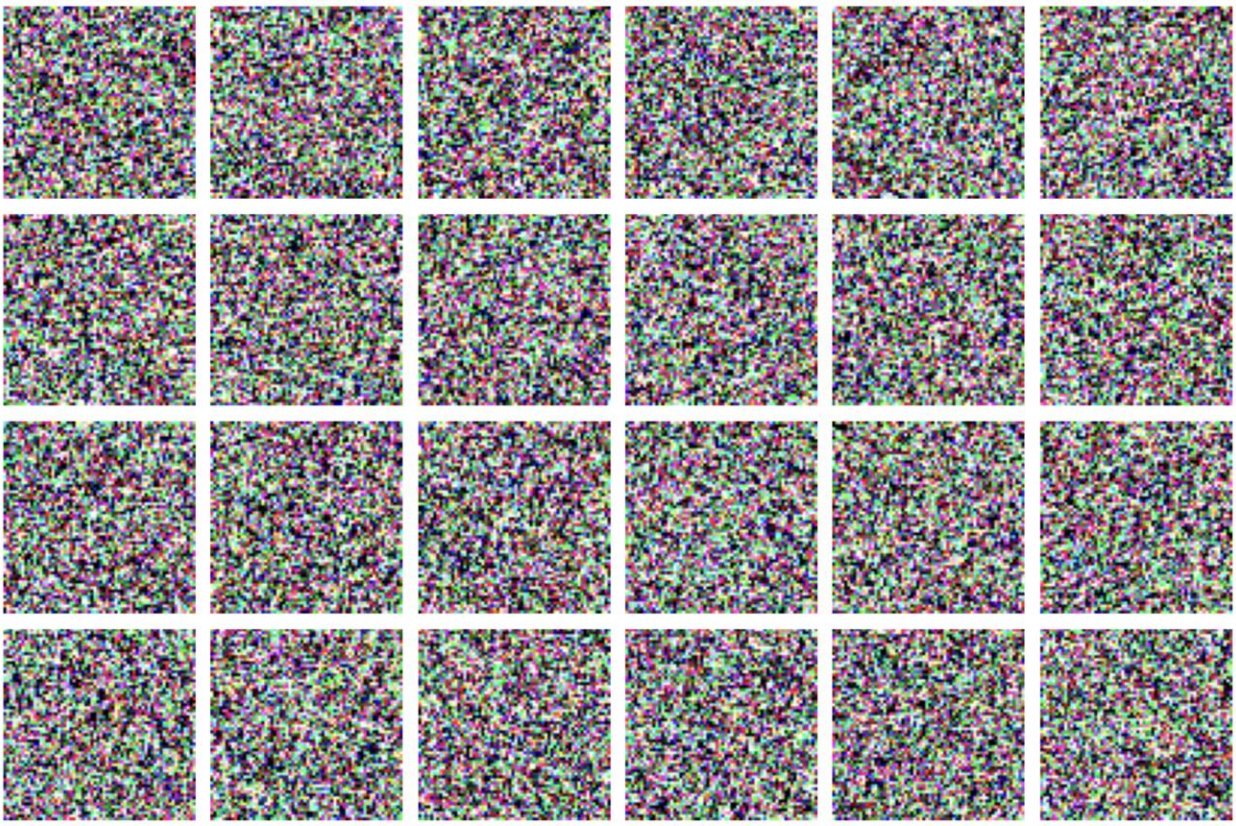
1 st model:



```
model_flow.eval()

fid_value = calculate_fid(model_flow, dataloader, real_image_count=1000, fake_image_count=1000)
print(f"FID: {fid_value}")
```

```
100%|██████████| 16/16 [00:50<00:00,  3.17s/it]
100%|██████████| 16/16 [00:52<00:00,  3.25s/it]
FID: 442.5484216279382
```

It looks like the model is generating images that resemble random noise. Here are some comments and suggestions for improvement:

1. **Noise-like Output**: The images generated show a lot of noise (but still some face lines visible), indicating that the model may not be learning effectively from the data.
2. **Potential Causes**: This can be due to various factors including:
    - Insufficient training data or poor data quality.
    - Poor choice of hyperparameters such as learning rate or batch size.
    - Issues in the loss function or instability during training (unlikely though)

Suggestions for improvement:

