Advanced Video Enhancement using Real-ESRGAN: Mitigating Pixelation and Blur in Zoomed-In Footage

Swapnil Kannojia (1229614375), Ayushi Rajshekhar (1229974865)
Aprajita Bhowal (1225248297), Baibhav Phukan (1225408392)

Problem Statement

Numerous video recording devices grapple with a decline in video fidelity when employing zoom, often manifesting as pixelation and a loss of clarity. This project endeavors to create a system that addresses this issue by leveraging Real Enhanced Super-Resolution Generative Adversarial Networks (Real - ESRGAN). Real - ESRGAN presents a promising solution for producing high-resolution videos from lower-resolution inputs, leading to a substantial enhancement in the visual fidelity of zoomed-in footage.

Methodology

The project revolves around employing advanced computational methods to enhance video quality by addressing blurry segments. Initially, a meticulously curated dataset, DIV2K, featuring high-quality images ideal for image restoration, undergoes preprocessing, including data refinement and partitioning for distinct phases of experimentation. Following this, the development phase focuses on crafting an algorithmic framework capable of identifying regions within video frames exhibiting discernible blurriness. This algorithm, leveraging sophisticated computer vision techniques, determines thresholds indicative of significant blur, acting as the groundwork for subsequent enhancement processes.

The enhancement stage involves the utilization of Real - ESRGAN, an advanced image enhancement model, tailored for gradual refinement of identified blurry segments within the video frames. The optimization of Real - ESRGAN parameters ensures a progressive and coherent improvement, preserving visual quality throughout the video sequences. Integration efforts streamline the blur detection algorithm and Real - ESRGAN enhancement into a unified pipeline, fostering a seamless and efficient workflow for identifying and rectifying blurry sections within videos.

Thorough validation procedures entail extensive testing across a diverse spectrum of videos with varying levels of blur.Iterative refinement processes based on evaluation outcomes aim to fine-tune the system parameters and algorithms, with a focus on further enhancing its performance.

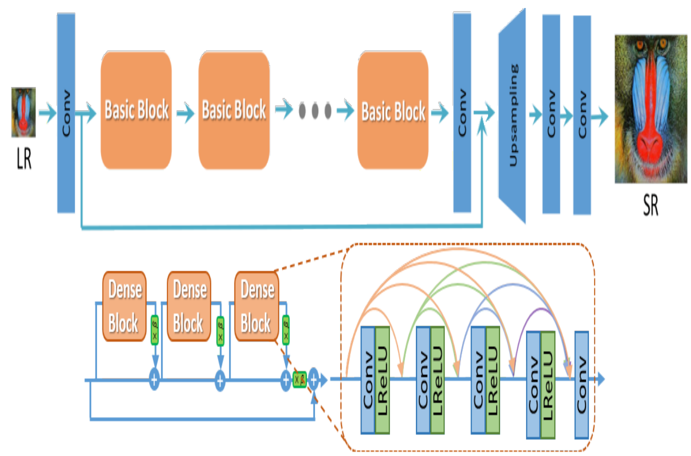


Figure 1: Real ESRGAN Model Architecture

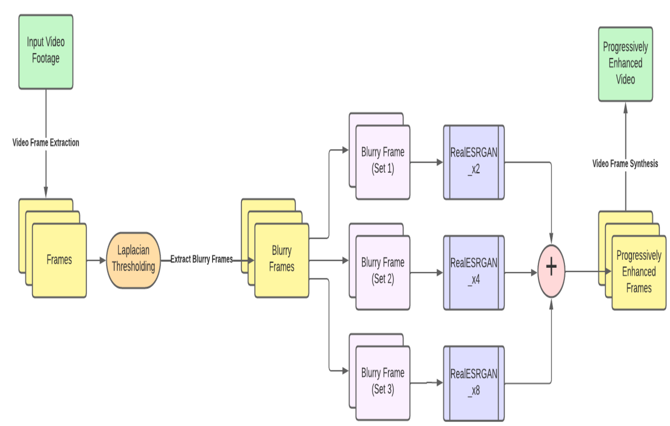


Figure 2: Project Pipeline

Results and Analysis:

The implemented architecture of RealESRGAN incorporates a selected model that proves instrumental in systematically enhancing video frames. This enhancement occurs on a frame-by-frame basis, utilizing a progressive approach. This progressive enhancement strategy ensures the adaptability of the model selection to cater to the diverse pixelation levels exhibited across different frames. The choice of the appropriate model is contingent upon the degree of pixelation present in each specific frame, guaranteeing optimal performance in the enhancement process.

To evaluate and discern the efficacy of various RealESRGAN models, including RealESRGANx2, RealESRGANx4, RealESRGANx8, RealESRGANx4plus, RealESRGANx4plusanime6B,and realesr-animevideov3, a comprehensive comparative study was conducted. The purpose of this study was to identify the model that consistently yielded superior enhancement results across a range of scenarios.

Among the evaluated models, RealESRGANx8 emerged as the most effective in terms of enhancing video frames. The choice of this particular model was substantiated by its superior performance in comparison to other variants. The relative comparison between the different RealESRGAN models is presented in the provided Table, named "Tablename," which encapsulates the key performance metrics and outcomes for each model. This systematic evaluation and model selection process contribute to the optimization of video frame enhancement, ensuring the delivery of high-quality results tailored to the specific characteristics of each frame.
This is the table:

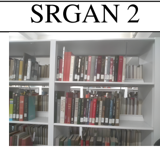
RealESRGAN_X4Plus	SRGAN 2
	
PSNR X4Plus: 33.47	PSNR SRGAN: 31.91

Figure 3: This image is a snippet from the video which compares the original input video and the enhanced output

Conclusions:

Addressing the challenge of diminishing video fidelity during zoom, our project employs Real Enhanced Super-Resolution Generative Adversarial Networks (Real-ESRGAN) to significantly elevate visual quality. The architecture seamlessly integrates RealESRGAN with a selected model, enabling progressive enhancement of video frames. This approach ensures the adaptive selection of the right model for each frame based on the degree of pixelation it exhibits.

A comprehensive comparative study across various ESRGAN models was conducted, revealing that RealESRGANx8 consistently yielded the best enhancement results. In essence, our project aspires to deliver an advanced and user-friendly solution for refining zoomed-in video content. This endeavor harmoniously blends state-of-the-art technology with a meticulous methodology, embodying our commitment to enhancing the overall video viewing experience.

References

- Wang, Xintao, Ke Yu, Shixiang Wu, Jinjin Gu, Yihao Liu, Chao Dong, Yu Qiao, and Chen Change Loy. "Esrgan: Enhanced super-resolution generative adversarial networks." In Proceedings of the European conference on computer vision (ECCV) workshops, pp. 0-0. 2018.
- Rabbi, Jakaria, Nilanjan Ray, Matthias Schubert, Subir Chowdhury, and Dennis Chao. "Small-object detection in remote sensing images with end-to-end edge-enhanced GAN and object detector network."
- Lamba, Mohit, and Kaushik Mitra. "Restoring extremely dark images in real time." In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 3487-3497. 2021.