Comparative Study of Face Restoration Methods: GFPGAN and CodeFormer

Monica Stef, Enqi Liang, Aidan Sibley

1. Problem Statement

Image restoration is a fundamental task in computer vision, with its applications in everyday photo enhancement to more specific areas such as cultural heritage preservation, law enforcement, and identity verification. In this domain, human face restoration is one of the most challenging problems being that faces are delicately structured, and humans are extremely sensitive to distortions. Even small inaccuracies can completely skew the whole restored image. Traditional image restoration techniques often struggle with this balance of producing over-smoothed results that fail to preserve unique identity features.

Recent advances in deep learning, specifically in Generative Adversarial Networks (GANs) and transformer architectures, have enabled breakthroughs in face restoration. Two of the most notable contributions in this area are GFPGAN (Wang et al., 2021) and CodeFormer (Zhou et al., 2022). Both methods have the goal of restoring high-quality, realistic faces from degraded inputs, however they take different approaches. GFPGAN leverages the powerful latent priors of StyleGAN to synthesize natural looking facial details, while CodeFormer uses a transformer based architecture allowing for adjustable restoration. Despite their popularity, there has not been an in-depth comparison of these two methods across both synthetic datasets and real-world old photographs.

The goal of this project is to implement and evaluate GFPGAN and CodeFormer side by side, focusing on their effectiveness in restoring degraded or historical face images. The evaluation will deal with not only pixel accuracy, but also perceptual quality and identity preservation. By the end of the term, the project will provide a robust understanding of the strengths and weaknesses of these two models and deliver a working demo that highlights their capabilities.

2. Related Work

The face restoration problem has been studied for decades, with early approaches relying on handcrafted priors and filtering techniques. However, these methods often failed to generalize to diverse conditions. The emergence of deep convolutional neural networks has made significant advances in the field.

- **GFPGAN (Wang et al., 2021)**: GFPGAN introduced a method that integrates pre-trained StyleGAN priors with a restoration network. By combining GAN-based priors with degraded image features, GFPGAN generates realistic faces while maintaining natural texture.
- **CodeFormer (Zhou et al., 2022)**: This model uses a transformer-based approach for face restoration. Its key innovation is the introduction of a fidelity-quality trade-off parameter which enables users to achieve both identity preservation and visual enhancement. It has shown strong generalization to real-world old photographs.

- **DeblurGAN (Kupyn et al., 2019)**: An earlier GAN-based method for image deblurring, which set an important foundation for applying adversarial training to restoration tasks.
- **Perceptual Metrics (LPIPS, Zhang et al., 2018)**: Proposed as an alternative to pixel-level metrics like PSNR, this evaluates perceptual similarity by comparing deep feature embeddings, making it fit for restoration tasks.
- **Identity Preservation (Deng et al., 2019 – ArcFace)**: Methods such as ArcFace introduced robust facial recognition embeddings, which can also be used to measure how well restored images maintain identity.

Together, these works highlight the shift from purely pixel-based restoration to approaches emphasizing perceptual quality and identity. Building on this foundation, this project will compare GFPGAN and CodeFormer in a unified framework.

3. Datasets and Resources

This project will use a combination of synthetic and natural datasets:

- **CelebA-HQ and FFHQ**: Standard high-resolution face datasets. Degradations such as blur, noise, and compression will be synthetically applied to create paired training and testing data.
- **Historical photo collections (Flickr Public Domain, open archives)**: Real-world examples of degraded photographs, useful for qualitative evaluation.
- **Resources**: PyTorch for implementation, Google Colab or local GPU for training and inference, and pretrained models from the Hugging Face hub for both GFPGAN and CodeFormer.

4. Project Plan and Timeline

The project will be completed over several phases:

- **Weeks 1–2**: Literature review, finalize dataset collection, and replicate baseline results using pretrained GFPGAN.
- **Week 3**: Run CodeFormer pretrained model and perform initial comparisons with GFPGAN on synthetic degradations.
- **Week 4**: Evaluate both models on historical photos, prepare qualitative comparisons.
- **Week 5 (Midterm)**: Submit progress report with quantitative metrics (PSNR, SSIM, LPIPS) and identity evaluation using ArcFace embeddings.
- **Weeks 6–7**: Fine-tune both models on a subset of old photos and test generalization.
- **Weeks 8–9**: Conduct small user study to collect perceptual ratings of restored images.
- **Week 10**: Finalize experiments, perform error analysis, and prepare final report and demo.

5. Evaluation Metrics

The models will be evaluated using both quantitative and qualitative metrics:

Quantitative:

- PSNR and SSIM for pixel similarity.
- LPIPS for perceptual similarity.
- Cosine similarity of ArcFace embeddings to measure identity preservation.

Qualitative:
- Side-by-side visual comparisons.
- User study where participants rate realism and identity consistency.

This will ensure a fair evaluation that captures both technical and perceptual aspects of restoration.

6. Risks and Fallbacks
While the project is feasible within the term, a few risks exist:
    - **Compute limitations**: Fine-tuning large models may exceed available GPU resources.
        - Fallback: focus on inference with pretrained models and evaluation across datasets
    - **Dataset availability**: Historical photos may be limited or inconsistent in quality.
        - Fallback: emphasize synthetic degradations for quantitative testing
    - **Subjectivity of user studies**: Human ratings may vary.
        - Fallback: emphasize objective identity metrics

7. Expected Contributions
The expected outcomes of the project are:
    1. A comparison of GFPGAN and CodeFormer for face restoration
    2. Insights into the trade-offs between fidelity, perceptual quality, and identity preservation
    3. Demonstration of model performance on real-world old photographs
    4. A visual demo showcasing before and after restoration results

This project will not only compare and contrast two cutting-edge methods, but also provide practical insights into their applicability for cultural preservation and personal photo restoration projects.

Sources:
GFPGAN (Wang et al., 2021): https://github.com/TencentARC/GFPGAN
CodeFormer (Zhou et al., 2022): https://github.com/sczhou/CodeFormer
DeblurGAN (Kupyn et al., 2019):
https://openaccess.thecvf.com/content_ICCV_2019/papers/Kupyn_DeblurGAN-v2_Deblurring_Orders-of-Magnitude_Faster_and_Better_ICCV_2019_paper.pdf
Perceptual Metrics (LPIPS, Zhang et al., 2018):
https://openaccess.thecvf.com/content_cvpr_2018/papers/Zhang_The_Unreasonable_Effectiveness_CVPR_2018_paper.pdf
Identity Preservation (Deng et al., 2019 – ArcFace):
https://openaccess.thecvf.com/content_CVPR_2019/papers/Deng_ArcFace_Additive_Angular_Margin_Loss_for_Deep_Face_Recognition_CVPR_2019_paper.pdf