

ABSTRAK

Face inpainting merupakan teknik yang bertujuan untuk menghasilkan kembali area wajah yang hilang atau tertutup berdasarkan informasi visual yang masih tersedia. Perkembangan metode berbasis *generative prior* menunjukkan kemampuan yang baik dalam menghasilkan wajah yang realistis dengan memanfaatkan generator *pretrained* sebagai sumber pengetahuan visual. Meskipun demikian, masih terdapat peluang untuk meningkatkan kualitas representasi fitur pada proses pemetaan citra menuju ruang laten generator melalui integrasi mekanisme perhatian yang ringan dan efisien. Oleh karena itu, penelitian ini mengkaji pengaruh integrasi *Coordinate Attention* pada model *face inpainting* berbasis StyleGAN3 terhadap kualitas hasil yang dihasilkan.

Penelitian ini menggunakan dataset CelebA-HQ yang terdiri atas 30.000 citra wajah beresolusi 256×256 piksel. Model yang dikembangkan mengadopsi arsitektur *encoder-generator* dengan StyleGAN3 sebagai *generative prior* yang digunakan dalam kondisi *frozen*. *Coordinate Attention* diintegrasikan pada bagian *context encoder* untuk meningkatkan representasi fitur spasial sebelum proses pemetaan ke ruang laten. Evaluasi dilakukan melalui studi ablasi pada beberapa konfigurasi penempatan *Coordinate Attention* serta dibandingkan dengan model *baseline* tanpa *Coordinate Attention*. Kinerja model diukur menggunakan metrik *Peak Signal-to-Noise Ratio* (PSNR), *Structural Similarity Index Measure* (SSIM), dan *Learned Perceptual Image Patch Similarity* (LPIPS) pada *full image* dan *hole region*. Selain itu, dilakukan pengujian hipotesis menggunakan *Wilcoxon Signed-Rank Test* untuk mengevaluasi signifikansi perbedaan performa antar model.

Hasil penelitian menunjukkan bahwa integrasi *Coordinate Attention* memberikan peningkatan performa yang konsisten dibandingkan model *baseline*. Konfigurasi terbaik diperoleh pada model A4 dengan penempatan *Coordinate Attention* pada lapisan *encoder* tingkat tinggi, yaitu *enc3* dan *enc4*, yang menghasilkan nilai LPIPS pada *hole region* sebesar 0,26001 dibandingkan model *baseline* sebesar 0,26929. Peningkatan juga terlihat pada metrik PSNR dan SSIM baik pada *full image* maupun *hole region*. Dari sisi kompleksitas komputasi, penambahan *Coordinate Attention* hanya meningkatkan GFLOPs dan jumlah parameter dalam jumlah yang relatif kecil. Hasil pengujian statistik menunjukkan nilai $p\text{-value} < 0,05$ pada seluruh metrik evaluasi, yang mengindikasikan bahwa perbedaan performa antara model *baseline* dan model dengan *Coordinate Attention* signifikan secara statistik. Temuan ini menunjukkan bahwa *Coordinate Attention* dapat digunakan sebagai mekanisme peningkatan representasi fitur yang ringan dan efektif pada model *face inpainting* berbasis StyleGAN3.

Kata Kunci: *Face Inpainting*, StyleGAN3, *Coordinate Attention*, *Generative Prior*

ABSTRACT

Face inpainting is a technique that aims to restore missing or occluded facial regions based on the remaining visual information. Recent advances in *generative prior*-based approaches have demonstrated promising capabilities in generating realistic facial images by leveraging pretrained generators as sources of visual knowledge. Nevertheless, there remains an opportunity to improve feature representation quality during the mapping process from image space to the latent space of the generator through the integration of lightweight and efficient attention mechanisms. Therefore, this study investigates the effect of integrating Coordinate Attention into a StyleGAN3-based *face inpainting* model on the quality of the generated results.

This study utilizes the CelebA-HQ dataset, which consists of 30,000 facial images with a resolution of 256×256 pixels. The proposed model adopts an *encoder-generator* architecture with StyleGAN3 serving as a frozen *generative prior*. Coordinate Attention is integrated into the *context encoder* to enhance spatial feature representation before the latent mapping process. Evaluation is conducted through an ablation study involving several Coordinate Attention placement configurations and is compared against a baseline model without Coordinate Attention. Model performance is assessed using Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index Measure (SSIM), and Learned Perceptual Image Patch Similarity (LPIPS) on both *full image* and *hole region* evaluations. In addition, a Wilcoxon Signed-Rank Test is performed to evaluate the statistical significance of performance differences between models.

The experimental results show that the integration of Coordinate Attention provides consistent performance improvements over the baseline model. The best performance is achieved by the A4 configuration, which applies Coordinate Attention to the higher-level encoder layers, namely enc3 and enc4, resulting in a hole-region LPIPS score of 0.26001 compared to 0.26929 for the baseline model. Improvements are also observed in PSNR and SSIM for both *full image* and *hole region* evaluations. From a computational perspective, the addition of Coordinate Attention introduces only a marginal increase in GFLOPs and model parameters. Statistical hypothesis testing yields p-values below 0.05 across all evaluation metrics, indicating that the performance differences between the baseline and Coordinate Attention models are statistically significant. These findings suggest that Coordinate Attention can serve as a lightweight and effective feature representation enhancement mechanism for StyleGAN3-based *face inpainting* models.

Keywords: *Face Inpainting, StyleGAN3, Coordinate Attention, Generative Prior*