UNIVERSITAS
GADJAH MADA
Pengembangan Arsitektur Dual-Stage Virtual Try-On Berbasis CatVTON Menggunakan Post-Processor Residual Refinement Module
Nathanael Aurelino Sulistyo, Faizah, S.Kom., M.Kom.; Aina Musdholifah, S.Kom., M.Kom. Ph.D
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

DAFTAR ISI

HALAMAN JUDUL **i**

HALAMAN PENGESAHAN **iii**

PERNYATAAN **iv**

PRAKATA **v**

DAFTAR ISI **vi**

DAFTAR TABEL **x**

DAFTAR GAMBAR **xi**

DAFTAR KODE **xii**

INTISARI **xiii**

ABSTRACT **xiv**

I PENDAHULUAN **1**

1.1 Latar Belakang 1
1.2 Rumusan Masalah 4
1.3 Batasan Masalah 4
1.4 Tujuan Penelitian 6
1.5 Manfaat Penelitian 6

II PENELITIAN TERKAIT **8**

2.1 Variational Autoencoders (VAE) 8
2.2 *Latent Diffusion Model* (LDM) 9
2.3 Arsitektur DS-VTON sebagai Inspirasi Pendekatan *Dual-Stage* 9
2.3.1 Arsitektur *Dual U-Net* dan Strategi *Mask-Free* 10
2.3.2 Tahap Resolusi Rendah (*Low-Resolution Stage*) 10
2.3.3 Tahap Resolusi Tinggi dan Proses *Blend-Refine Diffusion* 11
2.4 *Canny Edge Detection* 12
2.5 Proses Inferensi CatVTON 13

UNIVERSITAS
GADJAH MADA
Pengembangan Arsitektur Dual-Stage Virtual Try-On Berbasis CatVTON Menggunakan Post-Processor Residual Refinement Module
Nathanael Aurelino Sulistyo, Faizah, S.Kom., M.Kom.; Aina Musdholifah, S.Kom., M.Kom. Ph.D
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

2.5.1 *Agnostic Person Image Generation* + *Concatenation* 13
2.5.2 *Latent Space Transformation* 14
2.5.3 Proses Difusi & *Denoising* 16
2.5.4 Rekonstruksi Citra 16
2.6 *Cross-Attention* 17
2.7 *BlurPool* 18
2.7.1 Perbandingan Varian *Filter* pada *BlurPool* 19
2.8 *Residual Learning* 20
2.9 *Resize-Convolution* 21
2.9.1 Perbandingan dengan *Transposed Convolution* 22
2.10 *L1 Loss* 23
2.11 *Perceptual Loss* 23
2.12 *Group Normalization* (GN) 24
2.13 *Structural Similarity Index* (SSIM) 25
2.13.1 Intuisi Metrik 26
2.14 *Learned Perceptual Image Patch Similarity* (LPIPS) 26
2.14.1 Intuisi Metrik 27
2.15 *Fréchet Inception Distance* (FID) 27
2.15.1 Intuisi Metrik 28
2.16 Tinjauan Pustaka 28

III METODE DAN RANCANGAN PENELITIAN 35
3.1 Deskripsi Umum Penelitian 35
3.2 Akuisisi Data 37
3.3 Pra-pemrosesan 38
3.3.1 Kurasi Sampel Bertekstur Berbasis *Edge Density* 38
3.3.2 Konversi Resolusi dan Standardisasi Format 40
3.4 *Coarse Image Generation* dengan CatVTON 41
3.5 Pembagian Data Latih dan Validasi 41
3.6 Pemulihan Tekstur dengan *Residual Refinement Module* (RRM) 41
3.6.1 Ekstraksi Fitur dengan *Dual-Stream Encoder* 42
3.6.2 Fusi Fitur dengan *Cross-Attention Fusion* 45
3.6.3 Rekonstruksi *Residual* dengan *Residual Learning Decoder* 47
3.7 Konfigurasi *Loss Function* 51
3.7.1 L1 *Loss* ($\mathcal{L}_{L1}$) 51

