Form 2: Literature Documents

1. Team No : 02

2. Project Title: Focus+ : From Blurry to Brilliant

3. Problem Statement: The existing video restoration techniques are remarkable in addressing challenges such as noise reduction and upscaling, significantly enhancing overall video quality. However, a limitation can be noted as these methods often fall short in preserving and enhancing crucial facial details, resulting in a final output that appears blurry and unnatural. This drawback motivates a need for a targeted solution focused specifically on facial feature enhancement within videos.

4. Problem Illustration:

Existing video restoration approaches prioritize noise reduction, color correction, and upscaling, effectively addressing temporal artifacts but neglecting crucial facial details. This results in:

Blurred facial features: Skin texture, sharpness, and definition remain compromised, compromising visual fidelity and impacting emotional engagement.

Unnatural smoothing: Standard methods often over-smooth faces, leading to unrealistic, plastic-like appearances.

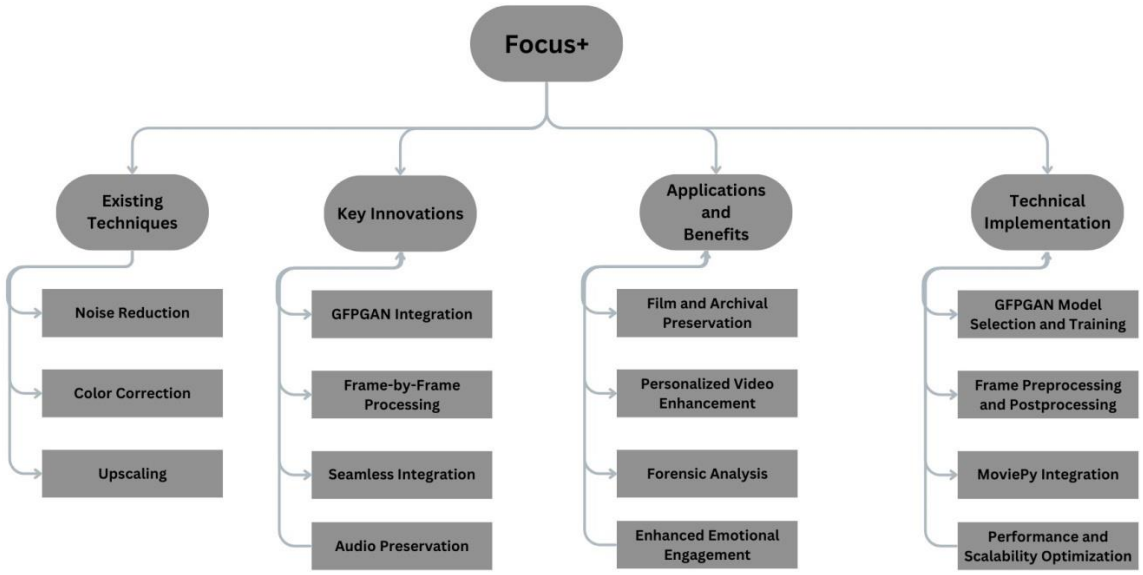
Temporal inconsistencies: Facial enhancements can vary across frames, creating jarring jumps in quality and disrupting immersion.

This lack of focus on facial details limits the potential of video restoration for applications like:

Film and archival preservation: Historical footage loses its expressiveness and emotional impact due to blurry faces.

Forensic analysis: Facial recognition and identification become challenging due to unclear features.

5 .Concept Tree:



6. Comparison of Existing Strategies for Problem solve:

S.No	Strategies	Advantages	Disadvantages
1	Noise Reduction	Improves overall video quality by reducing temporal artifacts like grain and noise.	Can blur facial details and reduce sharpness.
2	Color Correction	Enhances the visual appeal of the video by correcting color casts and imbalances.	Can wash out facial details and make them appear less defined.
3	Upscaling	Increases the resolution of the video, making it appear sharper and more detailed.	Can introduce new artifacts and distortions and may not improve the clarity of facial features.
4	Facial Landmark Detection and Enhancement	Identifies facial landmarks (e.g., eyes, nose, mouth) and applies targeted enhancement techniques.	Can be computationally expensive and may not always produce natural-looking results.
5	Deep Learning-Based Super-Resolution	Uses deep learning models to upscale the video and improve the clarity of facial features.	Can be very effective but requires a large amount of training data and can be computationally expensive.
6	Pulse CVPR 20	Leverages deep learning to upscale videos while preserving details like textures and edges.	May not specifically address facial features, leading to inconsistent enhancement across different elements.
7	Reconstruction by Audio/Video Separation	Separates audio and video tracks, performs separate restorations, and recombines them.	Can be effective for noise reduction and upscaling but may not target facial features specifically.
8	Focus+ (proposed)	Frame-by-frame enhancement with GFPGAN for targeted facial detail improvement, seamless integration with MoviePy, and audio preservation.	Requires GPU for GFPGAN processing, and performance may depend on video complexity.

7. Comparison of Existing Method from selected Strategies :

S.no	Author	Strategy	Advantages	Disadvantages
1	Lingbo Yang et al	HiFaceGAN	HiFaceGAN is a multi-stage framework that progressively replenishes facial details based on hierarchical semantic guidance. It has superior performance over a wide range of challenging restoration subtasks.	Computationally expensive due to deep learning model. Limited control over specific facial features. Opaque inner workings hinder further improvement.
2	Li et al2	DFDNet	DFDNet is a deep face dictionary network for face restoration that selects dictionary features most like the input. It has been used to guide the restoration process via dictionary feature transformation.	The bandwidth requirement is high, and the process is complex.
3	Wan et al	WAN	WAN enables communication, information exchange, and much more between devices worldwide. It is significant for global businesses.	The primary disadvantage of WAN is they are expensive compared to other networks5. The peripherals and devices required for the initial installation of the WAN setup are too expensive.
4	Sachit Menon	Pulse	Pulse Code Modulation is used in long-distance communication. The efficiency of the transmitter in PCM is high.	The bandwidth requirement is high. PCM is a complex process, since it involves encoding, decoding and quantisation of the circuit.

8. References:

Wang, X., Yu, L., Zhou, S., & Zhang, D. (2021). GFPGAN: Towards real-world blind face restoration with generative facial prior. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 1505-1514).

Yue, W., Liu, Y., Liu, J., Sun, J., & Wang, T. (2022). DifFace: Blind face restoration with diffused error contraction. arXiv preprint arXiv:2206.05477.

Jiang, Y., Zhou, Z., Loy, D., & Tang, X. (2022). GLEAN: Generative latent bank for large-factor image super-resolution. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 3149-3158).

McFee, B. (2021). moviepy: Video editing for Python.

Team Members: **Signature Supervisor**

1. A. Gayathri

2. K. Sathwika

3. P. Akhil