 penalties in this study reflects this gap and aims to disentangle their effects from other regularisation forms.

2.8 Evaluation Metrics for GANs

Quantitative comparison of GANs most commonly employs Frechet Inception Distance (FID) (Heusel et al.; 2017) and Inception Score (IS) and found that FID is likely to be more sensitive to small differences in sample quality than IS is, especially when IS values are widely clustered. Nonetheless, both metrics have limitations: FID assumes that the extracted feature activations follow a Gaussian distribution, and IS does not directly compare against the real data distribution.

2.9 Summary and Research Gap

GAN stabilisation literature has indicated the existence of a large number of advanced techniques, which include gradient penalties, spectral normalisation, and construction of new architectures, which can significantly boost image realism and convergence. But such methods tend to be complex to apply, computational demanding and often sensitive to details of implementation making them to lack accessibility and reproducibility in low resource or educational settings.

In comparison, L1 and L2 regularisation is understudied in an adversarial settings. While well established in supervised learning for controlling overfitting and improving generalisation, their impact in GAN training particularly with respect to where they are applied within the architecture has received little systematic attention. Previous work seldom examines placement specific application to components such as generator outputs, discriminator inputs, or internal weights, nor does it test whether such effects transfer across GAN variants.

This study addresses this gap through a controlled evaluation of placement specific L1 and L2 regularisation in DCGAN and CGAN models, holding all other factors constant to isolate the influence of regularisation. The findings reveal that targeted L1 weight regularisation on early discriminator layers yields modest yet consistent improvements in FID for DCGAN and demonstrates some transferability to CGAN, though overall performance gains remain limited under short training periods and basic architectures. These results underscore both the potential and the constraints of L1 and L2 regularisation in GANs, and provide a reproducible framework for future investigations into lightweight stabilisation techniques.

3 Methodology

This section outlines the experimental framework used to evaluate the impact of L1 and L2 regularisation strategies on the training performance and stability of Generative Adversarial Networks (GANs). The process is based on the CRISP-DM process model adapted to academic experimentation in the deep learning area.

3.1 Data Understanding

This study uses CIFAR-10, which is a common benchmark dataset to evaluate generative models. The dataset comprises 60,000 colour images of resolution 32×32 pixels, evenly distributed across 10 object classes. Only the 50,000 training images were used for model training, following standard GAN practice (Radford et al.; 2016). Using the entire training split ensures maximum diversity for the generator and discriminator, which is essential for modelling.

3.2 Data Preparation

In the DCGAN experiments, no class filtering or subsetting was applied to all the CIFAR-10 images, which were normalized to the range of $[-1, 1]$ in order to fit the Tanh output of the generator. In the CGAN experiments, in addition to the same normalization, class conditioning was introduced by first mapping the class labels to learnable embeddings via an nn.Embedding layer. For the generator, the label embedding was concatenated with the input noise vector before being passed through the transposed convolutional layers. For the discriminator, the label embedding was spatially replicated to match the input image size and concatenated with the image tensor along the channel dimension.

3.3 Modelling

The modelling part entailed training a Deep Convolutional GAN (DCGAN) and Conditional GAN (CGAN) with a broad range of regularisation settings. Both of the penalties, namely L1 and L2 were independently applied at four distinct architectural locations: generator weights, generator outputs, discriminator inputs, and discriminator weights. A total of eleven configurations (2 baselines + 9 regularised variants) were evaluated per GAN architecture.

3.4 Evaluation

Quantitative and qualitative measures were utilized to measure the model performance so that adequate assessment of generative quality and training stability could be achieved.

Quantitative Metrics Two widely adopted metrics were employed:

- **Frechet Inception Distance (FID)**: Computes the Frechet distance between the feature distributions of real and generated images. Lower scores indicate higher similarity and better generation quality.
- **Inception Score (IS)**: Evaluates image quality and diversity based on predicted class probabilities. Higher scores are preferred.

Qualitative Analysis Visual inspection was conducted through:

- **Generated image samples** at regular epochs, using fixed latent noise vectors to monitor generative progression.
- **Loss curves** for both generator and discriminator, used to diagnose training dynamics and convergence behaviour.

