

Laporan Final Project Kecerdasan Buatan (Lanjut)

Implementasi CycleGAN untuk Transformasi Citra
Wajah Realistik Menjadi Kartun pada Dataset Tidak
Berpasangan



Kelompok 2
Anggota Kelompok:

23.11.5580 | Calvin Perdana Michael Laoli

23.11.5574 | Yosua Aldrin Garanta

23.11.5617 | Abdi Wicaksono B. S

23.11.5566 | Surahman

Latar Belakang

Fenomena penggunaan avatar kartun di media sosial kini semakin menjamur karena dianggap mampu memberikan kesan artistik dan berbeda dari foto profil biasa. Namun realitanya, tidak semua orang memiliki keahlian seni digital untuk membuat ilustrasi wajah sendiri, dan proses pembuatan manual pun memakan waktu yang tidak sebentar. Cela inilah yang coba diisi oleh teknologi Computer Vision dan Deep Learning untuk mengotomatisasi proses artistik tersebut [2]. Meskipun riset terdahulu seperti CartoonGAN [1] pernah dilakukan, penerapan teknologi ini di lapangan masih menemui kendala.

Hambatan terbesar dalam mengembangkan AI "pelukis" ini terletak pada ketersediaan data. Algoritma yang umum digunakan, seperti Pix2Pix, menuntut syarat yang berat yaitu adanya dataset berpasangan (paired). Implikasi teknisnya, model ini menuntut ketersediaan dataset berpasangan (paired), di mana citra asli harus disandingkan dengan versi kartun pada pose yang persis sama. Realitanya, mengumpulkan data presisi semacam ini dalam skala besar adalah hal yang nyaris mustahil dilakukan [4].

Guna menyiasati kelangkaan tersebut, penerapan Unpaired Image-to-Image Translation melalui arsitektur CycleGAN [6] dinilai sebagai langkah yang paling realistik. Nilai lebih dari metode ini terletak pada keluwesannya dalam menyerap karakteristik visual lintas domain tanpa mengharuskan adanya pasangan data yang selaras. Studi ini secara khusus menguji ketangguhan model dalam melakukan alih gaya wajah yang akurat sekaligus mempertahankan fitur identitas subjek, kendati proses pelatihannya dihadapkan pada keterbatasan data kartun yang minim dan tidak proporsional (imbalanced dataset)."

Metode

1. Alur Pengerjaan (Project Flow)

Alur pengerjaan proyek ini terdiri dari beberapa tahapan utama:

1. **Pengumpulan Data:** Mengunduh dataset wajah asli dan dataset referensi kartun.
2. **Preprocessing:** Mengubah ukuran citra (*resizing*) menjadi 256x256 piksel, normalisasi nilai piksel ke rentang [-1, 1], dan pembagian data latih (*train*) dan uji (*test*).
3. **Perancangan Model:** Membangun arsitektur Generator dan Discriminator menggunakan framework PyTorch.
4. **Pelatihan (Training):** Melatih model secara iteratif untuk meminimalkan *Loss Function*.
5. **Evaluasi:** Menguji model menggunakan data baru dan menganalisis hasil visual.

2. Algoritma: CycleGAN

CycleGAN memperkenalkan Cycle Consistency Loss. Konsep utamanya adalah jika sebuah citra X diubah ke domain Y (menjadi kartun), lalu dikembalikan lagi ke domain X, maka hasilnya harus kembali menjadi citra awal.

Total Loss penjumlahan dari:

- *Adversarial Loss*: Agar gambar terlihat nyata bagi Discriminator.
- *Cycle Consistency Loss*: Agar struktur objek tetap terjaga.
- *Identity Loss*: Agar warna dan komposisi tetap terjaga.

3. Arsitektur Model

- **Generator**: Menggunakan arsitektur berbasis **ResNet (Residual Networks)** dengan 9 blok residual. ResNet dipilih karena kemampuannya mempertahankan informasi fitur wajah yang mendalam selama proses transformasi [3].
- **Discriminator**: Menggunakan arsitektur **PatchGAN** berukuran 70x70. Berbeda dengan klasifikasi biasa, PatchGAN menilai keaslian gambar berdasarkan tekstur pada patch-patch lokal gambar, bukan keseluruhan gambar, sehingga hasil tekstur kartun lebih tajam.

Dataset

Dataset yang digunakan dalam penelitian ini bersumber dari Kaggle dan terdiri dari dua domain berbeda:

1. **Domain A (Source - Real Face) [5]**:
 - Diambil dari dataset **CelebA** (CelebFaces Attributes).
 - Jumlah data latih yang digunakan: **4.000 citra**.
 - Karakteristik: Foto wajah selebriti dengan berbagai pose, pencahayaan, dan latar belakang.
2. **Domain B (Target - Cartoon Face)**:
 - Diambil dari dataset ilustrasi kartun/lukisan wajah.
 - Jumlah data latih yang digunakan: **193 citra**.
 - Karakteristik: Gambar 2D dengan tekstur *flat*, garis tepi tegas, dan warna *vibrant*.

Catatan Pengolahan Data:

Terdapat ketidakseimbangan ekstrem antara data wajah (4.000) dan kartun (193). Untuk mengatasinya, dalam proses loading data, dataset kartun mengalami random oversampling (dipanggil berulang kali) dalam satu epoch agar seimbang dengan iterasi dataset wajah.

