

Laporan Final Project Kecerdasan Buatan (Lanjut)

Batik Generator Lightweight GAN



Kelompok 14
Anggota Kelompok:
1. 23.11.5595 | Wirawan Ariyoso
2. 23.11.5596 | Yuafiq Alfin Al Khasan
3. 23.11.5609 | Adhitya Surya Handika

1. Latar Belakang

Batik adalah warisan budaya Nusantara yang tidak hanya bernilai seni tinggi, tetapi juga sarat akan simbolisme filosofis dari masyarakat pembuatnya. Kompleksitas visualnya—mulai dari ragam motif hingga pola warna yang luas—sering kali menyulitkan sistem visi komputer dalam mengelola dan mencari citra spesifik. Kendati demikian, riset terkini menunjukkan bahwa metode Deep Learning, terutama Convolutional Neural Networks (CNN), mampu mengungguli pendekatan konvensional dalam menangkap detail unik dari citra batik tersebut [1].

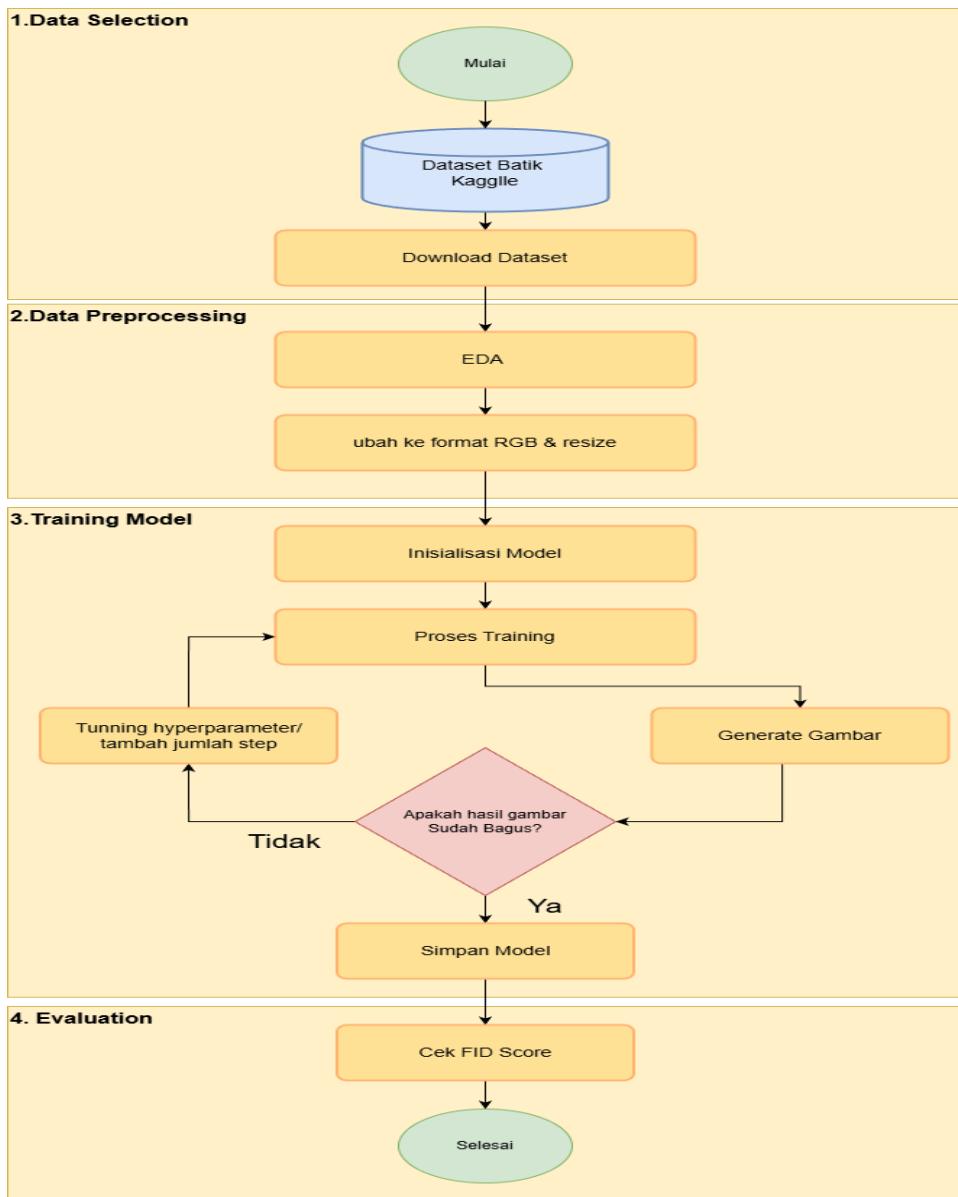
Di sisi lain, ranah sintesis citra batik masih menghadapi kendala teknis yang cukup pelik, meskipun teknologi ekstraksi fitur sudah semakin maju. Secara teori, Generative Adversarial Networks (GAN) memang menawarkan kemampuan pembelajaran representasi secara unsupervised [2]. Sayangnya, model generatif modern kerap beroperasi layaknya "kotak hitam" dengan parameter yang sulit dipahami [3]. Bahkan model mutakhir sekelas StyleGAN2 membutuhkan dataset yang sangat besar agar hasilnya realistik dan terhindar dari isu entanglement [3]. Akibatnya, pada kondisi data terbatas, performa GAN sering kali merosot tajam karena diskriminator justru mengalami overfitting, yang berujung pada ketidakstabilan pelatihan dan rendahnya kualitas visual [4].

