High-Fidelity Face Reconstruction from Degraded Imagery

Final Year Design Project – I (CSE 4000A)

Date: 8th of August, 2025

011221503	Jiyasmin Akter Sinthiya
011221149	Eila Afrin
011221130	Md. Shakib Al Islam
011221235	Rifat Ibne Yousuf
011221129	Mark Protik Mondol

United International University

Dept. of Computer Science & Engineering

Table of Contents

Introduction

Context

Surveillance cameras are very important for finding evidence in crimes. But sometimes the video they record is blurry or low quality. In real life, CCTV or video footage is often not clear it can be low resolution, noisy and hard to see. If we look at just one blurry frame, it usually does not have enough detail to recognize a person. But a video has many frames, each showing the face from slightly different angles, movements, and lighting. If we use these frames together in a smart way, we can recover more details of the face than from just one frame.

Problem statement

How can we reconstruct a high quality, identity consistent face image from noisy, blurry or low resolution CCTV/video footage so the result increases the chance of correct identification while minimizing hallucinated, misleading details?

Brief background

Modern AI tools for fixing images like super resolution, deblurring, and denoising and special face generation models (GANs and diffusion models) can make blurry faces look clear and realistic. But the problem is, these tools don't always keep the real person's identity, so it is still hard to know who they are. A better way could be to combine information from many frames of a video (to get extra real details) with identity focused restoration methods (to make sure the face stays true to the person) and then use strong testing plus human review. This could help create face reconstructions that are good enough for real forensic work.

Objectives

Main goal

Build and show a complete system that takes blurry or noisy CCTV footage or video clips and produces the clearest, most accurate face reconstruction possible while keeping the person's true identity. The system should also show how confident it is in the results, clearly explain when and why it might fail, and include rules to make sure it is used ethically.

Specific goals

1. Build the video pipeline – Take video frames, find where the person is in each frame, and either choose the best frame or combine several frames to get the most detailed view of their face.

2. Keep the true identity – Use identity-preserving methods like ArcFace or InsightFace, and combine identity information from multiple frames to make sure the reconstructed face still looks like the real person.

3. Score the results – Create a confidence score for each reconstructed face by combining how stable the identity features are with how good the frame quality is.

4. Make it easy to test – Build a simple demo interface and share reproducible code. Also run both automated tests and a small human study to check how well it works and when it fails.

5. Address ethics and privacy – Write a detailed section about ethics and privacy, including bias checks and clear rules for using human review in the process.

Scope

In scope

- Work with pre-recorded CCTV or video clips (10–60 seconds) from a single camera.
- Include face detection, face tracking, combining information from multiple frames, restoring the face, checking identity, and building a demo interface.
- Test the system using public datasets (like surveillance and face datasets) and run a small, controlled human study with people who give consent.
- Do an ethical review, check for bias, and give recommendations for using the system responsibly.
- Real time use on large, live city wide CCTV.
- Making final legal identifications the system will only assist and always include human review.
- Creating or enforcing rules for using large biometric face databases.

Methodology

1. Data & Preprocessing

- Use public face and surveillance datasets:
    - High quality face datasets (FFHQ, CelebA-HQ) for training and as priors. Also we can build our own custom dataset by collecting human portrait images with full consent from the individuals.
    - Surveillance style datasets (SCface, ChokePoint, UCCS) for testing.
    - Include a small real CCTV clip with consent from the participants.
- Process videos by:
    - Extracting individual frames.
    - Detecting faces (using RetinaFace or MTCNN).
    - Aligning faces using landmarks.
    - Tracking faces across frames (DeepSORT or SORT).
- Create a synthetic degradation pipeline to add realistic noise, blur, compression, occlusion, and low resolution to high quality images for controlled testing.

2. Multi-Frame Fusion & Restoration

- Try multi frame restoration models (EDVR, BasicVSR/BasicVSR++, VRT) to merge nearby frames into a higher quality image or detect and crop the face first, align those crops across frames (using optical flow or landmarks), then run multi frame fusion on the aligned faces.

3. Face Specific Enhancement & Identity Preservation

- Use face enhancement models (GFPGAN, DFDNet, Real-ESRGAN) on the fused face crops.
- Train or fine tune with identity loss so the restored face is as close as possible to the real person's face.
    - Measure this with cosine distance between embeddings from ArcFace or InsightFace.
- If a high quality reference photo is available, use it to guide the model to keep consistent fine details.

4. Embedding Aggregation & Scoring

- Extract face embeddings from each restored frame.
- Combine them with a weighted average, giving more weight to sharper and more confident frames.
- Create a trust score for each person's track based on:
    - Number of high-quality frames.
    - Stability of the embeddings.
    - Model's own confidence.
    - Measured restoration quality.

5. Evaluation

- Quantitative tests:
    - PSNR, SSIM, LPIPS, FID (when possible).
    - Identity cosine similarity with ground truth images.
- Qualitative tests:
    - Side by side image galleries.
    - Blind human matching tests with consented participants to check real world identifiability.
- Analyze failures based on pose, occlusion, lighting, ethnicity, and distance from camera.

6. UI & Demo

- Build a simple web app to show:
    - The original video clip.
    - Detected faces and tracks.
    - The best restored face per track.
    - Top similar matches from a small local watchlist.
    - The trust score for each reconstruction.