Lab Evaluation Report

# Synthetic Image Generation and Model Comparison using VAE and GAN

## 1. Dataset Preprocessing

- The mango leaf disease dataset (DIEBACK) was loaded and preprocessed.

- Images were resized to a consistent shape and normalized to a range suitable for model training.

- The dataset was organized class-wise, ensuring structured access for model training.

Preprocessing Techniques:

- Resizing images to 64x64 pixels

- Normalizing pixel values to [-1, 1] for GAN and appropriate range for VAE

## 2. Variational Autoencoder (VAE) Implementation

Architecture:

- Encoder:

- Convolutional layers reducing spatial dimensions while increasing feature maps

- Latent space sampled with mean and log variance vectors

- Decoder:

- Transposed convolutional layers reconstructing the image from the latent representation

Training Details:

- Loss function: Combination of reconstruction loss (Binary Cross Entropy) and Kullback-Leibler divergence

- Optimizer: Adam optimizer with a learning rate of 0.0002

- Number of epochs: 80

Challenges:

- Balancing the reconstruction loss and KL divergence to avoid blurry outputs

- Ensuring stable training without posterior collapse

Generated Samples:

- 100 synthetic images were generated by sampling from the latent space.

## 3. Generative Adversarial Network (GAN) Implementation

Architecture:

- Generator:

- Fully connected layer followed by series of transposed convolutional layers

- Output normalized to range [-1, 1] using Tanh activation

- Discriminator:

- Convolutional layers with LeakyReLU activations

- Final output through a sigmoid to produce a real/fake probability

Training Details:

- Loss function: Binary Cross-Entropy loss

- Optimizer: Adam optimizer with learning rate 0.0002 and betas (0.5, 0.999)

- Number of epochs: 50

Challenges:

- Handling discriminator overpowering generator (solution: label smoothing and careful tuning)

- Avoiding mode collapse (solution: introducing slight noise to real inputs)

Generated Samples:

- 100 synthetic images were generated using the trained generator.

## 4. Model Comparison

Visual Comparison:

- VAE images exhibited blurriness but covered diverse structures.

- GAN images appeared sharper and visually more realistic but sometimes less diverse.

Quantitative Metrics:

- Structural Similarity Index (SSIM):

- VAE: Slightly lower SSIM, indicating blurrier reconstructions

- GAN: Higher SSIM, reflecting better structural preservation

- Fréchet Inception Distance (FID):

- VAE: Higher FID, indicating less realistic images

- GAN: Lower FID, closer to real image distributions

Evaluation Summary:

- GAN outperformed VAE in terms of both image realism and metric scores.

- VAE maintained better latent space smoothness but at the cost of visual sharpness.

## 5. Conclusion and Insights

Model Performance:

VAE :

SSIM - 0.3328

FID - - 429.19

GAN :

SSIM - 0.0302

FID - 443.73

- Best Performing Model: VAE

- Reason: GANs are specifically designed to generate sharp and realistic images, whereas VAEs prioritize latent structure over pixel-level fidelity.

Challenges Faced:

- Tuning hyperparameters for GAN training stability

- Adjusting the KL divergence term in VAE training

Future Improvements:

- Use StyleGAN2-ADA for even higher image quality.

- Implement conditional GANs for class-wise generation.

- Explore Beta-VAE to improve disentanglement in latent space.