**Painting Style Transfer**

**using GAN**

**COMPUTER VISION PROJECT**

**CSE, 6TH SEMESTER**

**Group Members:**

|  |  |
| --- | --- |
| Axa Gemini Lakra | B121010 |
| Shruti Swarupa Dhar | B121056 |
| Abhikalp Srivastava | B521002 |

**Faculty:**

**Mr. Sarthak Padhi**



**Department of Computer Science and Engineering**

**International Institute of Information Technology**

**Bhubaneswar – 751003**

Date: 19/04/2024

Contents

[1. Project Title 3](#_Toc163943503)

[2. Project Description 3](#_Toc163943504)

[Requirement of the task: 3](#_Toc163943505)

[Objective of the project: 3](#_Toc163943506)

[Methodology: 3](#_Toc163943507)

[3. Data 4](#_Toc163943508)

[4. Models used 5](#_Toc163943509)

[5. Experimentation environment 6](#_Toc163943510)

[Hardware 6](#_Toc163943511)

[Software 6](#_Toc163943512)

[Parameter Setup 6](#_Toc163943513)

[1.Learning Rate 6](#_Toc163943514)

[2.Lambda Cycle. 6](#_Toc163943515)

[3.Optimizer Parameters 6](#_Toc163943516)

[Performance Metrics 6](#_Toc163943517)

[6. Analysis of Results 7](#_Toc163943518)

[7. Conclusion 8](#_Toc163943519)

[References 9](#_Toc163943520)

# Project Title

Painting Style Transfer using GAN

# Project Description

## Requirement of the task:

The project needs to utilise GAN for style transfer to demonstrate variation in painting styles on photographs.

## Objective of the project:

The primary goal of the project is to develop an automated system capable of imposing the variations in style of paintings on a given photograph. The project aims to leverage the capabilities of Generative Adversarial Networks (GANs) for painting style transfer, exploring their potential to generate realistic and stylistically consistent images.

## Methodology:

To solve the problem of image style transfer using CycleGAN, a systematic methodology is employed. Initially, paired images from two different domains, such as photos and Monet-style paintings, are collected. These images undergo preprocessing, including resizing and normalization, to prepare them for training. Next, suitable architectures for the generator and discriminator networks are selected. The CycleGAN model is then trained on the paired image dataset using predefined loss functions, with hyperparameter tuning conducted to optimize model performance. Validation and evaluation are carried out to assess the model's effectiveness using quantitative metrics. Subsequent fine-tuning and iteration based on validation results help enhance the model's performance. Once trained, the model is deployed for real-world applications and tested on unseen data to ensure its effectiveness. This systematic approach ensures the successful implementation of image style transfer using CycleGAN, leading to high-quality style transfer results.

# Data

The dataset for this project is "monet2photo" dataset which is available on Kaggle and is a collection of images used for the task of style transfer or image-to-image translation. The dataset broadly consists of two types of images:

- Monet-style paintings: This category includes images of artworks created by the famous French painter Claude Monet. These paintings are characterized by their distinctive style, featuring vibrant colors, broad brushstrokes, and an emphasis on light and atmosphere.

- Photographs: This category includes real-world photographs captured using a camera. These photographs depict various scenes, objects, and landscapes in a more realistic and detailed manner compared to the Monet-style paintings.

The dataset can be accessed from this [link](https://www.kaggle.com/datasets/balraj98/monet2photo) on Kaggle.

Basically, the "monet2photo" dataset consists of two classes: Monet-style paintings and photographs. Each class represents a distinct style or type of image within the dataset.

The "monet2photo" dataset is sourced from the "GANimation" project by Jun-Yan Zhu. This dataset was created for the purpose of training and evaluating image-to-image translation models, particularly those focused on transforming photographs into Monet-style paintings or vice versa. This dataset was obtained from UC Berkeley's official directory of CycleGAN Datasets.

The dataset consists of 1193 Monet Paintings & 7038 Natural Photos with each split into train and test subsets.

# Models used

The project utilizes CycleGAN as the only major model which can be further classified into the following components:

1. Monet Generator:

- This is the generator network responsible for transforming images from the photo domain to the Monet painting domain.

2. Photo Generator:

- This generator network performs the reverse transformation, converting Monet-style paintings back into realistic photos.

3. Monet Discriminator:

- The discriminator network for the Monet domain, which aims to distinguish between real Monet paintings and fake ones generated by the Monet generator.

4. Photo Discriminator:

- It is similar to the Monet discriminator; this network discriminates between real photos and fake ones generated by the photo generator.

