The Blurred Intruder: Handling Low-Resolution CCTV Feeds

Author: **22CSB0C20 CHANDU CHERUPALLY (CSE-A)**
Course: **Biometrics and Pattern Recognition**

1. Introduction

In real-world surveillance scenarios, facial images are often captured in **poor lighting**, **long-range**, or **low-resolution** conditions such as those from **CCTV cameras**.
Traditional face recognition systems—trained on high-quality, front-facing images—tend to perform poorly when the probe (test) image is of **low resolution** and the gallery (enrolled) image is of **high resolution**.

This project addresses a key challenge:

How can we reliably recognize individuals from low-quality surveillance images when the enrolled database contains only high-quality photos?

The objective of this work is to implement and analyse various methods that handle **resolution mismatch** between gallery and probe images. By simulating low-resolution probes and experimenting with **bicubic interpolation** and **super-resolution**, this project demonstrates the effect of different enhancement methods on face recognition accuracy.

This problem has high relevance in domains like **law enforcement**, **forensic identification**, and **automated surveillance**, where face recognition must work even in degraded visual conditions.

2. Methodology

2.1 Overview

The project follows a systematic pipeline:

1. Build a **gallery** of high-resolution enrolled faces.
2. Create **probe images** by artificially degrading gallery faces to simulate CCTV-like low-resolution captures.
3. Compare several processing methods for handling resolution mismatch:
    - **Naive Comparison:** Directly compares high-res and low-res faces.
    - **Bicubic Probe:** Upscales low-resolution probes using bicubic interpolation.
    - **Down-sample Gallery:** Degrades gallery faces to match the low resolution of probes.

- **Super-Resolution (SR) Probe:** Uses deep learning–based super-resolution for probe enhancement.

2.1.1 Datasets Used

The **Labelled Faces in the Wild (LFW)** dataset was chosen because it is a well-established benchmark for evaluating face recognition systems under **unconstrained and realistic conditions**. The images in LFW are collected from the web, containing natural variations in **lighting, pose, expression, and background**, which closely resemble challenges faced in CCTV environments. Its **moderate size and easy accessibility** make it ideal for experimentation and performance comparison across different enhancement and recognition methods without requiring heavy computational resources.

2.2 Algorithmic Flow

Step 1 — Face Detection and Cropping:
Each image undergoes face detection using the face-recognition library (HOG-based model). The detected face is cropped and resized to a consistent embedding size (160×160).

Step 2 — Feature Extraction:
Using face_recognition.face_encodings(), a **128-dimensional facial embedding** is computed for each face. These embeddings are later used to measure **cosine similarity** between probe and gallery features.

Step 3 — Resolution Simulation and Enhancement:
Each method applies a different preprocessing strategy on probes:

- **Naive:** Downscale probe to 24×24 and resize to 160×160 using nearest-neighbor interpolation (no enhancement).
- **Bicubic Probe (** *Preferred and Implemented* **):** Downscale probe to 24×24, then upscale to 160×160 using bicubic interpolation to restore gradient smoothness.
- **Down-sample Gallery:** Degrade high-res gallery to match probe resolution before comparison.
- **Super-Resolution (SR) Probe:** Apply pretrained DNN super-resolution model for reconstruction.

Step 4 — Similarity Computation and Evaluation:
For every probe image, cosine similarity is calculated against all gallery embeddings.

- **Rank-1 Accuracy:** The percentage of probes whose best match corresponds to the correct identity.
- **ROC Curve and AUC:** Plot true positive rate vs. false positive rate to measure discriminability.

2.3 Parameter Setup

- **Embedding size:** 160 × 160 pixels
- **Simulated low resolution:** 24 × 24 pixels
- **Feature model:** face_recognition (Dlib-based 128D embeddings)
- **Dataset:** lfw_dataset_fixed
- **Evaluation metrics:** Rank-1 accuracy, AUC (Area under ROC curve)

3. Results and Analysis

3.1 Quantitative Results

	timestamp	method	seed	max_id	low_res	gallery_size	probes	rank1	auc
1	2025-10-22T21:26:01.753357	bicubic_probe	123	50	24	50	135	0.933333	0.996895
2	2025-10-22T21:29:24.713037	downsample_gallery	123	50	24	48	131	0.908397	0.995617
3	2025-10-22T21:38:07.022495	sr_probe	123	50	24	50	135	0.933333	0.996895
4	2025-10-22T22:32:00.935901	downsample_gallery	123	200	24	183	567	0.809524	0.991335
5	2025-10-22T22:38:59.142103	sr_probe	123	200	24	193	581	0.857143	0.996647
6	2025-10-23T13:31:06.920087	bicubic_probe	123	200	24	193	584	0.847603	0.996336
7	2025-10-23T13:50:41.944615	naive	123	200	24	193	451	0.609756	0.984001
8	2025-10-23T13:57:58.653670	bicubic_probe	123	200	24	193	584	0.847603	0.996336
9	2025-10-23T16:16:10.692722	bicubic_probe	123	50	24	50	135	0.933333	0.996895
10	2025-10-23T16:17:28.665317	naive	123	50	24	50	108	0.750000	0.984570
11	2025-10-23T17:12:26.117180	naive	123	50	24	50	108	0.750000	0.984570