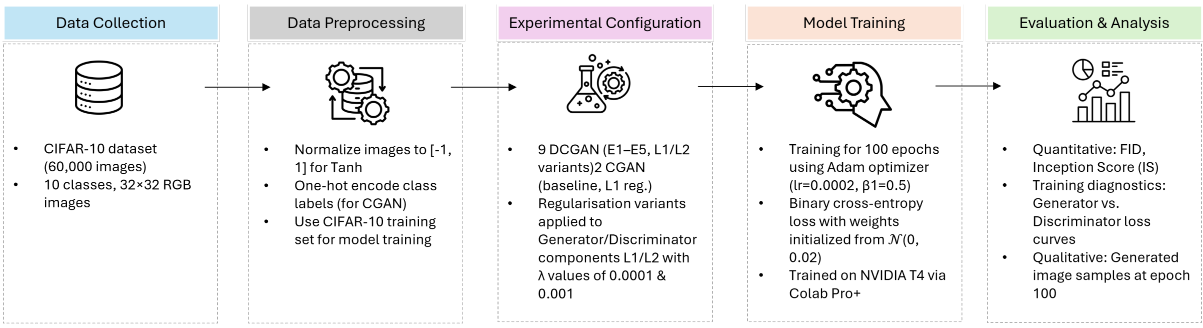


Figure 1: Overview of Experimental Methodology

3.5 Reproducibility and Deployment

In this research, the concept of the deployment will not be realized through production deployment since the research project is not implemented in a real world environment; it follows the principles of the CRISP-DM framework that allows reproducibility as the type of academic deployment. Each trained model, code, and evaluation scripts have been organized and documented to permit replication of the results by others. Each experiment has an archived set of sample outputs, and configuration files, to aid transparency in comparison and future reuse. This is to make sure that results are not simply substantiated but that they are also extendable in subsequent studies or benchmarking activities.

4 Design Specification

4.1 Model Architectures

4.1.1 DCGAN

The Deep Convolutional GAN (DCGAN) served as the baseline generative model. The generator maps a 100 dimensional noise vector to a $32 \times 32 \times 3$ RGB image through four transposed convolutional layers (kernel sizes 4×4 or 3×3, stride 1 or 2), each followed by batch normalisation and ReLU activation, and concluding with a Tanh output layer. The discriminator mirrors this design with four strided convolutional layers, batch normalisation, and LeakyReLU activations, ending with a sigmoid output to produce a scalar real/fake probability. Both networks were trained with the Adam optimiser (learning rate 0.0002, $\beta_1 = 0.5$) and binary cross-entropy loss.

4.1.2 CGAN

The Conditional GAN (CGAN) extends DCGAN by conditioning both generator and discriminator on class labels. A learnable embedding layer maps each label to a 10 dimensional vector, which is concatenated with the noise vector for the generator, and spatially replicated and concatenated with the input image for the discriminator. Other than this conditioning mechanism, the convolutional structure, activation functions, and training hyperparameters are identical to DCGAN, ensuring a controlled comparison.

4.2 Experimental Design and Regularisation Strategies

There were eleven experiments carried out (E1-E7) to test the influence of L1 and L2 regularisation strategies used in different positions within the DCGAN and CGAN architectures. The aim of these experiments was to see the impact of varying regularisation on training stability, convergence behaviour and generative performance.

Table 1 summarises the experimental configurations.

Table 1: Summary of regularisation experiments across DCGAN and CGAN architectures

Exp ID	Model	Reg. Type	λ_{L1}	λ_{L2}	Placement Description	Justification
E1_Baseline	DCGAN	None	0	0	No regularisation	Establishes a reference point for assessing the effect of regularisation
E2_L1_GenOut	DCGAN	L1	0.0001	0	Direct output from G	Constrains pixel level outputs → smoother, coherent images before discriminator
E2_L2_GenOut	DCGAN	L2	0	0.001	Direct output from G	Constrains pixel level outputs → smoother, coherent images before discriminator
E3_L1_DiscIn	DCGAN	L1	0.0001	0	Input fed into D (fake images)	Regularises discriminator input → improves robustness, stability
E3_L2_DiscIn	DCGAN	L2	0	0.001	Input fed into D (fake images)	Regularises discriminator input → improves robustness, stability
E4_L1_GenW	DCGAN	L1	0.0001	0	All transposed conv layers	Limits generator capacity → reduces overfitting
E4_L2_GenW	DCGAN	L2	0	0.001	All transposed conv layers	Limits generator capacity → reduces overfitting
E5_L1_DiscW	DCGAN	L1	0.0001	0	First 2 conv layers	Regularises early features → more stable, general feedback
E5_L2_DiscW	DCGAN	L2	0	0.001	Last 2 conv layers	Regularises early features → more stable, general feedback
E6_CGAN_Baseline	CGAN	None	0	0	No regularisation	Provide a conditional GAN reference for comparison with regularised versions
E7_CGAN_L1_DiscW	CGAN	L1	0.0001	0	First 2 conv layers	Transferred the best configuration from DCGAN

Regularisation was applied by augmenting the loss function of the generator or discriminator with an additional penalty term. For weight based regularisation, PyTorch's built in parameter norm utilities were used. For output and input based regularisation, the L1 or L2 norm was computed directly on the generator outputs or discriminator input batches, respectively, prior to backpropagation. The regularisation strength was fixed at $\lambda = 0.0001$ for L1 and $\lambda = 0.001$ for L2.

4.3 Design Rationale

DCGAN and CGAN were selected because of their simplicity of architecture, interpretability and established position related to generative modelling studies. The models achieve a level of expressiveness that is sufficient to study the effect of regularisation, where such effects are not confounded by the effects of large scale architectural optimisation.

CIFAR-10 dataset has been chosen as a benchmark because it has a moderate level of complexity, its computational complexity is not high, and it is widely used to evaluate GANs. Its standardised nature together with its frequent use in other studies makes it easy to compare with existing studies.

These choices ensure that any observed differences in training stability or generative performance can be attributed primarily to the placement and type of regularisation, rather than confounding variables such as dataset difficulty or architectural novelty.