Masalah efisiensi data dan stabilitas ini menjadi landasan utama penggunaan arsitektur yang lebih ringan. Dalam kondisi komputasi rendah dan sampel terbatas, diskriminator gagal memberikan gradien informatif bagi generator, sehingga memicu fenomena mode collapse. Untuk mengatasi hal ini, diperlukan pendekatan yang mampu menstabilkan dinamika pelatihan, salah satunya melalui aturan pembaruan laju pembelajaran Two Time-Scale Update Rule (TTUR) yang terbukti secara matematis dapat membawa model menuju local Nash equilibrium yang stasioner. Penggunaan metrik Fréchet Inception Distance (FID) menjadi standar evaluasi yang lebih konsisten dibandingkan Inception Score untuk mengukur kemiripan distribusi antara citra batik asli dan buatan[5].

Proyek ini mengadopsi arsitektur Lightweight GAN yang dirancang untuk mencapai konvergensi cepat pada resolusi tinggi (hingga 1024x1024 piksel) dengan sumber daya minimal. Keunggulan arsitektur ini terletak pada modul Skip-Layer channel-wise Excitation (SLE) yang memperkuat aliran gradien antar lapisan tanpa beban konvolusi yang berat, serta penggunaan diskriminator berbasis auto-encoder yang dilatih secara self-supervised untuk mengekstrak fitur yang lebih deskriptif[6]. Integrasi teknik Differentiable Augmentation (DiffAugment) juga diterapkan untuk meningkatkan efisiensi data dengan menerapkan augmentasi yang terdiferensiasi pada

sampel riil maupun buatan, sehingga menjaga keseimbangan optimasi antara generator dan diskriminator. Melalui kombinasi strategi ini, generator diharapkan mampu menghasilkan motif batik yang secara estetika valid dan secara statistik mendekati distribusi data asli meskipun dilatih dengan dataset yang terbatas[4].

