Artify: Video 2 Video Neural Style Transfer

Adwait Agashe | Abhishek Goyal | Muskaan Beriwal

Table of Contents

Introduction

In our project, we aim to build upon Neural Style Transfer (NST) by extending its capabilities from static images to dynamic video content. Through our attempt, we have tried to adapt this technique to work with video data, allowing for the extraction of styles from one video and applying them to another.
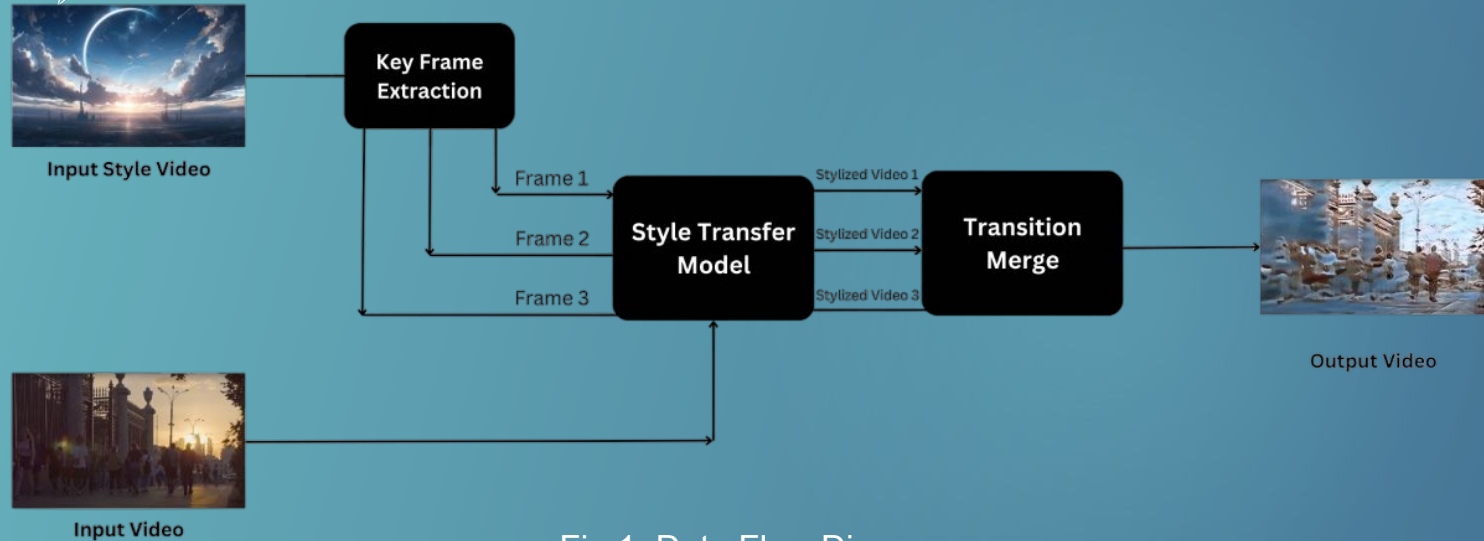

Fig 1: Data Flow Diagram

Goals and Objectives

1- Develop a foundational model:

Creating a foundational model capable of extracting the artistic style from an input image and then applying that style to an output image or video.

2- Implement video style transfer algorithms:

Using various algorithms, we extract a style image from a video input which serves as the style input to be applied on another video, while preserving temporal coherence & video quality.

3- Efficiency & Performance Optimization:

The optimization process aims to ensure that style transfer can be performed quickly, allowing users to experience dynamic style changes without significant delay.

4- Quality Evaluation:

We use comprehensive metrics and assessment criteria to evaluate the quality of the generated videos, including both automated metrics & manual assessments.

Approach

Creation of a Stylized Image Model

-> Develop a foundational model which accepts an input image and a style image, and generates a stylized image as its output.

-> Implemented the following 3 models: Gatys et al., Segmentation - based, and Arbitrary Image Stylization - and optimized model performance.

-> We expand each model's functionality to accommodate video inputs. This enhancement enables us to generate a coherent and synthesized stylized video.

-> Experimented with the same 3 models on different loss functions to minimize flickering, ensuring a consistent stylized output.

Extension to Video Inputs

Video Style Extraction and Optimization

-> We further our model by using video input for the extraction of style in an image.

-> We extracted key-frames from the input style video by segmenting the video and picking a frame from each which serve as the style inputs. These are applied to the target video in transition.

Style Transfer by Gatys et al

Motivation

Establish a baseline model and implement basic neural style transfer for input image and style image

Model Structure

Minimise the Style Loss from style image and Content loss form the Original Image to Reconstruct the output image.