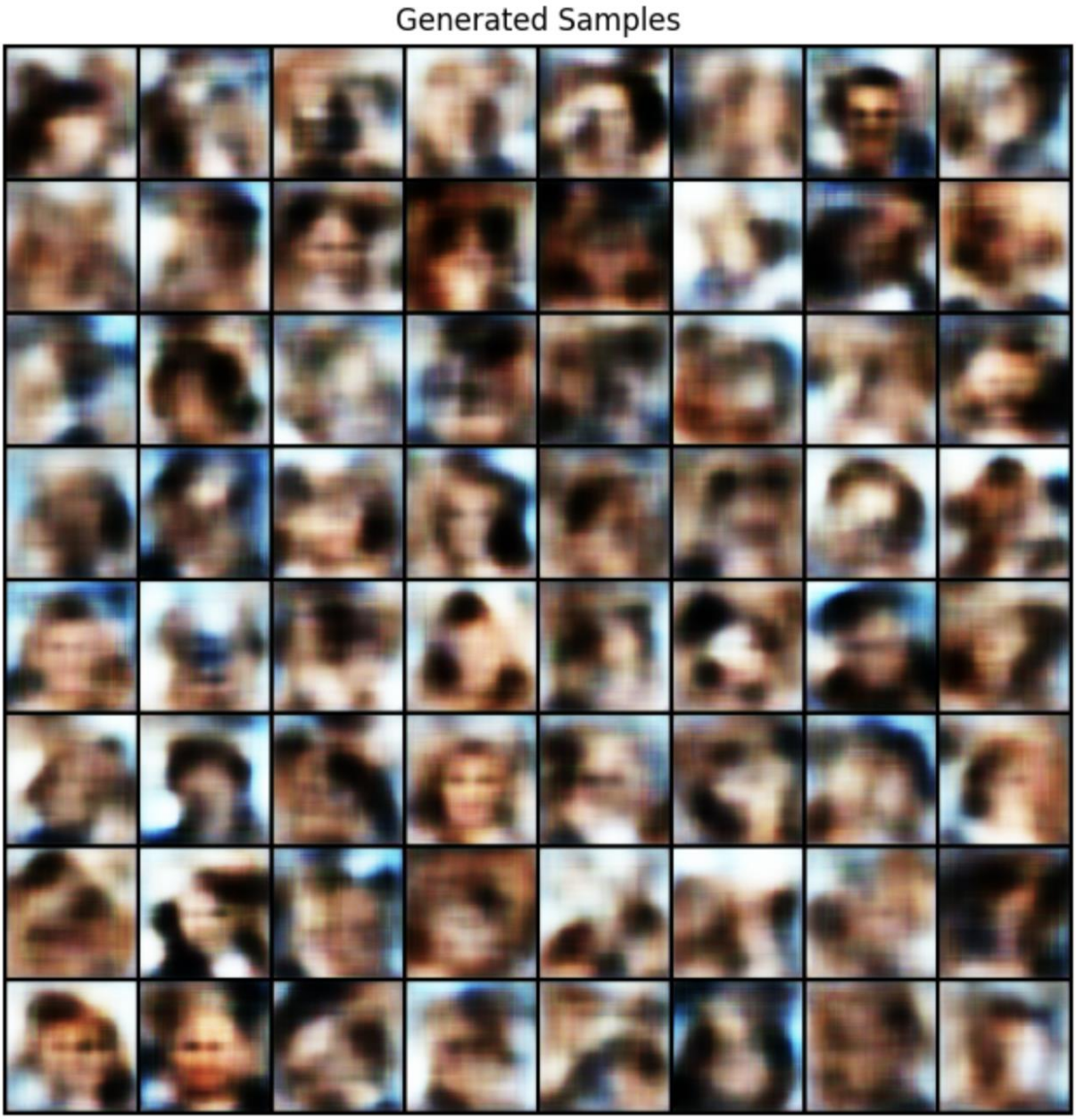
Loss function:

- Ensure the loss function is appropriate for the task. For image generation tasks, loss functions like Mean Squared Error (MSE) or cross-entropy might need to be revisited.
- If using a Variational Autoencoder (VAE), ensure adequate weight on reconstruction vs. KL divergence.

Training process:

- Check for any overfitting or underfitting. Regularization techniques like dropout or weight decay can help.
- Ensure the model has fully converged. If needed, train the model for more epochs.

Model 2:




```
FID Score: 29.6586
```

Shows a lot better results in terms of face recognition, shapes of eyes, hair etc

So, Model 2 is able to generate facial images that, while blurry, contain recognizable features. This indicates an improved learning of data manifold compared to Model 1.

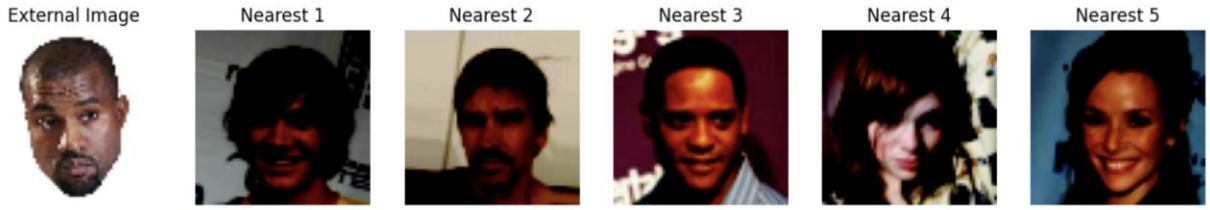
Quality: Faces are not sharp and appear blurry, which suggests that while the model captures structural information, it may need enhancements for detail and resolution.

Model 3 now:

- **Output**: The images are consistent with clear facial features, showing a high degree of similarity across samples. Each image seems to represent an averaged or prototypical face.
- **Characteristics**: Indicates strong learning of specific facial features, but lacks diversity and variance among samples.

  Improving data diversity and adjusting the model can help produce more varied images while maintaining clarity.

  Fun section:



  The nearest variants are kind of reasonable, but 3rd picture looking like the closest one. However, model picks 1st because of dark shades and probable worse clarity of the face expression)