MIE 1517

Project Proposal

Team 8

Xingjian Li 1011563096
Zixuan Xia 1011232818
Yu Tao 1011679060
Hankun Wang 1011796344

Project: 2D Animation Resolution Enhancement (Multi-Frame Super-Resolution)

Goal:

The goal of this project is to enhance the resolution of 2D animation videos by using deep learning-based super-resolution techniques, specifically focusing on upscaling 360p animation to 1080p resolution. The project will apply state-of-the-art models such as **SRCNN** (Super-Resolution CNN), **ESRGAN** (Enhanced Super-Resolution Generative Adversarial Networks) or **Real-ESRGAN** to enhance the visual quality of low-resolution animation while preserving critical animation features like textures, lines, and colors.

Key objectives of this project:

- Extract frames from the provided 360p (low-resolution) animation video.
- Use the 1080p version as a high-resolution reference for training and evaluation.
- Train a deep learning model to upscale 360p frames to 1080p resolution.
- Apply the model to upscale the 360p frames to the desired quality level.
- Reconstruct the enhanced frames into a high-resolution 1080p animation video.
- Evaluate the results by comparing the upscaled animation against the original 1080p version using **PSNR** (Peak Signal-to-Noise Ratio) and **SSIM** (Structural Similarity Index).

Technology:

- **Software Development Environment:** Python, Jupyter Notebook / Google Colab
- **Deep Learning Frameworks:** PyTorch
- **Video Processing Tools:** OpenCV, FFmpeg
- **Deep Learning Models:** SRCNN, ESRGAN, Real-ESRGAN

Datasets:

- **The original video link:** https://www.bilibili.com/video/BV1ua4y1c7wk/?spm_id_from=333.337.search-card.all.click&vd_source=8a01a4be1976bdc67f39e330f0a3f22d
- **Google Drive:** https://drive.google.com/file/d/1LWGHU_pa0oK0XyAYE2y_dJccVKY4tihS/view?usp=sharing

  https://drive.google.com/file/d/1jXYTLWq3XWkfIaMnspL6fcXOWDDtEFdq/view?usp=sharing

- The 360p and 1080p videos will be paired up as inputs and outputs(targets), where both of them will be converted to frames(images) for model training, validation and testing.
- The length of the videos (360 & 1080 p) is 724 seconds, of which the first 70% (~508s) will be used as training data and the remaining 30% will be validation (15%, ~ 108s) and test sets (15%, ~108s).

Expected results:

By the end of this project, the following results are expected:

1. **Enhanced Resolution:** The 360p animation will be upscaled to 1080p resolution, retaining high levels of detail, sharpness, and clarity in the resulting frames.
2. **Visual Fidelity:** The upscaled 360p frames will closely match the high-resolution 1080p frames, preserving the original artistic style and ensuring that textures, lines, and animations are not distorted in the process.
3. **Quantitative Quality Improvement:** Using PSNR and SSIM, the upscaled video will show significant improvements in visual quality when compared to the original low-resolution video. We aim to achieve high PSNR/SSIM scores closer to the 1080p reference.
4. **Temporal Consistency:** The enhanced animation will maintain frame-to-frame consistency, avoiding flickering or temporal artifacts that can arise from frame-by-frame upscaling. The model will ensure smooth transitions and coherence across the entire video.
5. **Final High-Resolution Animation:** A polished, high-quality 1080p animation video will be produced that is visually sharper, clearer, and more suitable for professional or content production purposes.

Reasons (Why our project will be successful):

1. Proven AI Models – We use SRCNN, ESRGAN and/or Real-ESRGAN, which are well-tested for high-quality super-resolution.
2. High-Quality Data – With both 360p and 1080p versions, our model can learn effectively from real examples.
3. Reliable Evaluation – We measure improvements using PSNR, SSIM, and visual comparisons to ensure quality.
4. Efficient Workflow – Tools like FFmpeg and OpenCV streamline video processing, keeping our approach fast and effective.

Conclusion:

This project focuses on leveraging AI/DL to enhance 2D animation resolution, specifically SRCNN, ESRGAN and/or Real-ESRGAN, to upscale 360p animation to 1080p while preserving quality and details. By extracting frames, applying super-resolution models, and evaluating results with PSNR and SSIM, we ensure high visual fidelity.

The outcome will provide a practical solution for improving low-resolution 2D animations, with applications in upscaling and restoration. Additionally, this project lays the foundation for extending deep learning-based super-resolution techniques to other video types, such as 3D animation, opening opportunities for broader applications in video enhancement.