Hasil Pengujian

1. Skenario Pengujian

Pelatihan dan pengujian dilakukan menggunakan perangkat keras lokal dengan spesifikasi:

- **GPU**: NVIDIA GeForce GTX 1660 Super (6GB VRAM).
- **Hyperparameters**:
 - *Optimizer*: Adam (Learning Rate = 0.0002, Beta1 = 0.5).
 - *Batch Size*: 1 (Dipilih karena keterbatasan VRAM dan efektivitas *Instance Normalization* pada style transfer).
 - *Epoch*: 10. (Menggunakan strategi *Early Stopping* untuk mencegah *overfitting* pada dataset kartun yang kecil).

2. Hasil Visualisasi

Berikut adalah hasil pengujian model Generator (G_AB) yang mengubah foto asli menjadi kartun:



(Gambar 1: Sebelah kiri adalah Foto Asli, Sebelah kanan adalah Hasil Generasi AI)

Analisa Hasil

Berdasarkan hasil visualisasi dan grafik pelatihan, didapatkan analisis sebagai berikut:

1. Kualitas Transformasi Gaya:

Model berhasil menangkap karakteristik utama gaya kartun dari 193 data referensi. Hal ini terlihat dari kulit wajah yang menjadi lebih halus (smoothing), hilangnya detail pori-pori realistik, serta penegasan garis kontras pada area mata dan alis yang menyerupai goresan lukisan.

2. Stabilitas Pelatihan:

Meskipun hanya dilatih selama 10 epoch, grafik loss menunjukkan konvergensi yang cepat. Generator Loss stabil di kisaran angka 3.0 - 4.0. Keputusan untuk menghentikan pelatihan di epoch 10 terbukti tepat; jika dilanjutkan lebih lama dengan data target yang terbatas (hanya 193), model berisiko mengalami overfitting (menghafal gambar kartun tertentu dan mengabaikan input wajah asli).

3. Preservasi Identitas:

Penggunaan Cycle Consistency Loss berfungsi dengan baik. Meskipun tekstur wajah berubah total menjadi kartun, struktur geometris wajah (bentuk hidung, senyum, posisi mata) tetap terjaga, sehingga identitas orang dalam foto masih dapat dikenali.

Kesimpulan

1. Proyek ini berhasil memvalidasi keandalan arsitektur CycleGAN dalam skema unpaired image-to-image translation. Terbukti, transformasi dari foto wajah ke kartun dapat dilakukan secara efektif tanpa perlu bergantung pada dataset berpasangan.
2. Ketidakseimbangan jumlah dataset (4.000 wajah vs 193 kartun) dapat diatasi dengan teknik *oversampling* dan pembatasan jumlah epoch (*early stopping*).
3. Model yang dihasilkan pada epoch ke-10 sudah cukup representatif untuk menghasilkan gambar kartun yang artistik dengan tetap mempertahankan identitas wajah asli.

Referensi

- [1] I. J. Goodfellow *et al.*, “Generative adversarial nets,” *Advances in Neural Information Processing Systems*, vol. 27.
- [2] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, “Unpaired image-to-image translation using cycle-consistent adversarial networks,” in *2017 IEEE International Conference on Computer Vision (ICCV)*, IEEE, Oct. 2017, pp. 2242–2251. Accessed: Jan. 01, 2026. [Online]. Available: <https://doi.org/10.1109/iccv.2017.244>
- [3] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, “Image-to-image translation with conditional adversarial networks,” in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Jul. 2017, pp. 5967–5976. Accessed: Jan. 01, 2026. [Online]. Available: <https://doi.org/10.1109/cvpr.2017.632>
- [4] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Jun. 2016, pp. 770–778. Accessed: Jan. 01, 2026. [Online]. Available: <https://doi.org/10.1109/cvpr.2016.90>
- [5] Z. Liu, P. Luo, X. Wang, and X. Tang, “Deep learning face attributes in the wild,” in *2015 IEEE International Conference on Computer Vision (ICCV)*, IEEE, Dec. 2015, pp. 3730–3738. Accessed: Jan. 01, 2026. [Online]. Available: <https://doi.org/10.1109/iccv.2015.425>
- [6] Y. Chen, Y.-K. Lai, and Y.-J. Liu, “CartoonGAN: Generative adversarial networks for photo cartoonization,” in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, IEEE, Jun. 2018, pp. 9465–9474. Accessed: Jan. 01, 2026. [Online]. Available: <https://doi.org/10.1109/cvpr.2018.00986>

Kontribusi & Distribusi Anggota Kelompok

Untuk memastikan transparansi penggeraan proyek, berikut adalah distribusi tugas anggota kelompok kami:

1. Abdi Wicaksono B.S:

- Implementasi kode program (*coding*) menggunakan Python dan PyTorch.
- Melakukan preprocessing dataset dan proses pelatihan (*training*) model menggunakan perangkat keras lokal.
- Melakukan *debugging* dan tuning hyperparameter model.
- Pembuatan PPT

2. Yosua Aldrin Garanta:

- Melakukan eksplorasi dan pengumpulan dataset (CelebA dan Cartoon) dari sumber open source (Kaggle).
- Pembuatan Video untuk menjelaskan Codingan dan Dataset

3. Calvin Perdana Michael Laoli:

- Menyusun dokumentasi laporan bab metode dan arsitektur sistem.
- Membantu testing produk akhir

4. Surahman:

- Melakukan pengujian model (*testing*) terhadap data foto baru.