Introduction

Rapid advancements have been made in artificial intelligence (AI). The creation of Generative Adversarial Networks (GANs) is among the most noteworthy accomplishments in this field. By using preexisting face data to identify patterns, these networks are able to produce realistic human faces.

Computer vision has undergone a revolution thanks to Generative Adversary Networks (GANs), which produce realistic and high-quality images, including ones with human faces. In this report, we examine the use of GANs for producing lifelike human faces. We examine the underlying technology of GANs, their difficulties, and potential directions for future research.

Realistic face generation has broad uses in virtual reality, identity verification, and entertainment and gaming.

Why GAN?

Most commercial neural nets can be easily tricked into misclassifying objects with only a small amount of noise added to the original data. Surprisingly, the addition of noise increased the model's confidence in the incorrect prediction relative to the correct prediction. Due to the extreme problem of overfitting caused by the fact that most machine learning models are trained on small amounts of data, adversaries of this type emerge. Furthermore, there is almost a linear mapping from the input to the result. Even a small change in a single point in the feature space could have significant effects because, although the borders dividing the various classes may seem to be linear, they are actually composed of linearity's.

Background of the Project

In this project, the use of Generative Adversarial Networks (GANs) to generate realistic human faces is investigated. The study makes use of the TensorFlow and Keras libraries to implement a GAN architecture in Python. The training dataset is made up of preprocessed and resized human face photos from the designated directory. A discriminator network assesses the realism of the generated faces, which are produced by a generator network. The report explores the implementation's specifics and offers a thorough analysis of the outcomes that were produced.

Two neural networks, a discriminator and a generator, compete with one another to form an ANN. The generator creates synthetic data, and the discriminator evaluates whether the samples it generates are real. Over time, GANs improve through this adversarial training process and produce increasingly realistic outputs.

Objective

The objective of generating realistic human faces using GANs is to create synthetic images that closely resemble real human faces. This has applications in various fields, including art, entertainment, and research.

Data Pre-Processing

The collected data has all the facial images of the celebrities and the set is normalized and used for preprocessing.

The images are preprocessed where the maximum number of images to be processed are present. Their original height and width are given. The picture is modified by cropping as per the requirement and by removing the extra pixels. After the preprocessing of the image, they are again added back to the images set and resizing, cropping and other actions are performed on them.

Displaying the images

The images are displayed and run in the loop as per the code written. Pictures are augmented, cropped and then run in loop for the visualization purposes. Through the implementation of these procedures, you guarantee that the gathered face photos are properly pre-processed, standardized, and enhanced in order to train a GAN model. Before training the GAN, you can examine the pre-processed images visually with the display function.

Dataset

The facial recognition models that can be trained and tested on this dataset work particularly well when it comes to identifying features of the face, like brown hair, wearing glasses, or smiling. Many different people in a variety of poses, cluttered backgrounds, and rich annotations are all present in the images.

Evaluation Metrics

How are GANs operated?

GANs compete between two neural networks to learn the probability distribution of a dataset.

New data instances are created by one neural network, known as the Generator, and are then assessed for authenticity by another, known as the Discriminator. Specifically, the discriminator determines whether or not each instance of data it examines is an authentic part of the training dataset.

The generator feeds the discriminator with fresh, artificial, or fake images in the interim. The intention behind this is that despite being fake, they too will be accepted as genuine. Utilizing transposed convolution, the inverse of convolution, the fictitious image is produced from a 100-dimensional noise (uniform distribution between -1.0 and 1.0).

The generator's objective is to produce acceptable images so that a person can lie without being discovered. The discriminator's task is to detect images that originate from the generator as fraudulent.

The actions a GAN performs are as follows:

After receiving a random number, the generator outputs an image.

Together with a stream of images obtained from the real, ground-truth dataset, this generated image is fed into the discriminator.

The discriminator accepts input in the form of both real and phony images, and outputs probabilities—a number between 0 and 1 that indicates phony and 1 that indicates a prediction of authenticity.

Thus, you possess a dual feedback loop:

-We know the ground truth of the images, so the discriminator is in a feedback loop with it.

-The generator is currently receiving feedback.