2. Metode:



Gambar1: Fowchart penggeraan proyek

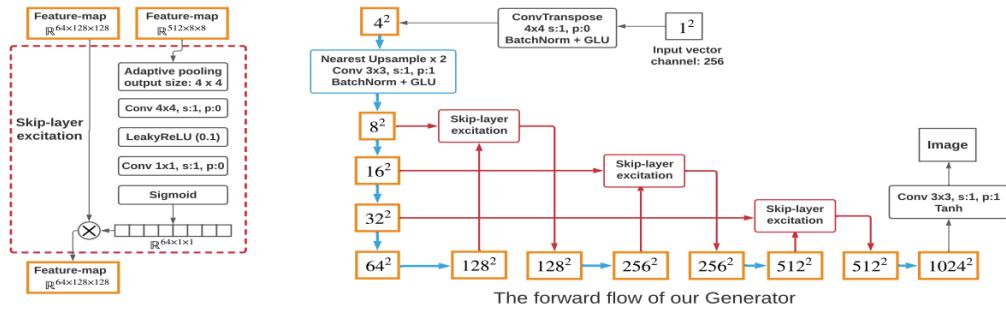
Implementasi proyek Batik Generator ini dirancang melalui sebuah pipeline riset yang sistematis, dimulai dari fase akuisisi data hingga tahap optimasi model menggunakan paradigma transfer learning. Proses diawali dengan ekstraksi dataset dari repositori publik Kaggle melalui pustaka kagglehub untuk mendapatkan basis data mentah yang terdiri dari berbagai motif batik.

Mengingat integritas data visual sangat krusial dalam pelatihan jaringan saraf generatif, dilakukan audit dataset dan pra-pemrosesan mendalam (Exploratory Data Analysis - EDA) untuk menstandardisasi input; citra dengan format warna yang tidak konsisten seperti RGBA, Grayscale, atau Palette dikonversi secara sistematis menjadi ruang warna RGB. Agar kompatibel dengan arsitektur model, citra masukan dinormalisasi secara spasial menjadi resolusi 256x256 piksel menggunakan teknik resize dan center crop. Eksperimen komputasi dijalankan pada dua lingkungan berbeda, yakni platform Google Colab (GPU Tesla T4 15GB) serta perangkat lokal yang didukung GPU RTX 3060 (12GB).

Strategi pelatihan sendiri dirancang dalam dua fase utama. Fase awal berfokus pada base training menggunakan dataset Batik Parang dan Megamendung untuk mempelajari representasi tekstur dasar. Setelah itu, proses dilanjutkan ke tahap transfer learning melalui fine-tuning pada dataset target (Batik Kawung). Pendekatan ini bertujuan untuk mempercepat konvergensi sekaligus memitigasi kendala keterbatasan jumlah sampel pada domain target.

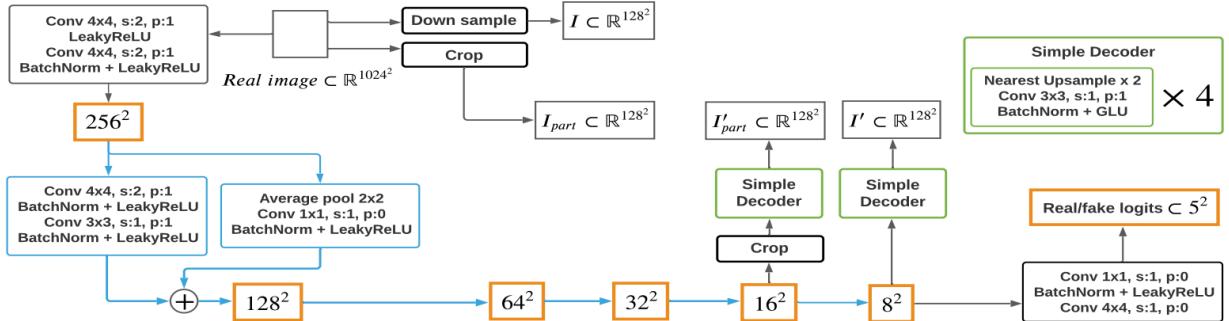
Secara algoritmik, proyek ini mengadopsi kerangka kerja **Lightweight GAN** yang dirancang untuk efisiensi komputasi tinggi namun mampu menghasilkan sintesis citra berkualitas pada resolusi tinggi. Dalam menjaga stabilitas dinamika adversarial, model ini mengintegrasikan teknik *Differentiable Augmentation* (DiffAugment) yang menerapkan augmentasi terdiferensiasi secara simultan pada sampel riil dan buatan, sehingga mencegah diskriminator mengalami *overfitting* tanpa mengubah distribusi target secara permanen[4]. Proses optimasi menggunakan algoritma Adam dengan penerapan *Two Time-Scale Update Rule* (TTUR), yang secara teoretis menjamin model mencapai *local Nash equilibrium* melalui penggunaan laju pembelajaran yang berbeda antara generator dan diskriminator. Evaluasi performa model dilakukan secara kuantitatif melalui pemantauan skor *Fréchet Inception Distance* (FID) secara berkala, yang berfungsi untuk mengukur kedekatan statistik antara distribusi citra asli dan citra hasil sintesis[6].