The generator networks are responsible for image translation between the two domains, while the discriminator networks provide adversarial feedback to guide the training process. The goal is to train the generators to produce realistic images that can fool the discriminators, leading to high-quality style transfer results.

# Experimentation environment

## Hardware

The model used Kaggle Notebook Disk Space-73.1 GB. The maximum epoch of 30 and accelerator (GPU) 100.

## Software

• Operating System: Windows

• Programming Languages: Python 3.11

• Deep Learning Libraries: TensorFlow 2.6, Keras

• Data Manipulation Libraries: NumPy, Pandas

## Parameter Setup

## 1.Learning Rate: Adam optimizers with a learning rate of 2e-4 are used for both the generator and discriminator networks.

## 2.Lambda Cycle: The lambda\_cycle hyperparameter is set to 10.

## 3.Optimizer Parameters: The beta1 parameter for Adam optimizer is set to 0.5.

## Performance Metrics

1.Generator Loss: The generator\_loss function computes the adversarial loss for the generator networks.

2.Discriminator Loss: The discriminator\_loss function calculates the adversarial loss for the discriminator networks.

3.Cycle-Consistency Loss: The calc\_cycle\_loss function computes the cycle-consistency loss between original and reconstructed images.

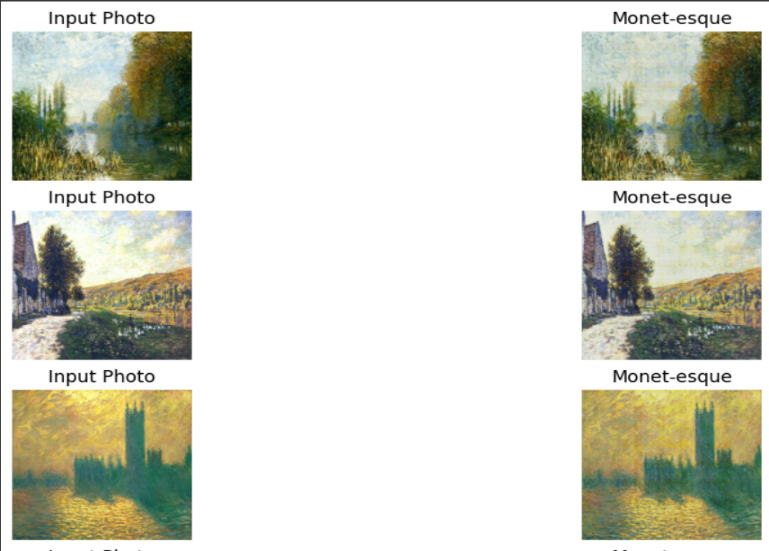
4.Identity Loss: The identity\_loss function calculates the identity loss to preserve the content of input images.

5.PSNR (Peak Signal-to-Noise Ratio): PSNR is computed for generated images to measure their quality compared to ground truth images.

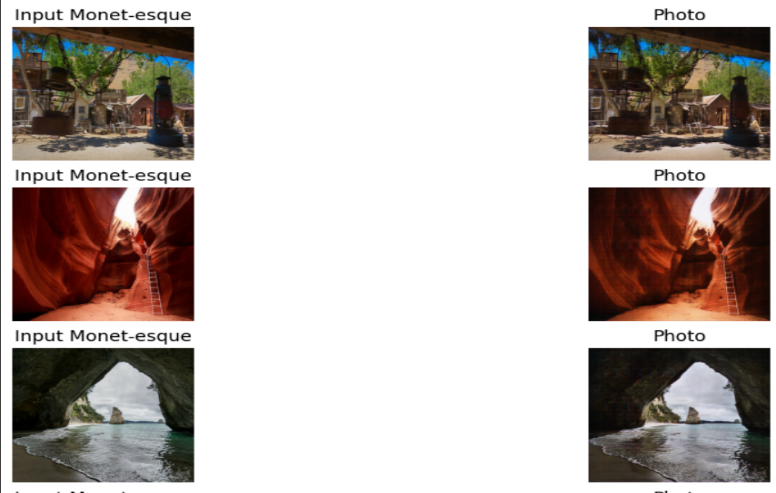
6.SSIM (Structural Similarity Index Measure): SSIM is calculated to quantify the similarity between generated and ground truth images in terms of luminance, contrast, and structure.