5 Implementation

5.1 Codebase and Environment

The project was structured as modular Python scripts and Jupyter notebooks where each experimental configuration (E1- E7) is located in a directory of its own. The directories contained a definition and training scripts of the corresponding DCGAN and CGAN architectures, visualisation scripts of loss curve and generated samples, visualisation functions and scripts to calculate Frechet Inception Distance (FID) and Inception Score (IS).

All code was implemented using:

- **Language:** Python 3.10.0
- **Frameworks:** PyTorch 2.6.0+cu124, NumPy 2.0.2, Matplotlib 3.10.0, Torchvision 0.21.0+cu124
- **Platform:** Google Colab Pro+ with NVIDIA Tesla T4

5.2 Training Configuration

The following settings were applied consistently across all experiments to ensure fair comparisons and reproducibility:

- **Optimiser:** Adam optimiser with learning rate 0.0002, $\beta_1 = 0.5$, $\beta_2 = 0.999$, following best practices from Radford et al. (Radford et al.; 2016)
- **Training schedule:** 100 epochs, batch size of 128, latent vector dimensionality of 100
- **Loss function:** Binary Cross Entropy (BCE) loss
- **Weight initialisation:** Custom normal distribution initialisation (mean 0, standard deviation 0.02)
- **Reproducibility:** Fixed random seeds used across all experiment setups

5.3 Generated Outputs

Each experiment directory is structured as follows:

- images/ - generated samples saved at specified epochs

- metrics/losses.csv - training losses
- metrics/loss_curve.png - plotted loss curve
- metrics/eval_metrics.txt - calculated FID and Inception Score
- eval_data/real/ - real CIFAR-10 images used for metric calculation
- eval_data/fake/ - generated images used for metric calculation
- checkpoint.pt - model state and optimizer state
- fixed_noise.pt - fixed latent vectors for consistent sampling
- config_manual.txt - experimental configuration details
- training_time.txt - total training duration

6 Evaluation

6.1 Quantitative Summary of Results

The below charts present the Frechet Inception Distance (FID) and Inception Score (IS) of all the experiments done over the DCGAN and CGAN models. The baseline DCGAN model (E1) had a FID of 47.63 which will be used as a basis of comparison. E5_L1_DiscW provided the lowest FID (44.51) compared to any of the configurations, thus having the best generative performance. Inception Scores were generally comparable across the DCGAN variants, indicating similar overall diversity. Nevertheless, because the training duration was relatively short, the generated images had lower sharpness, making qualitative differences less noticeable. FID was more sensitive in this respect to differentiating configurations. Based on these results, E5_L1_DiscW was applied to the CGAN architecture, yielding a modest improvement over the CGAN baseline, though performance remained substantially below that of the best DCGAN variant.

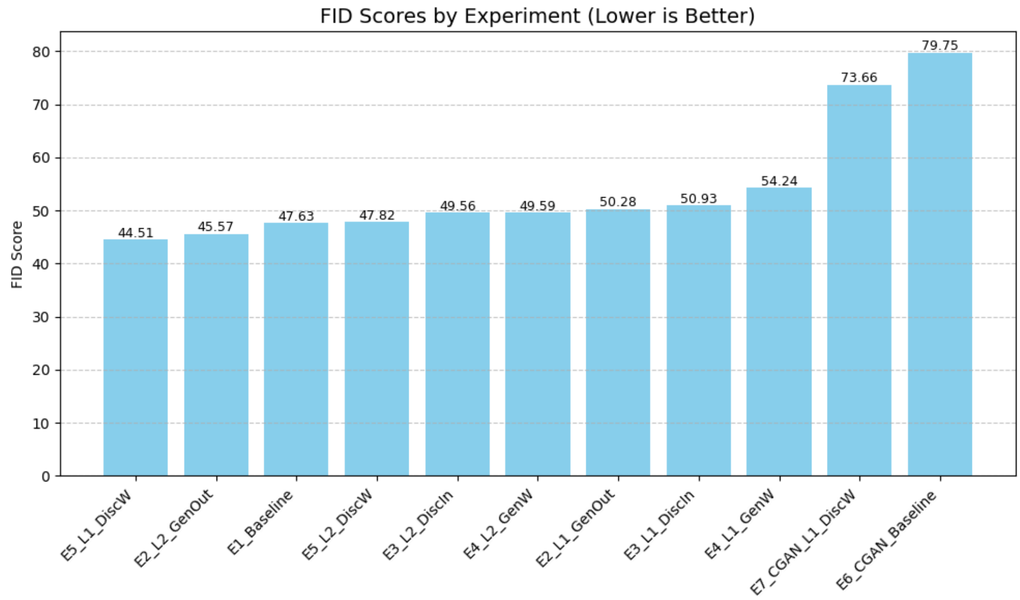


Figure 2: FID scores across all DCGAN and CGAN experiments. Lower FID indicates better similarity to real image distribution. The DCGAN baseline (E1) achieved 47.63, while E5_L1_DiscW gave the best score (44.51), showing a modest improvement with L1 regularisation on early discriminator weights. CGAN models performed worse overall, with the CGAN baseline reaching 79.75, but applying E5_L1_DiscW slightly reduced FID to 73.66.

