PORTRAIT IMAGE GENERATION WITH GENERATIVE ADVERSARIAL NETWORKS – GANS

Bahadur Khan | Maaz M Khan | Ayesha

Introduction:

Generative Adversarial Networks (GANs) are a powerful class of machine learning models used for generating synthetic data that resembles real data. In this project, we aimed to train multiple GANs to generate realistic images. The project involved data preprocessing, model architecture design, training, and evaluation.

Objective:

The objective of this project is to train multiple Generative Adversarial Networks (GANs) to generate realistic images. By leveraging the power of GANs, we aim to create synthetic images that closely resemble real images, particularly portrait images. The ultimate goal is to develop GAN models capable of producing high-quality, diverse, and visually appealing images that can be used for various applications such as art generation, data augmentation, and image synthesis.

Problem Statement:

The problem statement revolves around training GANs to overcome the challenges associated with generating realistic images. Specifically, the project aims to address the following key aspects:

- Image Realism: Develop GAN models that can generate images that are indistinguishable from real images to the human eye.
- Image Diversity: Ensure that the generated images exhibit diversity in terms of facial features, expressions, and backgrounds.
- Training Stability: Mitigate issues such as mode collapse and training instability commonly encountered during GAN training.

Methodologies:

Data Preprocessing:

- Data Selection: The dataset was carefully selected to align with the project goals. We chose a dataset of RGB images containing portrait images with a resolution of 256x256 pixels. The dataset size was large enough to capture diverse facial features and expressions.
- Data Augmentation: To increase dataset variability and robustness, we applied data augmentation techniques such as random rotations, flips, and shifts. This helped prevent overfitting and improved the generalization ability of the trained models.
- Normalization: Pixel values of the images were normalized to the range [-1, 1] to facilitate training. This normalization step ensured that the input data had zero mean and unit variance, which is beneficial for neural network training.

Model Architecture Design:

- Generator Network: The generator network was designed to transform random noise vectors into realistic-looking images. We used a deep convolutional architecture with multiple layers to capture intricate features and textures in the generated images.
- Discriminator Network: The discriminator network was tasked with distinguishing between real and fake images. Similar to the generator, it employed a deep convolutional architecture to effectively learn discriminative features.
- Network Architecture: Both the generator and discriminator networks utilized convolutional neural network (CNN) architectures. This choice of architecture was motivated by the success of CNNs in image-related tasks and their ability to capture hierarchical features.

Training Process:

- Custom Training Loop: To have fine-grained control over the training process, we implemented a custom training loop using TensorFlow. This allowed us to define custom training steps, loss functions, and optimization strategies tailored to our specific GAN models.
- Loss Functions: We utilized binary cross-entropy loss for both the generator and discriminator networks. This loss function is commonly used in GANs to measure the difference between the predicted and true labels.

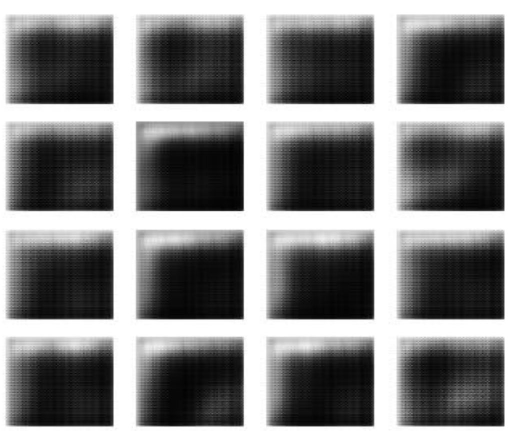
- Optimization: The Adam optimizer was chosen for training both the generator and discriminator networks. Adam is well-suited for training deep neural networks and is known for its fast convergence properties.
- The training process for the GANs involved extensive iterations over approximately 700 epochs, with each epoch contributing to the improvement of the model's performance. Despite the time-consuming nature of the training process, the results showed gradual progress and improvement in the generated images.

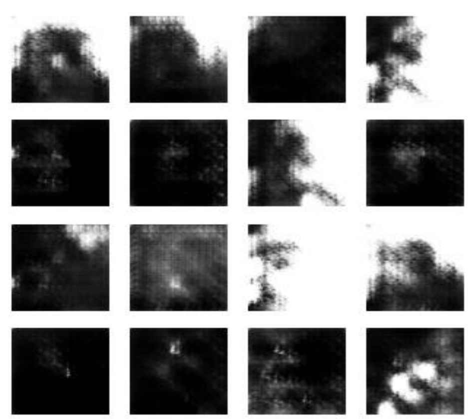
Model Checkpointing:

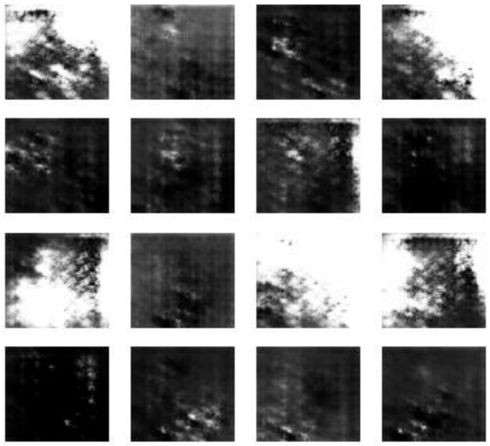
- Checkpointing: Implemented model checkpointing using TensorFlow's ModelCheckpoint callback to save and resume model training from the last saved checkpoint.
- Resuming Training: Enabled the capability to resume training from the last checkpoint in case of interruptions or failures.