Arsitektur model yang diimplementasikan memiliki desain minimalis namun sangat efektif dalam menangkap dependensi fitur yang kompleks. Pada sisi **Generator (G)**, diintegrasikan modul **Skip-Layer channel-wise Excitation (SLE)** yang menghubungkan aktivasi dari lapisan resolusi rendah langsung ke peta fitur resolusi tinggi. Desain ini memungkinkan aliran gradien yang lebih kuat sepanjang jaringan tanpa beban komputasi konvolusi yang berat, sekaligus memfasilitasi pemisahan otomatis antara konten spasial dan atribut gaya.



Gambar2: Struktur modul skip-layer excitation. Kotak kuning mewakili *feature-maps* (kami menunjukkan ukuran spasial dan menghilangkan nomor saluran), kotak biru dan panah biru mewakili struktur up-sampling yang sama, kotak merah berisi modul SLE seperti yang diilustrasikan di sebelah kiri.

Sementara itu, **Diskriminatator (D)** dikonfigurasi sebagai *feature-encoder* yang dilengkapi dengan dekoder tambahan untuk tugas rekonstruksi citra secara mandiri (*self-supervised*). Pendekatan berbasis *auto-encoder* ini memaksa diskriminatator untuk mengekstraksi representasi fitur yang lebih deskriptif dan menyeluruh dari seluruh area gambar, yang terbukti krusial dalam menstabilkan pelatihan pada dataset kecil.



Gambar3: Struktur dan alur maju Diskriminatator. Kotak biru dan panah mewakili struktur residual down-sampling yang sama, kotak hijau berarti struktur dekoder yang sama.

3. Dataset

Dataset yang digunakan dalam proyek ini bersumber dari repositori publik Kaggle bertajuk "dataset-fix" (versi 1) yang dikurasi oleh kontributor hydiexe. Secara makro, dataset tersebut mencakup total sekitar 3.000 citra yang terdistribusi ke dalam 20 kelas motif batik berbeda. Guna menguji efektivitas metode transfer learning pada arsitektur lightweight, penelitian memfokuskan eksperimen pada dua subset spesifik yang masing-masing berjumlah 150 citra.

4. Hasil Pengujian

a. Sekenario Pengujian

Dalam mengevaluasi performa model, pengujian dikonstruksi melalui dua paradigma pelatihan yang dirancang untuk memvalidasi efektivitas transfer pengetahuan pada dataset terbatas (few-shot learning). Seluruh pengujian dijalankan pada resolusi 256x256 piksel dengan batch size 16 (untuk megamendung memakai batch size 64), serta mengaktifkan teknik augmentasi terdiferensiasi (differentiable augmentation) untuk memitigasi risiko overfitting pada diskriminatore.

- Skenario A (Pelatihan dari Nol)

Model dilatih menggunakan dataset Batik Parang (150 citra) dan dataset Batik megamendung(102 citra) dari titik nol selama 20.000 iterasi (steps). Skenario ini bertujuan untuk membangun baseline kemampuan model dalam mempelajari representasi fitur tekstur dasar, karakteristik kain, dan repetisi pola batik tanpa inisialisasi bobot sebelumnya.