Architecture

- Generator: The generator network's purpose is to create realistic human faces out of random noise, or latent vectors. It is made up of convolutional and dense layers that gradually upsample the input noise to produce images with the desired dimensions. LeakyReLU activations and Conv2DTranspose layers for feature map upsampling are included in the architecture.
- Discriminator: The discriminator network assesses how realistically the produced faces are rendered. Dropout is used for regularization, and convolutional layers with LeakyReLU activations make up the structure. Depending on whether the input image is generated or real, the final layer generates a binary classification output.

Training Parameters

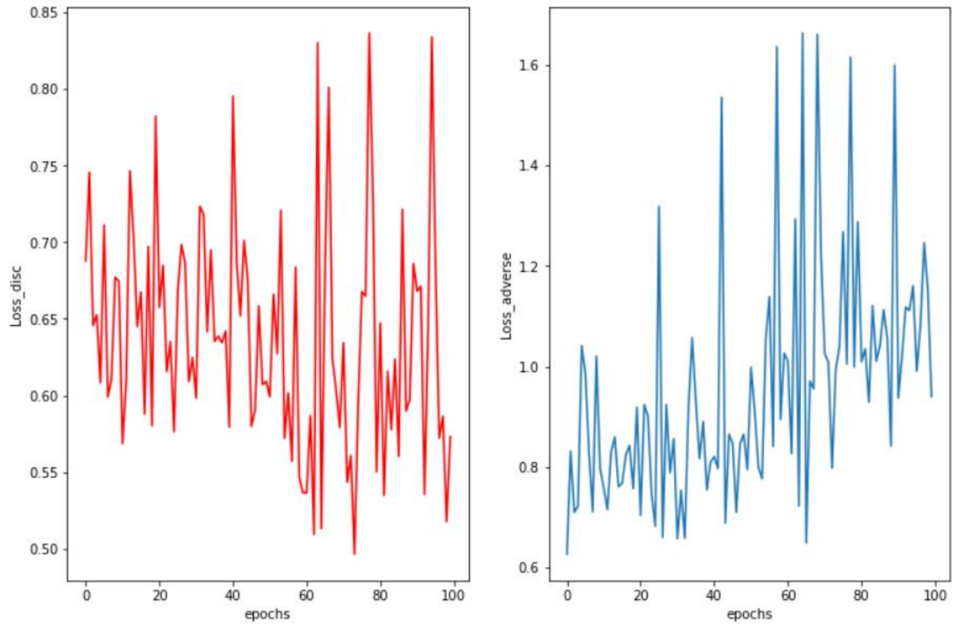
The architecture begins with a densely connected layer and proceeds with a series of transposed convolutional layers, using a common GAN pattern to combine the feature maps until the desired output resolution is achieved. To ensure that the generated image's pixel values fall within the range [-1, 1], the final layer employs the tanh activation function, as is common with GANs for image generation tasks.

The generator and discriminator are trained simultaneously in a competitive manner. The generator aims to create more realistic images to fool the discriminator.

Challenges

- The tendency of GANs to overfit their training data over time is known as overfitting. As a result, a lot of images with little variation may be produced. Frequently, regularization, data augmentation, and gradient penalty are employed by GANs to tackle this problem.
- During the initial phases of training, convergence of GANs can pose a considerable difficulty. Sometimes, in order to get convincing images, the models may need to be trained for several hundred thousand iterations.
- Discriminator Complexities: The discriminator may become proficient at producing false images as a result of its training to distinguish between real and fake images. In this instance, the generator ages out of use, resulting in the Discriminator Dominance phenomenon. Spectral normalization, Wasserstein loss, and learning rate manipulation are a few tactics that can be used to combat this.
- Interpretation: Because of their complexity and non-linearity, GANs are frequently regarded as "black-box" models despite their success. Although research is moving toward greater interpretability, there is still much to learn about the underlying mechanisms of GANs.

In conclusion, GAN’s - generative models in general can be highly entertaining, as well as confusing. They signify an additional advancement towards a society where artificial intelligence plays a bigger role. A few of the many applications for GANs include generating examples for image datasets, producing realistic photographs, translating images to text, translating images to images, translating semantic images to photos, generating new human poses, aging faces, video prediction, producing 3D objects, and many more scenarios.



The model is successfully trained and the two values acquired here for 100 epochs are d_loss: 0.6092, a_loss: 0.7952.

d_loss: 0.5731, a_loss: 0.9400