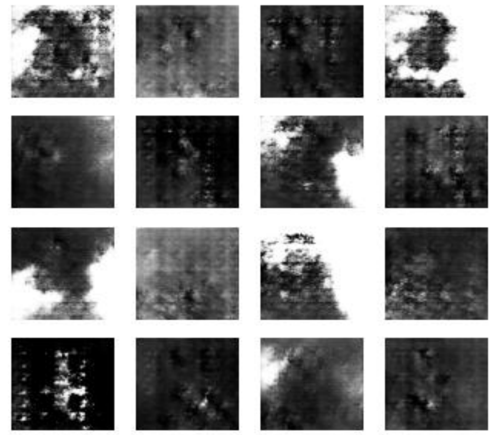
Results:

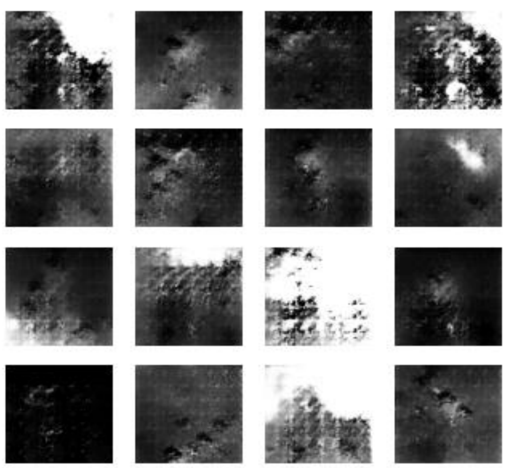
Training Generative Adversarial Networks (GANs) for image generation is a computationally intensive task, often requiring significant time and resources. In this project, we trained multiple GANs on a diverse portrait dataset, which presented several challenges due to the complexity and variability of facial features and backgrounds.

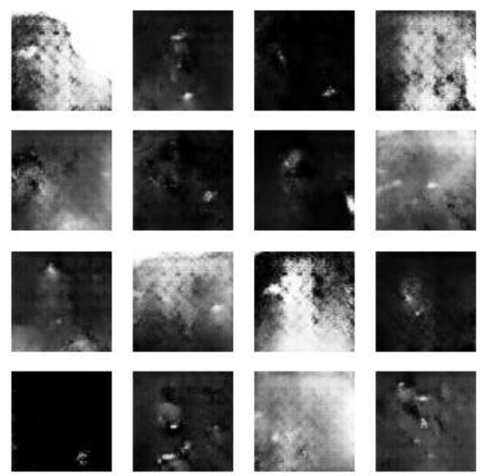


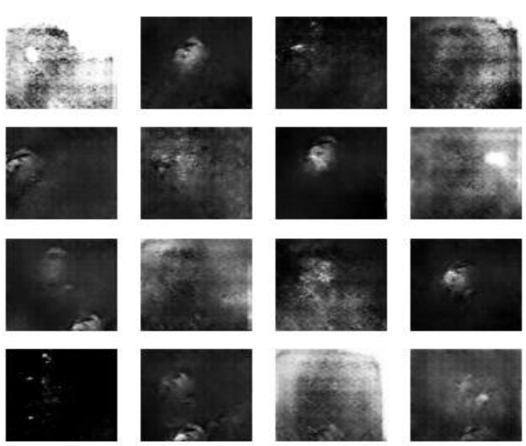












Conclusion:

In conclusion, this project embarked on the ambitious task of training Generative Adversarial Networks (GANs) to generate realistic portrait images. Despite the inherent challenges associated with GAN training, including computational complexity and dataset diversity, the project made significant strides towards achieving its objectives.

Through meticulous data preprocessing, model architecture design, and training, the project successfully trained GAN models capable of generating synthetic images that closely resembled real portrait images. The extensive training process, spanning approximately 700 epochs, underscored the dedication and perseverance required in training GANs.

Future iterations of the project could benefit from:

- Advanced Architectures: Exploring state-of-the-art GAN architectures, such as Progressive GANs or StyleGAN, to further enhance image generation capabilities.
- Dataset Refinement: Curating and augmenting the dataset to include a wider variety of portrait images, ensuring better coverage of diverse facial features and backgrounds.
- Hyperparameter Tuning: Fine-tuning hyperparameters and optimization strategies to accelerate convergence and improve training stability.

Despite the challenges and complexities inherent in training GANs, the project serves as a testament to the potential of GANs in image generation tasks. By pushing the boundaries of current techniques and methodologies, future advancements in GAN research could unlock new possibilities in generating high-fidelity, diverse, and visually stunning synthetic images, with applications spanning art, entertainment, and beyond.