Usha Nandhini M

Final Project

PROJECT TITLE

Grayscale Image Transformer

To convert grayscale images to colorized versions using GAN

AGENDA

1. PROBLEM STATEMENT
2. PROJECT OVERVIEW
3. END USERS
4. SOLUTION AND VALUE PROPOSITION
5. WOW FACTOR OF THE SOLUTION
6. MODELLING
7. RESULTS

PROBLEM STATEMENT

Despite the vast amount of historical and monochrome imagery available, there is a significant challenge in revitalizing these images to **convey their original vibrancy and context accurately. Manual colorization processes** are **time-consuming** and often result in subjective outcomes. Thus, there is a need for **an efficient and accurate automated tool for colorizing black and white images** while preserving their authenticity and historical significance.

PROJECT OVERVIEW

The project aims to develop an automated **Grayscale image transformer tool utilizing WGAN** (Wasserstein Generative Adversarial Network) . Leveraging state-of-the-art techniques, the tool will provide users with a seamless and efficient way to **transform monochrome images into vibrant, colored counterparts** while preserving their original context and authenticity.

The tool will cater to diverse applications such as **historical image restoration, artistic enhancement, and multimedia content creation**.

Through rigorous testing, user feedback incorporation, and continuous improvement, the project seeks to deliver a **reliable and versatile** solution **for colorizing black and white images**, contributing to **advancements** in **image processing technology** and **historical preservation efforts.**

WHO ARE THE END USERS?

1. **Historians and Researchers:**
   Utilizing colorization for deeper insights into historical imagery.

2. **Artists and Designers:**
   Enhancing creative projects with colorized sketches and illustrations.

3. **Photographers and Filmmakers:**
   Restoring and adding visual interest to archival footage and photographs.

4. **General Users:**
   Preserving personal memories and sharing black and white images in a vibrant format.

SOLUTION AND ITS VALUE PROPOSITION

The solution leverages a **Wasserstein Generative Adversarial Network (WGAN)** to perform **grayscale image transformation.** It is trained on a dataset of grayscale images paired with their corresponding colorized versions. The model learns to generate realistic colorizations **by minimizing the Wasserstein distance** between the distribution of generated colorized images and the distribution of real colour images. This approach ensures that the generated colorizations **maintain the original context and authenticity** of the grayscale input images.

Value Proposition: Time-saving, Accuracy, Versatility, Enhancing Creativity, Accessibility

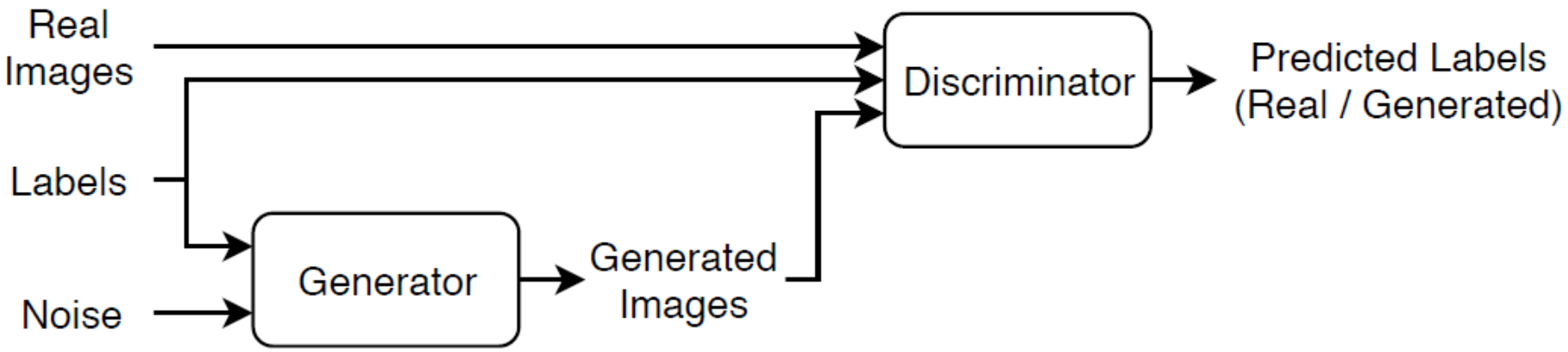
THE WOW IN YOUR SOLUTION

The colorization tool's "wow" factor, driven by **Wasserstein Generative Adversarial Network (WGAN),** lies in its seamless transformation of black and white images into vibrant, lifelike representations.

Unlike traditional GANs, **WGAN's stability** is ensured by using the **Wasserstein distance** as the **loss function**, allowing for smoother optimization. Additionally, WGAN's 1-Lipschitz constraint on the critic **enhances its ability to discern subtle differences**, resulting in **accurate and detailed colorizations** that preserve the original image's essence and detail.