Extension to the Original Model

We extended the Style Transfer model to work from a single input image to an input video
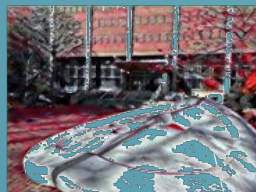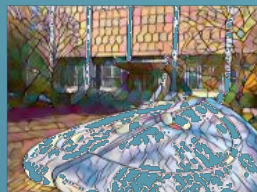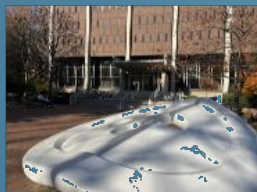
Drawbacks:

Slow Style Transfer for Videos

Reconstructing video frame by frame takes a lot of time

Background Flickering

Inconsistent styles in still objects and lots of changes in color across frames of the video

Model Outputs

Model Outputs

Output Video

Style Image

Input Video

Segmentation of Foreground and Background

Motivation

Try to get consistency in the foreground and background of the image and try to reduce the flickering in videos.

Structure

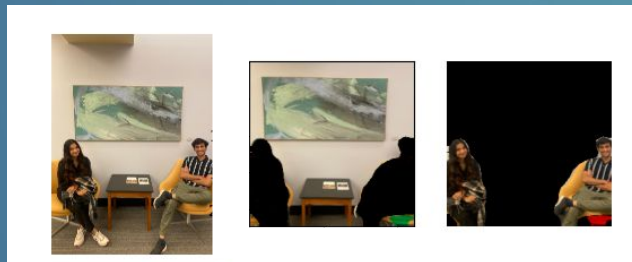Segment out Foreground and Background

Apply styles independently.

Merge to make the final stylized image.

Approaches Implemented

1. Segmenting foreground and background using strict edges
2. Segmenting foreground and background using blurring at edges.
3. Blurring instead of segmenting out.
4. Resizing foreground and background to size of image

Model Outputs
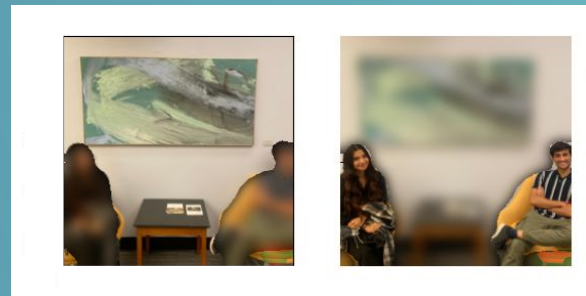
Segment Foreground and Background

Input Image   Background segmented   Foreground segmented
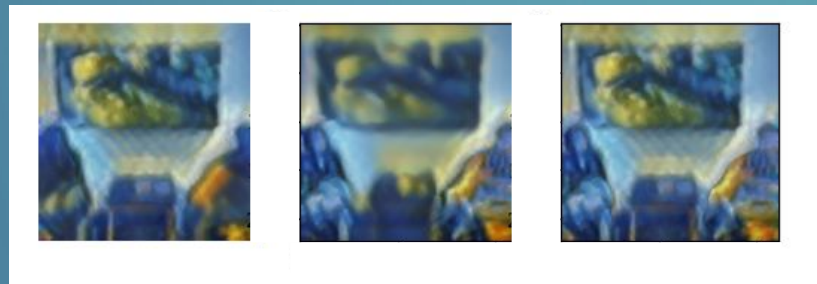
Blurring both images

Background blurred   Foreground blurred

Final Output and Merging

Background stylized   Foreground stylized   Combined Stylized Output

Model Outputs

Arbitrary Image Stylization

Motivation

Overcome the drawbacks of the baseline model and try to improve upon the time taken for each frame to be stylized

Model Structure

Using Style Prediction Network that supplies set of normalization constants for Style Transfer Network which minimizes content and style loss to generate stylized images
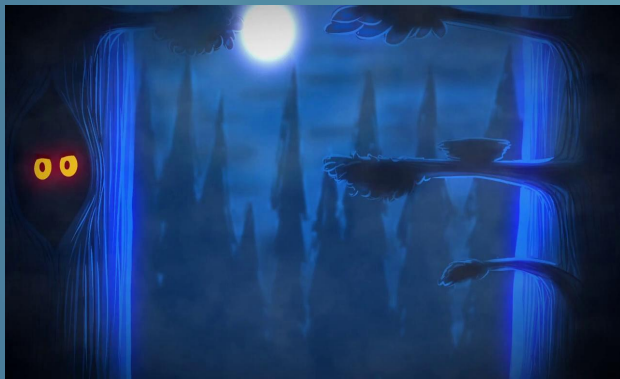
Extension to the Original Model

Implemented Transition Style Strategy by using overlaying style transfer with proportional mix ratio. The intensity of style 1 is gradually diminished while amplifying the influence of the subsequent keyframes style

Applying Multiple Styles

Key Frames from Video

Video to Video Style Transfer Outputs

Video to Video Style Transfer Outputs

Future Work

1. Incorporating Optical Flow Information to improve temporal coherence

2. Inpainting in segmented Foreground and Background images for more seamless results

3. Extracting Keyframes by quantitatively identifying style change and synchronizing style changes

4. Identifying layer in NN with maximum variation across video frames and normalizing it to reduce video flickering

Thank You