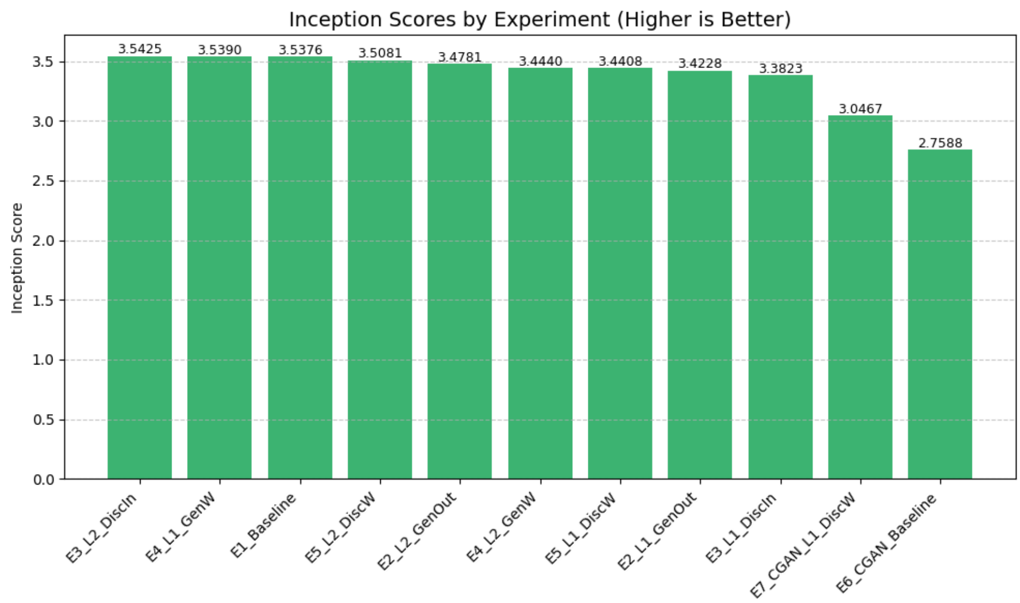


Figure 3: Inception Scores (IS) across experiments. Higher IS indicates better diversity and recognisability of generated images. Scores were largely similar across DCGAN variants, showing that regularisation did not significantly affect sample diversity. CGAN models underperformed, with the baseline achieving the lowest IS (2.76) and showing only a slight improvement after applying E5_L1_DiscW.

Table 2: Summary of FID and IS scores across experiments

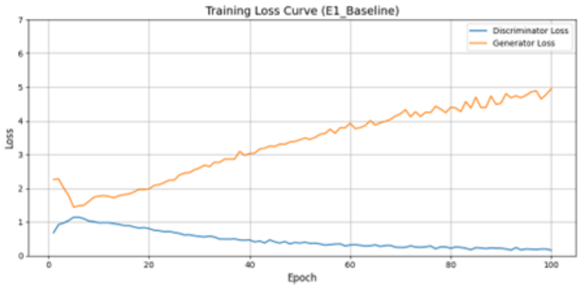
Exp ID	Model	FID	IS	Notes
E1_Baseline	DCGAN	47.63	3.54 ± 0.08	Unregularised baseline
E2_L1_GenOut	DCGAN	50.28	3.42 ± 0.07	L1 on Generator output
E2_L2_GenOut	DCGAN	45.57	3.48 ± 0.08	L2 on Generator output
E3_L1_DiscIn	DCGAN	50.93	3.38 ± 0.08	L1 on Discriminator input
E3_L2_DiscIn	DCGAN	49.56	3.54 ± 0.09	L2 on Discriminator input
E4_L1_GenW	DCGAN	54.24	3.54 ± 0.11	L1 on Generator weights
E4_L2_GenW	DCGAN	49.59	3.44 ± 0.06	L2 on Generator weights
E5_L1_DiscW	DCGAN	44.51	3.44 ± 0.10	L1 on Discriminator weights
E5_L2_DiscW	DCGAN	47.82	3.51 ± 0.08	L2 on Discriminator weights
E6_CGAN_Baseline	CGAN	79.75	2.76 ± 0.05	Unregularised CGAN baseline
E7_CGAN_L1_DiscW	CGAN	73.66	3.05 ± 0.06	L1 regularised CGAN (E5 transferred)

6.2 Loss Curve Analysis

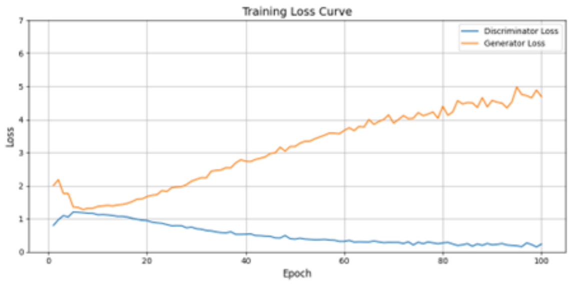
The training loss curves across all experiments displayed broadly consistent patterns: discriminator loss generally decreased over time, indicating improved capacity to distinguish real from fake samples, while generator loss steadily increased, reflecting the growing challenge of deceiving a more capable discriminator. However, a number of configurations showed some significant variations in training dynamics.