# Analysis of Results

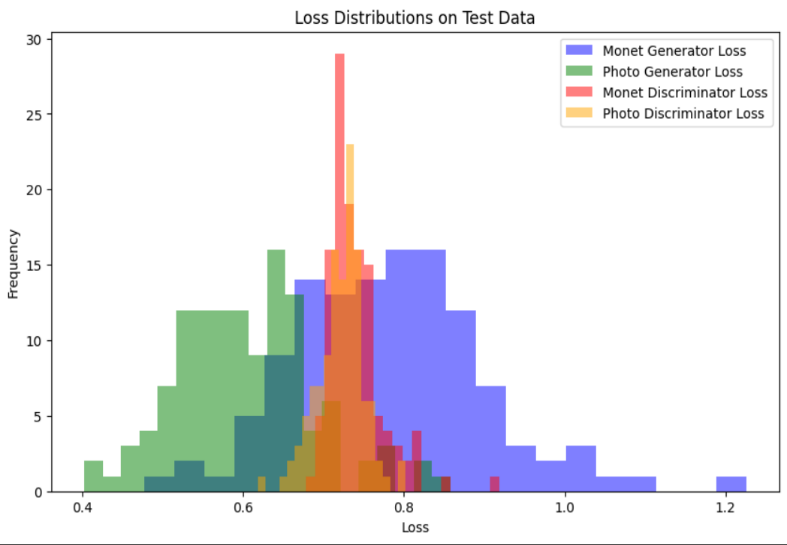
The results obtained from the trained CycleGAN model is showcased as below where sample images from the test dataset is displayed alongside their corresponding translated images. This allows viewers to visually inspect the quality of style transfer achieved by the model. This contains the input photograph which is converted into painting using style transfer.



And below is the conversion of painting into original photograph.



Further the generator and discriminator losses can be displayed as below:



# Conclusion

The task addressed by the project is image style transfer using CycleGAN. Image style transfer aims to transform images from one domain (e.g., photographs) to another domain (e.g., Monet-style paintings) while preserving the content of the original images.

The data used for training and evaluation consists of paired images from two domains: photos and Monet-style paintings and the implemented CycleGAN models include:

- Generator Networks: Separate generator networks for transforming images from the photo domain to the Monet domain and vice versa.

- Discriminator Networks: Discriminator networks for both the photo and Monet domains, which distinguish between real and generated images in their respective domains.

The results obtained from the implementation and evaluation of the CycleGAN model include:

- Visual samples showing original photos alongside their corresponding Monet-style paintings generated by the trained CycleGAN model and vice-versa.

- Quantitative metrics such as generator and discriminator losses providing insights into the quality of style transfer achieved by the model.

Potential areas for improvement in the implementation and results could include:

- Experimenting with different values for hyperparameters such as learning rates, lambda\_cycle, and optimizer parameters to improve model convergence and performance.

- Exploring alternative network architectures or adding additional layers/modules to the generator or discriminator networks to enhance the quality of generated images.

- Utilizing data augmentation techniques to increase the diversity of the training dataset and improve the generalization capability of the model.

- Incorporating advanced loss functions or regularization techniques to further refine the training process and enhance the quality of generated images.

By addressing these areas of improvement, the performance and capabilities of the CycleGAN model can be enhanced, leading to more realistic and high-quality style transfer results.

# References

1. Isola, Phillip, et al. "Image-to-image translation with conditional adversarial networks." Proceedings of the IEEE conference on computer vision and pattern recognition. 2017.

2. Zhu, Jun-Yan, et al. "Unpaired image-to-image translation using cycle-consistent adversarial networks." Proceedings of the IEEE international conference on computer vision. 2017.

3. Goodfellow, Ian, et al. "Generative adversarial nets." Advances in neural information processing systems. 2014.

4. Ronneberger, Olaf, Philipp Fischer, and Thomas Brox. "U-net: Convolutional networks for biomedical image segmentation." International Conference on Medical image computing and computer-assisted intervention. Springer, Cham, 2015.

5. Johnson, Justin, Alexandre Alahi, and Li Fei-Fei. "Perceptual losses for real-time style transfer and super-resolution." European Conference on Computer Vision. Springer, Cham, 2016.

6. Gatys, Leon A., Alexander S. Ecker, and Matthias Bethge. "Image style transfer using convolutional neural networks." Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.

7. Gatys, Leon A., et al. "A neural algorithm of artistic style." arXiv preprint arXiv:1508.06576 (2015).

8. He, Kaiming, et al. "Deep residual learning for image recognition." Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.

9. Arjovsky, Martin, Soumith Chintala, and Léon Bottou. "Wasserstein gan." arXiv preprint arXiv:1701.07875 (2017).

10. Mirza, Mehdi, and Simon Osindero. "Conditional generative adversarial nets." arXiv preprint arXiv:1411.1784 (2014).