UNIVERSITAS
GADJAH MADA
Pengembangan Arsitektur Dual-Stage Virtual Try-On Berbasis CatVTON Menggunakan Post-Processor Residual Refinement Module
Nathanael Aurelino Sulistyo, Faizah, S.Kom., M.Kom.; Aina Musdholifah, S.Kom., M.Kom. Ph.D
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

3.7.2 LPIPS *Loss* ($\mathcal{L}_{\text{LPIPS}}$) 52
3.8 Konfigurasi Pelatihan 53
3.9 Evaluasi Model 54
3.9.1 Evaluasi Kualitas Citra 54
3.9.2 Evaluasi Efisiensi Komputasional 55
3.10 Skema Evaluasi 56
3.10.1 Evaluasi Kuantitatif 56
3.10.2 Evaluasi Kualitatif 57
3.10.3 Studi Ablasi Parameter *Delta Scale* (δ) 57

IV IMPLEMENTASI 58
4.1 Lingkungan Implementasi 58
4.2 *Library* Implementasi 58
4.3 Akuisisi Data 60
4.4 Pra-pemrosesan 61
4.4.1 Kurasi Sampel Berbasis *Edge Density* 61
4.4.2 Konversi Resolusi dan Standardisasi Format 62
4.5 *Coarse Image Generation* dengan CatVTON 64
4.5.1 Persiapan Lingkungan CatVTON 64
4.5.2 Pembentukan *Triplet Input* 65
4.5.3 Pembuatan *File* Pasangan Inferensi 65
4.5.4 Eksekusi Inferensi CatVTON 66
4.6 Pemuatan Hasil Konversi Resolusi dan Inferensi CatVTON 67
4.7 Pembagian Data Latih dan Validasi 68
4.8 Pemulihan Tekstur dengan *Residual Refinement Module* (RRM) 69
4.8.1 Ekstraksi Fitur dengan *Dual-Stream Encoder* 70
4.8.2 Fusi Fitur dengan *Cross-Attention Fusion* 72
4.8.3 Rekonstruksi *Residual* dengan *Residual Learning Decoder* 72
4.9 Konfigurasi *Loss Function* 74
4.10 Konfigurasi Pelatihan 75
4.11 Evaluasi Model 77
4.11.1 Evaluasi Kualitas Citra 77
4.11.2 Evaluasi Efisiensi Komputasional 80

V HASIL DAN PEMBAHASAN 82
5.1 Hasil Pelatihan 82

UNIVERSITAS
GADJAH MADA
Pengembangan Arsitektur Dual-Stage Virtual Try-On Berbasis CatVTON Menggunakan Post-Processor Residual Refinement Module
Nathanael Aurelino Sulistyo, Faizah, S.Kom., M.Kom.; Aina Musdholifah, S.Kom., M.Kom. Ph.D
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

5.2 Hasil Evaluasi Kuantitatif 83
5.2.1 Evaluasi Kualitas Citra pada *Test Set* Penuh 83
5.2.2 Evaluasi Kualitas Citra pada *Test Set* Terkurasi 85
5.2.3 Stratifikasi Perbaikan pada Sampel Bertekstur 87
5.2.4 Distribusi Sampel Citra Membaik dan Memburuk 88
5.2.5 Evaluasi Efisiensi Komputasional 89
5.3 Hasil Evaluasi Kualitatif 90
5.3.1 Evaluasi Visual pada Citra Bertekstur Kompleks 90
5.3.2 Evaluasi Visual pada Citra dengan Kesalahan Struktural 94
5.4 Studi Ablasi Parameter *Delta Scale* (δ) 96
5.5 Sintesis Temuan dan Efektivitas RRM 98

VI KESIMPULAN DAN SARAN 101
6.1 Kesimpulan 101
6.2 Saran 102

DAFTAR PUSTAKA 103