In particular, models with L1 or L2 regularisation applied to the generator or discriminator weights (e.g., E4_L1_GenW, E4_L2_GenW) exhibited flatter discriminator loss trajectories during early epochs. The training loss curves did not display abrupt divergence or extreme oscillations, suggesting numerically stable optimisation. However, given the low resolution of CIFAR-10 and the limited training duration, the generated samples were too coarse to reliably assess diversity, and subtle mode collapse cannot be ruled out.

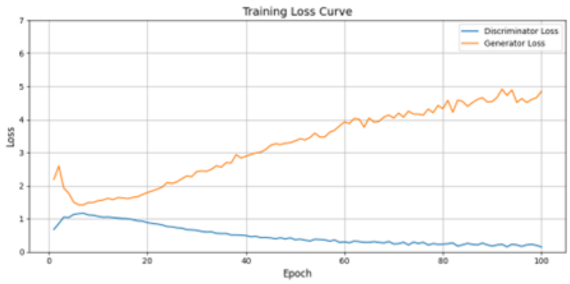
For the CGAN variants, a similar trend was observed. Both the baseline (E6_CGAN_Baseline) and the regularised model (E7_CGAN_L1_DiscW) showed decreasing discriminator loss and increasing generator loss over time. However, as with DCGANs, the regularisation did not significantly alter the overall convergence trajectory.



