

Addressing Racial Bias in AI-Driven Photo Restoration: Enhancing Fidelity for African Facial Features

Abstract

AI-powered photo restoration tools have made it easier to preserve historical images by repairing damage, improving resolution, and adding color to grayscale photographs. However, these models often show racial bias, altering African facial features like skin tones, textures, and bone structures to fit Eurocentric standards. This paper examines these differences in common generative models, linking them to the lack of African representation in training sets like FFHQ and CelebA, where diverse ethnic groups make up less than 10% of the samples. We suggest a fine-tuning process that uses a curated dataset of 5,000 public domain African heritage images, enhanced with synthetic damage (such as scratches and fading) through adversarial training. Our method uses StyleGAN3 with perceptual loss functions tailored for ethnic-specific traits to ensure cultural fidelity.

The approach incorporates edge-detection modules to keep facial shapes intact and a custom loss function that combines FID and SSIM for quantitative assessment. Tests on a sample of 500 damaged photos show that baseline models change African features in 72% of cases (for instance, lightening skin tones by 15-25% in HSV space). In contrast, our fine-tuned model reduces this to 18%, improving FID from 0.42 to 0.21 and SSIM from 0.78 to 0.92.

This work emphasizes the importance of inclusive datasets in computer vision and provides a guide for fair AI in imaging. Future directions include integrating multiple methods for video restoration and creating ethical guidelines for dataset curation.

Keywords: AI bias, photo restoration, racial equity, generative models, computer vision

References

- [1] J. Buolamwini and T. Gebru, "Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification," Proc. Machine Learning Research, vol. 81, pp. 1-15, 2018.
- [2] T. Wang et al., "GFPGAN: Towards Real-World Blind Face Restoration with Generative Facial Prior," in Proc. CVPR, 2021, pp. 5898-5907.
- [3] A. Krizhevsky et al., "ImageNet Classification with Deep Convolutional Neural Networks," in Proc. NeurIPS, 2012, pp. 1097-1105.
- [4] T. Karras et al., "A Style-Based Generator Architecture for Generative Adversarial Networks," in Proc. CVPR, 2019, pp. 4401-4410.
- [5] R. Rombach et al., "High-Resolution Image Synthesis with Latent Diffusion Models," in Proc. CVPR, 2022, pp. 10684-10695.