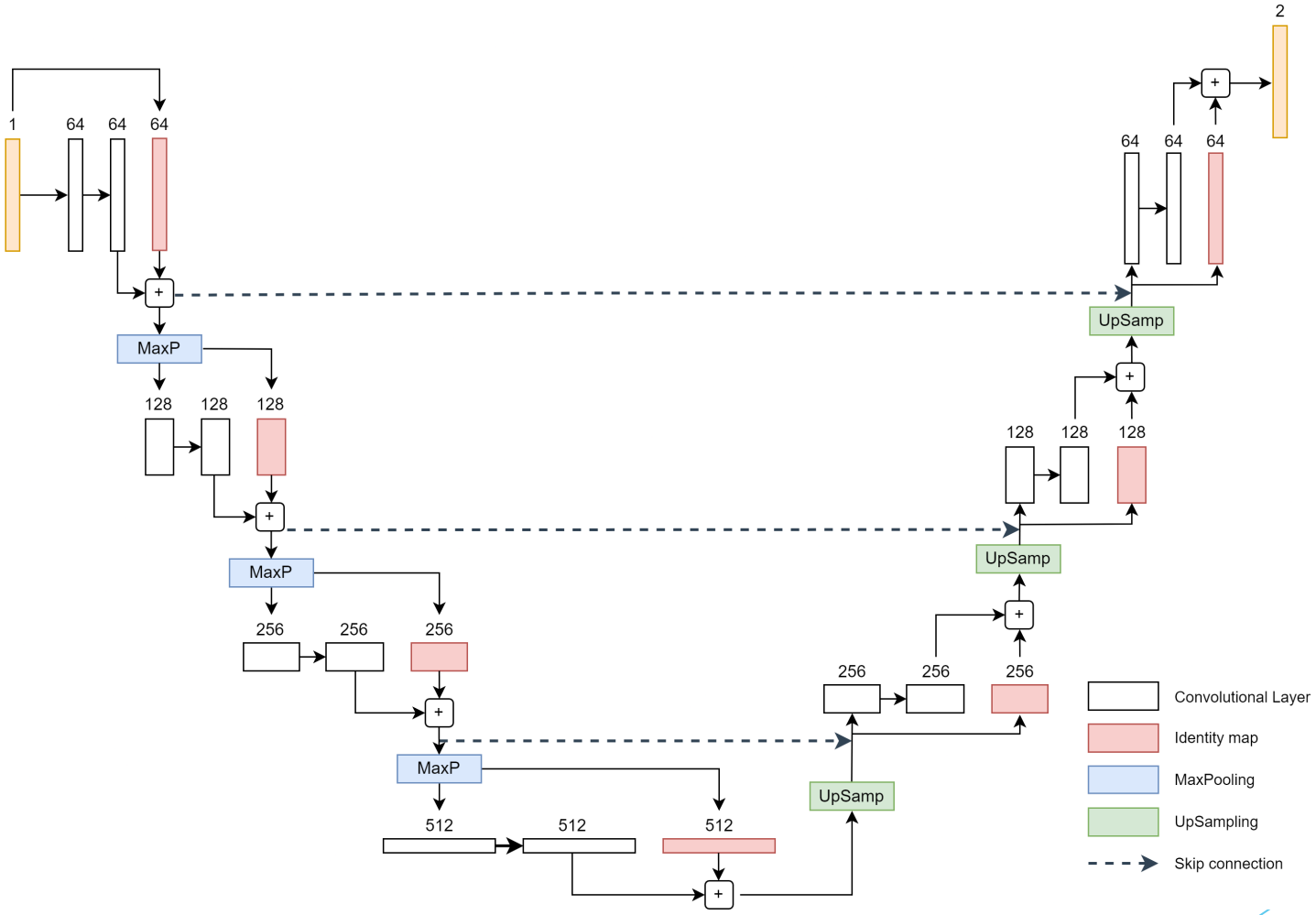
MODELLING

CONDITIONAL GENERATIVE ADVERSIAL NETWORK



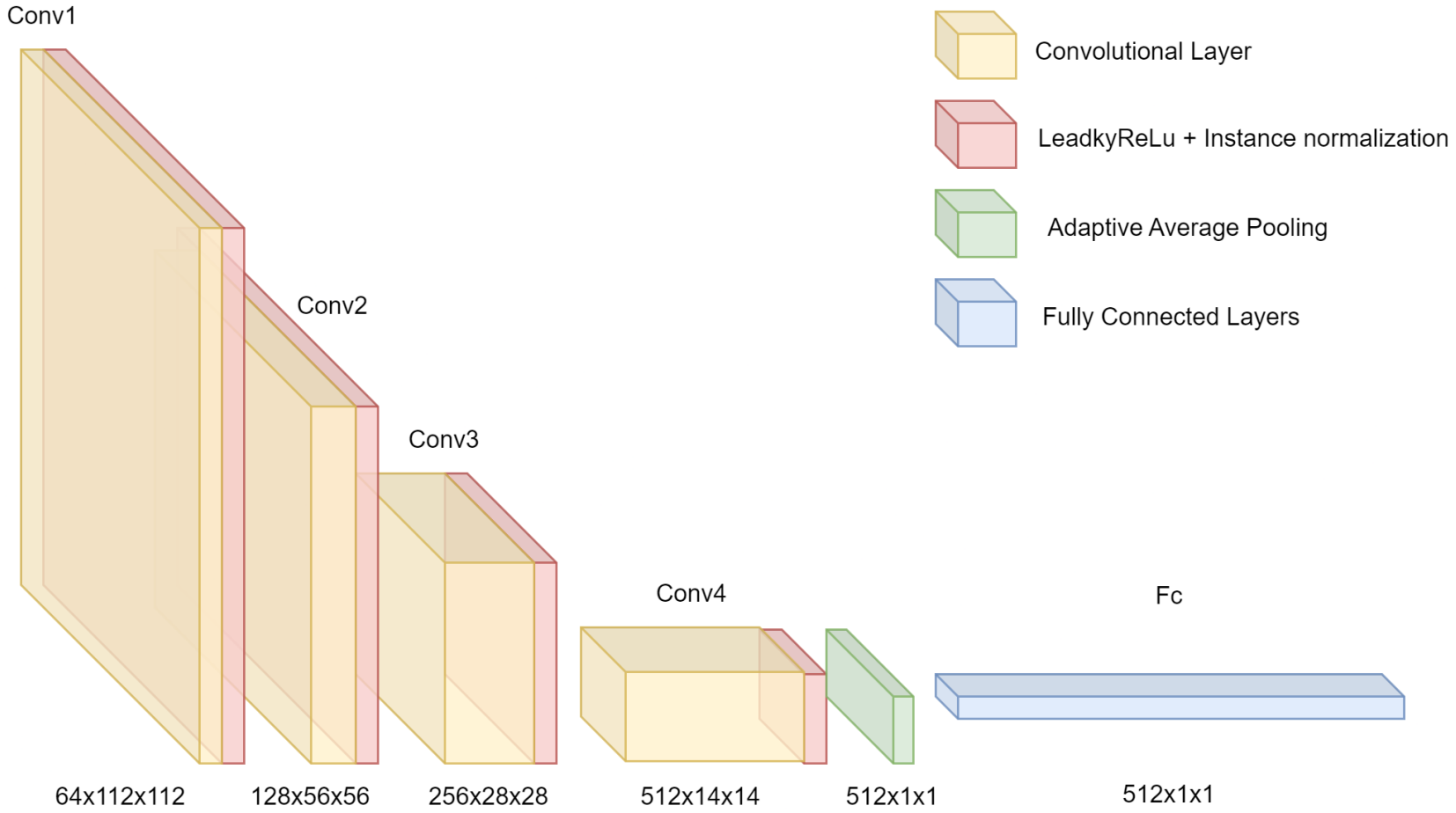
MODELLING (cont)

Generator (ResU-NET)



MODELLING (cont)

Discriminator (Critic)



RESULTS

1. Human Scoring:
    1. Human scorers label each image as real or fake.
    2. Utilizes platforms like Amazon Mechanical Turk for scoring.
    3. Provides qualitative insights into image realism.

2. Inception Score (IS):
    1. Uses a pre-trained Inception model to classify generated images.
    2. Measures both image quality and diversity.
    3. Higher entropy indicates better performance.

3. Frechet Inception Distance (FID):
    1. Compares real vs. generated image distributions.
    2. Lower values indicate better quality and diversity.
    3. More robust to noise and mode collapse.

RESULTS (CONT)

1. Inception score

```
# Calculate the Inception Score
mean_real, std_real = is_model.calculate_is(all_real)
print("Inception Score of real images: mean: {:.4f}, std: {:.
4f}".format(mean_real, std_real))
mean_is, std_is = is_model.calculate_is(all_preds)
print("Inception Score of fake images: mean: {:.4f}, std: {:.
4f}".format(mean_is, std_is))
```

```
Inception Score of real images: mean: 4.1913, std: 1.7591
Inception Score of fake images: mean: 4.2828, std: 1.9007
```

RESULTS (CONT)

2. Frechet Inception Distance (FID):

In [106]:

```
# Initialize the FID class
device = "cuda" # or "cpu" if you don't have a GPU
fid_calculator = FID(device)

# Calculate the FID
fid_value = fid_calculator.calculate_fid(all_real, all_preds)
print("FID: {:.4f}".format(fid_value))
```

```
FID: 36.9385
```

CONCLUSION

In conclusion, **GRAYSCALE IMAGE TRANSFORMER** is a powerful model for specific applications such as colorization.
Using Wasserstein GAN (WGAN) and a U-Net architecture based on residual blocks are two ways to improve the performance of the tool for colorization.

WGANs use **the Wasserstein distance metric** to train the generator and discriminator which can help stabilize the training process and **produce more realistic results**. A **U-Net architecture** which is **well-suited for image segmentation tasks** combined with Residual blocks allows the network to learn **fine details of the input image**, which can be particularly **useful for colorization tasks**.

This can help **improve the stability** and the ability **to learn fine details** of the input image, producing more **realistic results.**