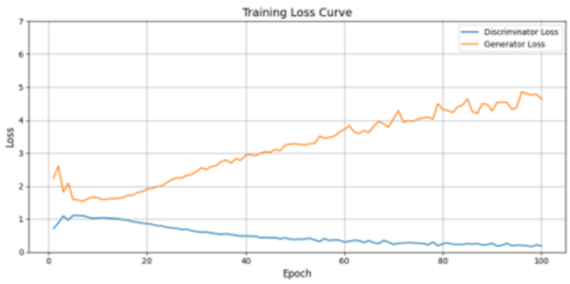
(a) E1_Baseline



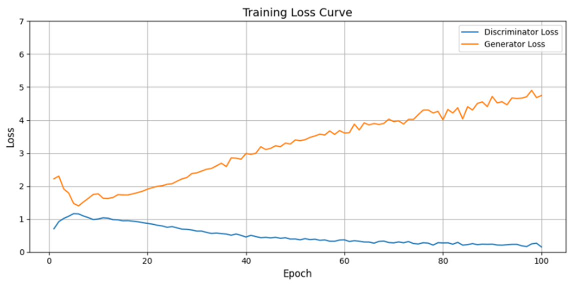
(b) E2_L1_GenOut



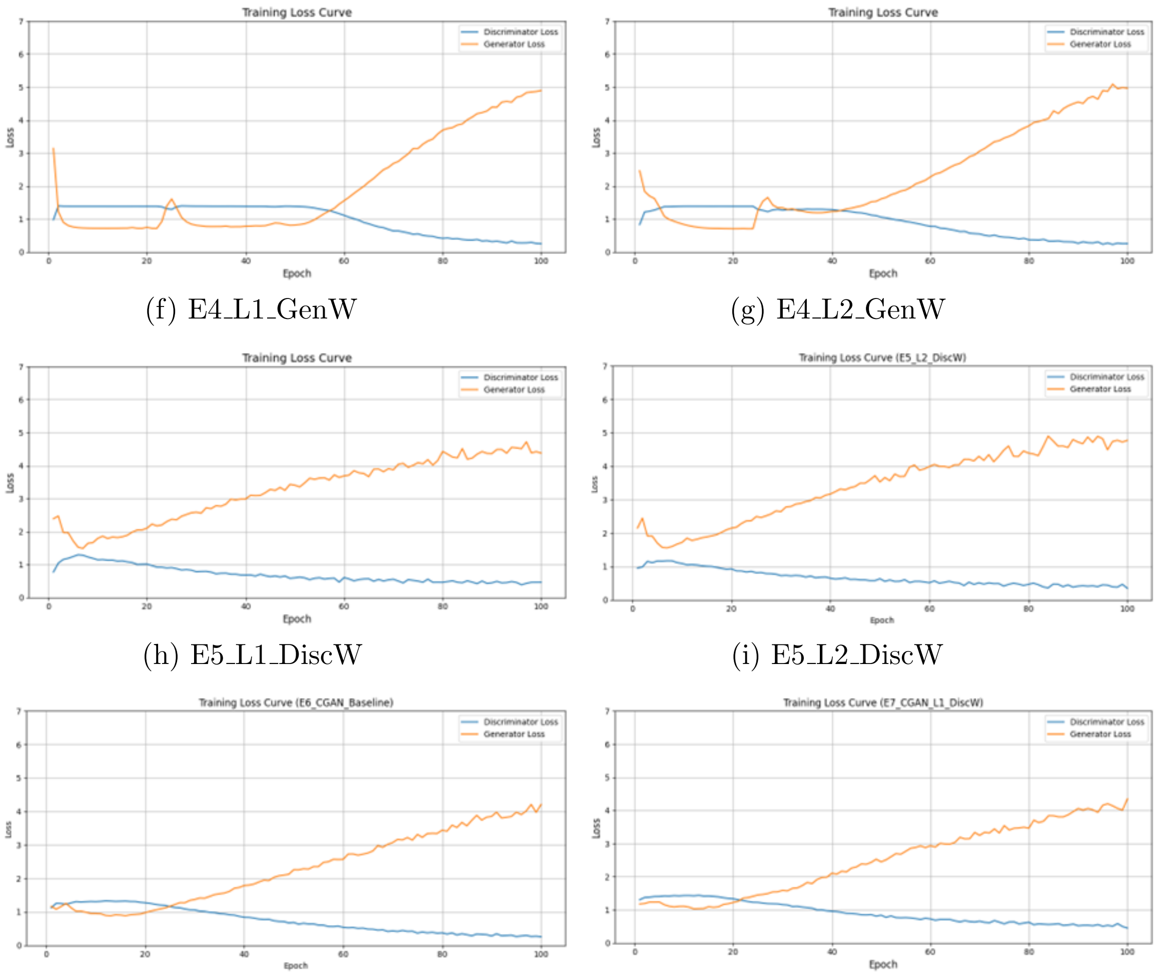
(c) E2_L2_GenOut



(d) E3_L1_DiscIn



(e) E3_L2_DiscIn



(f) E4_L1_GenW
(g) E4_L2_GenW
(h) E5_L1_DiscW
(i) E5_L2_DiscW
(j) E6_CGAN_Baseline
(k) E7_CGAN_L1_DiscW

Figure 4: Training loss curves across all experimental configurations

6.3 Evaluation of Regularisation Strategies

6.3.1 Baseline Models

E1 (DCGAN baseline) achieved an FID of 47.63 and an IS of 3.54 ± 0.08. E6 (CGAN baseline) recorded a substantially higher FID (79.75) and a lower IS (2.76 ± 0.05). These results show that the conditional model underperformed relative to its unconditional counterpart in this setting, which may reflect the added complexity of conditional generation.

6.3.2 Regularisation at Generator Output

E2_L1_GenOut produced a higher FID (50.28) than the baseline, while E2_L2_GenOut achieved a slightly lower FID (45.57). IS values for both were close to the baseline. These observations indicate that L1 regularisation on the generator's output reduced quantitative performance, whereas L2 regularisation had a modest positive effect on FID. The limited change in IS suggests no substantial impact on diversity. Further work would be required to determine whether the difference is related to changes in output smoothness or pixel level sparsity.

6.3.3 Regularisation at Discriminator Input

Both E3_L1_DiscIn (FID 50.93) and E3_L2_DiscIn (FID 49.56) performed worse than the baseline in terms of FID, while IS values remained similar. These results indicate that, in this configuration, applying L1 or L2 penalties to the discriminator's input did not improve quantitative performance. The mechanism underlying this effect was not explored and would require further investigation.

6.3.4 Regularisation at Generator Weights

E4_L1_GenW resulted in the highest FID (54.24) among all DCGAN variants, while E4_L2_GenW recorded a smaller increase over the baseline (49.59). IS values were comparable in both cases. These results suggest that, for this setup, weight regularisation on the generator did not yield performance gains and may have been detrimental in the case of L1. Additional experiments would be needed to assess whether layer specific or adaptive regularisation could mitigate this effect.

6.3.5 Regularisation at Discriminator Weights

E5_L1_DiscW produced the lowest FID (44.51) among all DCGAN variants and E5_L2_DiscW achieved slightly weaker results. These observations may indicate that applying L1 regularisation to discriminator weights can be beneficial in this setting, though further experiments would be required to confirm whether the effect is linked to sparsity in early layers.

6.3.6 Transfer to CGAN

Compared to the CGAN baseline (E6), E7 achieved a lower FID (73.66 vs. 79.75) and a higher IS (3.05 ± 0.06 vs. 2.76 ± 0.05). This outcome indicates an improvement within the conditional setting when using the E5_L1_DiscW regularisation. Both CGAN configurations, however, recorded higher FID scores than any DCGAN variant in this study. Further work would be needed to determine whether alternative or stronger regularisation approaches could reduce this gap.

6.4 Qualitative Results and Visual Comparison

The samples of images were stored with every 10 training epochs to make the qualitative monitoring of generative progression possible. The reason behind this approach was to complement the quantitative measures using a visual impression of convergence, the mode stability and the image fidelity over the time. Nevertheless, the image samples produced by the experiments did not show much visual difference due to the relative lack of resolution of the CIFAR-10 dataset (32×32 pixels), as well as the brief training time of 100 epochs. Subtle improvements in sharpness, structure, or class consistency were difficult to reliably detect by eye.

To help visualise the quantitative results, samples of representative images of the final training epoch (epoch 100) will be shown of the baseline model (E1) and the best performing regularised model (E5_L1_DiscW). However, these samples were not generated from fixed latent vectors, and therefore do not allow for a direct pixel wise comparison across models.

