

REPORT ON THE DESIGN AND IMPLEMENTATION OF **THE VIDEO CONVERSION TOOL**

PROBLEM STATEMENT

To develop a prototype of a video conversion process that can convert SD videos to HD videos using diffusion (image to image, Inpainting) type models, while preserving the context of the video and being efficient.

INTRODUCTION

The video conversion tool is designed to upscale and enhance video quality by converting Standard Definition (SD) videos to High Definition (HD). Specifically, videos of quality 640x480 px to 1280 x 720 px. This report outlines the workflow, processes, and models utilized into the code implementation. The tool utilizes diffusion models, to perform inpainting and upscale the video frames effectively.

WORKFLOW OVERVIEW

The video conversion process is divided into several key stages:

1. Frame Extraction
2. Frame Resizing and Padding
3. Frame Inpainting
4. Video Reconstruction

Frame Extraction

The initial step in the video conversion process is to extract individual frames from the input SD video. This is achieved using video processing tools that read the video file and capture each frame sequentially. The extracted frames are then stored for further processing.

- Tools Used: OpenCV
- Output: A list of frames extracted from the video

Frame Resizing and Padding

Given that the extracted frames might not match the target HD resolution, each frame undergoes a resizing and padding process. This step ensures that the aspect ratio is maintained while scaling the frames to the desired dimensions.

Process:

- Calculate the scaling factor based on the target width and height.
- Resize the frame using the scaling factor.
- Create a new frame with the target dimensions and pad the resized frame to fit centrally within it.

Frame Inpainting

The core enhancement process involves inpainting, which utilizes a diffusion model to improve and upscale the quality of the resized frames. The diffusion model, variant of the Stable Diffusion model designed for inpainting, fills in the missing or low-quality parts of the frames.

- Model Used: Stable Diffusion Inpainting Model
- Pre-trained Model: runwayml/stable-diffusion-inpainting
- Framework: Diffusers
- Hardware: Utilizes GPU acceleration

Process:

- Resize frames and corresponding masks to the input size required by the model.
- Use the diffusion model to perform inpainting on the frames.
- Resize the inpainted frames back to the original target dimensions.

Video Reconstruction

The final stage is to reconstruct the video from the inpainted frames. This involves compiling the processed frames back into a video file.

- Tools Used: OpenCV

CONCLUSION

This report provides an overview of the design and implementation of the video conversion tool, emphasizing the processes and models used to achieve high-quality video enhancement.

SHAURYA KUSHWAH

04119011921

AI-DS

USAR, GGSIPU