Training and Evaluation Report

1. Training Setup

Dataset

- Dataset: MNIST
- Image Size: 28 × 28
- Channels: 1 (Grayscale)
- Preprocessing:
    - Resize to 28 × 28
    - Normalize pixel values to [-1, 1]
    - Shuffle and batch data
    - Train/Test split = 95% / 5%

Model Configuration

Generator

- Input: Noise vector of dimension 100
- Architecture:
    - Linear → ReLU
    - Reshape (128 × 7 × 7)
    - ConvTranspose2D → ReLU
    - ConvTranspose2D → Tanh
- Output: (1, 28, 28) image

Discriminator

- Input: Image (1, 28, 28)
- Architecture:
    - Conv2D → LeakyReLU
    - Conv2D → LeakyReLU
    - Flatten → Linear → Sigmoid
- Output: Probability score

Training Parameters

Parameter	Value
Batch Size	64
Epochs	50
Noise Dimension	100
Optimizer	Adam
Learning Rate	0.0002
Beta1	0.5
Loss Function	Binary Cross Entropy
Label Smoothing	Real labels = 0.9

2. Training Procedure

a) Load training data using DataLoader.
b) Initialize Generator and Discriminator.
c) For each batch:
    - Generate fake images from random noise.
    - Compute Discriminator loss on real and fake images.
    - Update Discriminator weights.
    - Compute Generator loss using Discriminator predictions.
    - Update Generator weights.
d) Repeat for all epochs.
e) Save Generator checkpoints and sample outputs periodically.

3. Training Monitoring

- Generator and Discriminator losses logged per epoch.
- Training samples saved at fixed intervals.
- Final trained models saved to disk.

4. Evaluation Methodology

Feature Extraction

- Pretrained InceptionV3 used as feature extractor.
- Auxiliary classifier disabled.
- Final classification layer removed.
- Features extracted for real and generated images.

Evaluation Metrics

i. Fréchet Inception Distance (FID)

- Measures distance between real and generated feature distributions.
- Lower value indicates better similarity.

ii. Diversity Score

- Computed as mean standard deviation across generated features.
- Indicates variation in generated samples.

iii. t-SNE Visualization

- Dimensionality reduction applied to feature vectors.
- Used for visual comparison of real and fake embeddings.

5. Evaluation Results

Metric	Observation

FID Score	Reduced after training
Diversity Score	Stable
t-SNE	Overlapping real and fake clusters
Mode Collapse	Not observed