The exact Rank-1 and AUC values are based on runs with max_id=200 and low_res=24

3.2 Observations

- **Naive comparison**, surprisingly, achieved high accuracy on the small controlled dataset (75%), but in real low-quality CCTV scenarios, it typically collapses due to domain mismatch. The strict version often fails to extract valid embeddings.
- **Bicubic Probe** and **SR Probe** both improved the stability of embeddings when resolution was severely degraded, achieving similar AUCs (~0.997).
- **Dow-sample Gallery** yielded a lower accuracy (90.8%) but provides a consistent low-resolution feature space, which can be more robust when all probe images are extremely degraded.
- On a larger subset (200 identities), **performance dropped to ~81%**, confirming that real-world scalability amplifies the resolution mismatch effect.
- The **ROC curves** plotted for each method show clear separation between genuine and impostor distributions, confirming strong discriminative performance even after degradation handling.

3.3 Visual Analysis

Example visual outputs (saved in the examples/ directory) illustrate the impact of resolution handling:

- **Naive (strict)**: Highly pixelated faces lead to missing embeddings.
- **Bicubic Probe**: Smoother and more face-like structures improve feature extraction.
- **SR Probe**: Reconstructed images recover sharper facial details and enable more accurate embeddings.
- **Down-sample Gallery**: Creates uniform, low-frequency face templates, trading off detail for consistency.

3.4 Trade-offs

Method	Advantages	Limitations
Naive	Fast, simple baseline	Fails on real low-res probes
Bicubic Probe	Simple, effective, no model needed	Limited reconstruction quality
Downsample Gallery	Equalizes both domains	Loses gallery details, lower recognition power
SR Probe	Restores detail, high accuracy	Requires model loading, longer runtime

3.5 Why Prefer **Bicubic Probe** Over Other Methods

The bicubic probe method is preferred in the current context of handling CCTV images because it offers an ideal balance between performance, robustness, and practicality. CCTV footage often suffers from low resolution, noise, and varying lighting or camera qualities, and bicubic interpolation helps enhance probe images by preserving essential details like edges and textures while reducing artifacts. This leads to clearer and more consistent images for recognition. Unlike deep super-resolution methods, bicubic interpolation is lightweight, fast, and does not require GPU resources or training, making it suitable for real-time surveillance applications. It also provides consistent enhancement across different camera sources, ensuring stable recognition even in cross-camera scenarios. Additionally, the mild smoothing effect of bicubic interpolation reduces noise without losing critical facial or structural details, improving the accuracy of feature extraction. Its generalization ability is another advantage—it performs reliably across various CCTV conditions without retraining or dataset-specific tuning. Overall, the bicubic probe method offers strong accuracy, robustness, and computational efficiency, making it a practical and reliable choice for improving identification performance in real-world CCTV environments.

4. Conclusion

This project demonstrated the resolution mismatch problem in face biometrics and compared multiple strategies for handling it.
While direct (naive) matching is unreliable for true CCTV scenarios, **bicubic interpolation** and **super-resolution enhancement** significantly improve feature stability and identification accuracy.

The experiments showed that even simple interpolation techniques can enhance low-resolution face embeddings to a level close to high-resolution ones. Super-resolution offers further gains but at increased computational cost.

In conclusion, **bicubic and SR-based approaches** provide a practical balance between speed and accuracy for real-world surveillance applications. Future work can extend this to deep learning–based feature alignment and cross-resolution face synthesis.

5. References

1. Parkhi, O. M., Vedaldi, A., & Zisserman, A. (2015). *Deep Face Recognition*.
2. Huang, G. B. et al. (2007). *Labeled Faces in the Wild: A Database for Studying Face Recognition in Unconstrained Environments*.
3. Lim, S., et al. (2020). *Enhanced Super-Resolution for Low-Resolution Face Recognition*. IEEE T-BIOM.
4. Dlib & face_recognition Python Libraries — https://github.com/ageitgey/face_recognition
5. Attaching the ROC Images achieved in the performance evaluations of different methods.

   https://drive.google.com/drive/folders/1VMk3zb6cP1H0PY8XE6kgW6kOvu2YCTdb?usp=drive_link