AI Style Transfer Competition

Agenda Document

Introduction: The Power of Style Transfer

Style transfer is a powerful technique that blends the content of one image with the style of another, enabling the creation of visually compelling and artistically expressive results. It began with neural networks and has since evolved with advances like GANs and diffusion models, making it possible to generate and stylize images with remarkable control and creativity.

Today, by first generating meaningful base images using diffusion models and then applying personalized styles through custom-trained GANs, we unlock new possibilities in visual storytelling, awareness campaigns, and digital art. This competition aims to give students hands-on experience with these cutting-edge tools to explore how AI can be used not just for automation, but for expression, impact, and design.

Agenda Statement

Welcome to the **AI Style Transfer Challenge**, an exciting blend of **visual storytelling, creativity, and deep learning**! This competition will test your ability to not only implement advanced AI models but to also interpret abstract ideas, express them visually, and stylize them with generative art techniques.

The competition consists of **three progressive stages**, each requiring both technical prowess and creative problem-solving. Participants will work in teams using only **cloud GPU environments** such as Google Colab or Kaggle or their own Laptops or Remote GPU Access.

Problem Overview

You will be challenged to generate **visually stylized representations** of abstract phrases, many of which hint at **social, global, or psychological issues**. Your journey will take you from **conceptual understanding → image generation → artistic transformation**.

Phase 1 - Visual Interpretation Using Diffusion Models

Teams will receive a set of **unique, cryptic phrases** inspired by **social problems, world health, or environmental concerns**.

Your task:

- Interpret the **hidden** meaning of each phrase.
- Use a pre-trained **Diffusion Model** (or a similar text-to-image model) to generate an image that visualizes the concept.
- You can generate **1** or **more images** depending on how many phrases you choose to tackle.
- Minimum **1 Phrase** is to be Selected.
- Maximum **3 Phrases** can be Selected.
- These images will serve as your **Base Canvas** for the final **stylization phase**.

Phase 2 - Train a GAN Architecture from Scratch

You will be given **2 small style datasets**, each containing images with distinct visual themes (e.g., cartoon, painting, sketch).

Your task:

- Choose any one of the **style dataset** provided.
- Perform Necessary **Data Preparation** Steps.
- Implement a GAN model or a **Derivative** of GAN Model from **scratch** using PyTorch or TensorFlow.
- Train your GAN to learn the style of your chosen dataset.
- Test your Architecture If it is able to Learn the Style from the Dataset Provided.

Phase 3 - Style Transfer: Apply Your GAN to Stylize Your Diffusion Output

Now comes the fusion of creativity and training.

Your task:

- Take the **Canvas Image(s)** generated in **Phase 1**.
- Use your trained GAN Architecture to apply the **style** you learned from the dataset onto the **canvas image**.
- This results in a **Style-Transferred Interpretation** of the original **abstract phrase**
- Save the Style-Transferred Images.

What You Will Submit

Phase 1: Metrics to be Submitted

- Make a Folder in your Competition Directory Named “Phase - 1 Submission”
- Your Selected Phrases from the Provided Pool
- Customized Prompt for Each Phrase
- Name of the Diffusion Model Selected
- CLIP Score (Prompt-Image Semantic Similarity)
- BLIP Caption Similarity Score (Prompt vs Generated Caption)
- Generated Diffusion Image
- Add a Text File Containing All the Textual Information “BLIP”,”CLIP” etc

Phase 2 and 3: Metrics to be Submitted

- Make a Folder in your Competition Directory Named “Phase - 2 Submission”
- Add **Screen Shots** of the Metrics
- **Fréchet Inception Distance** (FID)
- **Inception Score** (IS)
- **Style Loss** (Gram Matrix Distance)
- **Style Classification Accuracy**
- **LPIPS (Learned Perceptual Image Patch Similarity)**
- Your Custom GAN Architecture
- Trained Model Weights
- A 2 - Page (Can be 1 pager as well) Document Highlighting :
    - Which Diffusion Model you Used and Why? Quantization ?
    - Explain your Custom GAN Architecture. Any innovation ? Optimization ? Mixed Precision ?
    - Is it possible to Train your Architecture on Multiple Style Datasets at the same time (e.g a dataset contains Multiple Images with different style)? Why or Why not ?
    - Is there a better Approach for style transfer than GANs and why ?
    - Where did you Observe in this competition the Limitation of GANs as compared to Diffusion ?
- Finally, Make another folder Named “Phase - 3 Submission”.
- Submit your Style Transferred Image Here
- An Inference JupyterNotebook where we just provide an image and Style is being mapped onto it.
- You are also supposed to provide with detail report of your code and its working.

Happy Coding :)