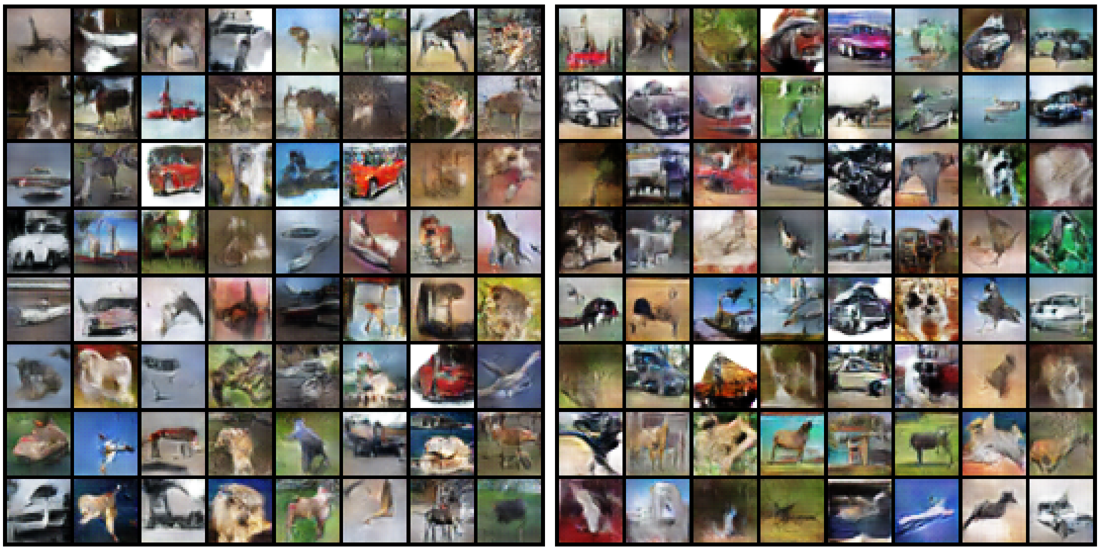


Figure 5: Generated samples at epoch 100 from baseline (left) and best performing model (right)

6.5 Discussion

Experimental findings prove that L1 and L2 regularisation in GANs have their performance dependent on their location and application. Application of regularisation on the generator outputs or discriminator inputs (E2, E3) did not bring significant improvements and on the contrary, in certain cases, it worsened performance, likely indicating that such localisations interfere with the fragile adversarial balance. Weight regularisation, specifically L1 on the first two layers of the discriminator (E5_L1_DiscW), on the other hand, produced the lowest FID score and thus this was chosen to be transferred to the conditional GAN scenario.

Even though the Inception Score (IS) was more or less constant across most DCGAN variants, the Frechet Inception Distance (FID) was more sensitive to changes in architecture or regularisation, making it the more reliable metric in this context. Nevertheless, the highest FID score obtained (44.5) can be considered moderate, meaning that regularisation substantially helped in some but not in all of the stability or noise suppression effects; nevertheless, it did not significantly enhance generative quality within these limits of the study.

The generalisation of the best regularisation setting to the CGAN model (E7) led to a bit of an improvement in FID (79.7 to 73.7), implying that the regularisation effect can be generalised to some extent. Nevertheless, both CGAN models underperformed relative to all DCGAN variants, likely due to the increased training complexity introduced by conditioning or insufficient adaptation of hyperparameters.

Limitations of the study include the relatively short training duration (100 epochs), fixed regularisation coefficients, and uniform application across full layers. Visual sample inspection was inconclusive, as most experiments generated qualitatively similar outputs due to early stage training. Future work should consider longer training, adaptive or layer specific penalties, and comparison with advanced GAN variants (e.g., StyleGAN or Diffusion Models) to better contextualise the impact of regularisation.

Overall, the results indicate that the interventions of regularisation, especially L1 on the early discriminator layers can add some small stability gains, its isolated use may be insufficient to achieve high quality image synthesis on CIFAR-10 under limited training.

7 Conclusion and Future Work

The paper systematically studied effects of approaches to regularisation: L1 and L2 regularisation in the context of GAN type architectures, especially their effects on training stability and generation quality when used between outputs, inputs, and internal weights. In a sequence of well managed experiments with DCGAN and CGAN models on the CIFAR-10 dataset, the results showed that a modest gain in weight based regularisation can be achieved, especially a L1 regularisation of the early layers of the discriminator. Not only did such a configuration have the lowest FID of any DCGAN variant, but it was also shown to be more easily transferred to conditional GANs, where it continued to outperform the CGAN baseline.

Although these gains were not radical, the findings highlight the significance of targeted regularisation, and its capability to stabilise adversarial practice without the need of a major architectural modification. In addition, the paper provides a reproducible experimental analysis structure of studying the effects of regularisation, and it should serve as a starting point of further research in GAN stability.

Future work could extend these insights in several directions. One direction would be to train longer than 100 epochs and see whether the effects of regularisation will be more pronounced with extended optimisation. Another involves investigating adaptive or layer specific regularisation schedules, which may better align regularisation strength with architectural depth or learning dynamics. Moreover, evaluating the methods on other more complicated datasets, e.g., CelebA or subsets of ImageNet, would lead to some evaluation of generalisability. Another direction of research that may present an opportunity of additional complementary advantages that are not on offer by the L1 and L2 norms is the exploration of alternative regularisation methods e.g. spectral norm or gradient penalties. Lastly, extending the methodology to newer generative models, such as StyleGAN or diffusion based systems, would also give more insight into the generalisability of these results.

In brief, L1 and L2 regularisation does not appear to be a solution to the problems with GAN training, but this study shows that L1 and L2 regularisation may have subtle benefits and is a step towards more sophisticated regularisation methods, in the context of generative models.

References

Arjovsky, M., Chintala, S. and Bottou, L. (2017). Wasserstein generative adversarial networks, *in* D. Precup and Y. W. Teh (eds), *Proceedings of the 34th International Conference on Machine Learning*, Vol. 70 of *Proceedings of Machine Learning Research*, PMLR, Sydney, Australia, pp. 214–223.
URL: *https://proceedings.mlr.press/v70/arjovsky17a.html*

Arora, S., Ge, R., Liang, Y., Ma, T. and Zhang, Y. (2017). Generalization and equilibrium in generative adversarial nets (gans), *in* D. Precup and Y. W. Teh (eds), *Proceedings of the 34th International Conference on Machine Learning*, Vol. 70 of *Proceedings of*

Machine Learning Research, PMLR, pp. 224–232.
URL: *https://proceedings.mlr.press/v70/arora17a.html*

Barratt, S. and Sharma, R. (2018). A note on the inception score, *arXiv preprint arXiv:1801.01973* .
URL: *https://arxiv.org/abs/1801.01973*

Brock, A., Donahue, J. and Simonyan, K. (2019). Large scale gan training for high fidelity natural image synthesis, *arXiv preprint arXiv:1809.11096* .
URL: *https://arxiv.org/abs/1809.11096*

Chen, T., Zhai, X., Ritter, M., Lucic, M. and Houlsby, N. (2019). Self-supervised gans via auxiliary rotation loss, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 12146–12155.

Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. and Bengio, Y. (2014). Generative adversarial nets, *Advances in Neural Information Processing Systems*, Vol. 27, Curran Associates, Inc., pp. 2672–2680.

Gulrajani, I., Ahmed, F., Arjovsky, M., Dumoulin, V. and Courville, A. C. (2017). Improved training of wasserstein gans, *in* I. Guyon, U. Von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan and R. Garnett (eds), *Advances in Neural Information Processing Systems*, Vol. 30, Curran Associates, Inc.

Heusel, M., Ramsauer, H., Unterthiner, T., Nessler, B. and Hochreiter, S. (2017). Gans trained by a two time-scale update rule converge to a local nash equilibrium, *in* I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan and R. Garnett (eds), *Advances in Neural Information Processing Systems*, Vol. 30, Curran Associates, Inc.

Karras, T., Laine, S. and Aila, T. (2021). A style-based generator architecture for generative adversarial networks, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **43**(12): 4217–4228.

Kurach, K., Lucic, M., Zhai, X., Michalski, M. and Gelly, S. (2019). A large-scale study on regularization and normalization in gans, *Proceedings of the 36th International Conference on Machine Learning (ICML)*, PMLR, pp. 3581–3590.
URL: *https://arxiv.org/abs/1807.04720*

Li, Z., Usman, M., Tao, R., Xia, P., Wang, C., Chen, H. and Li, B. (2023). A systematic survey of regularization and normalization in gans, *ACM Computing Surveys* **55**(11): 1–37.
URL: *https://doi.org/10.1145/3569928*

Lucic, M., Kurach, K., Michalski, M., Gelly, S. and Bousquet, O. (2018). Are GANs created equal? a large-scale study, *Advances in Neural Information Processing Systems*, Vol. 31, Curran Associates, Inc.

Mescheder, L., Geiger, A. and Nowozin, S. (2018). Which training methods for gans do actually converge?, *Proceedings of the 35th International Conference on Machine Learning*, Vol. 80, PMLR, pp. 3481–3490.
URL: *https://proceedings.mlr.press/v80/mescheder18a.html*

Miyato, T., Kataoka, T., Koyama, M. and Yoshida, Y. (2018). Spectral normalization for generative adversarial networks, *arXiv preprint arXiv:1802.05957* .
URL: *https://arxiv.org/abs/1802.05957*

Ng, A. Y. (2004). Feature selection, l1 vs. l2 regularization, and rotational invariance, *Proceedings of the Twenty-First International Conference on Machine Learning*, Association for Computing Machinery, Banff, Alberta, Canada, p. 78.
URL: *https://doi.org/10.1145/1015330.1015435*

Radford, A., Metz, L. and Chintala, S. (2016). Unsupervised representation learning with deep convolutional generative adversarial networks, *arXiv preprint arXiv:1511.06434* .
URL: *https://arxiv.org/abs/1511.06434*

Roth, K., Lucchi, A., Nowozin, S. and Hofmann, T. (2017). Stabilizing training of generative adversarial networks through regularization, *Advances in Neural Information Processing Systems*, Vol. 30, Curran Associates, Inc.

Salimans, T., Goodfellow, I., Zaremba, W., Cheung, V., Radford, A. and Chen, X. (2016). Improved techniques for training GANs, *Advances in Neural Information Processing Systems*, Vol. 29, Curran Associates, Inc., pp. 2234–2242.
URL: *https://proceedings.neurips.cc/paper/2016/hash/8a3363abe792db2d8761d6403605aeb7-Abstract.html*

Tseng, H.-Y., Jiang, L., Liu, C., Yang, M.-H. and Yang, W. (2021). Regularizing generative adversarial networks under limited data, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 7921–7931.