- Skenario B (Transfer Learning)

Model dilatih menggunakan dataset target Batik Kawung (150 citra). Namun, alih-alih memulai dari nol, model memuat bobot (weights) yang telah dikonvergensi pada domain Batik Parang dari Skenario A sebagai inisialisasi awal. Pelatihan kemudian dilanjutkan hingga mencapai total 30.000 iterasi guna menguji efisiensi adaptasi domain dari pola miring (slope) ke pola geometris bulat.

b. Hasil Pengujian

Kualitas sintesis citra diukur secara kuantitatif menggunakan metrik Fréchet Inception Distance (FID), yang secara statistik menghitung kedekatan antara distribusi fitur citra asli dan citra artifisial. Skor FID yang lebih rendah mengindikasikan tingkat realism dan variasi yang lebih baik.

Tabel Evaluasi Kuantitatif (Skor FID)

Skenario	Dataset	Jumlah Steps	Skor FID Akhir
A	Batik Parang	20.000	293
A	Batik Megamendung	20.000	150
B	Batik Kawug	30.000(20.000 dari base batik parang + 10.000)	189

Tabel Perbandingan Gambar asli Dengan Sintetis

Tabel ini menyajikan perbandingan kualitatif antara citra batik asli dari dataset dan citra sintetis yang dihasilkan oleh model Lightweight GAN.

Batik Parang		
Batik Megamendung		

Batik Kawug		
	Asli	Sintetis

5. Analisa Hasil

Evaluasi performa model Lightweight GAN dalam menghasilkan motif batik sintetis dilakukan melalui analisis komparatif terhadap dua metrik utama: skor kuantitatif Fréchet Inception Distance (FID) dan inspeksi visual kualitatif. Secara keseluruhan, hasil eksperimen menunjukkan bahwa arsitektur yang ringan dengan strategi pelatihan yang tepat mampu mengatasi tantangan few-shot learning pada dataset batik yang sangat terbatas.

a. Dampak Ukuran Batch pada Kualitas Generasi (Kasus Batik Megamendung)

Pencapaian paling signifikan terlihat pada eksperimen dataset Batik Megamendung yang mencatatkan skor FID terbaik, yaitu 150. Meskipun dilatih dengan jumlah data paling sedikit (hanya 102 citra) dibandingkan motif lainnya, model ini justru menghasilkan distribusi fitur yang paling realistik. Faktor krusial di balik kesuksesan ini adalah penerapan batch size 64.

Berbeda dengan skenario Batik Parang yang menggunakan batch size 16 (FID 293), penggunaan batch size yang lebih besar pada Megamendung memberikan estimasi gradien yang jauh lebih stabil selama proses backpropagation. Stabilitas ini memungkinkan diskriminator untuk mempelajari variasi fitur secara lebih holistik dalam satu iterasi, mencegahnya terjebak pada local minima atau menghafal data (overfitting).

b. Efektivitas Transfer Learning pada Adaptasi Domain (Kasus Batik Kawung)

Eksperimen pada Batik Kawung (Skenario B) membuktikan efektivitas metode transfer learning. Dengan menginisialisasi bobot generator dari model yang telah terlatih pada Batik Parang, model Kawung berhasil mencapai skor FID 189, jauh lebih baik dibandingkan pelatihan dari nol pada Batik Parang (FID 293).

Hal ini menunjukkan bahwa meskipun terdapat pergeseran domain yang cukup ekstrem dari pola garis diagonal (Parang) ke pola geometris bulat (Kawung) pengetahuan dasar mengenai tekstur kain dan struktur repetitif yang telah dipelajari sebelumnya dapat ditransfer dengan baik. Fine-tuning memungkinkan model untuk langsung berfokus pada adaptasi bentuk motif tanpa harus mempelajari fitur dasar gambar dari awal, sehingga mempercepat konvergensi dan meningkatkan kualitas visual akhir meskipun jumlah iterasi total aktual hanya 10.000 langkah.

c. Tantangan pada Pola Geometris Kompleks (Kasus Batik Parang)

Sebaliknya, Batik Parang yang dilatih dari nol (Skenario A) menghasilkan skor FID tertinggi (293), yang mengindikasikan kualitas generasi terendah. Kendala ini kemungkinan besar disebabkan oleh kombinasi batch size kecil (16) dan kompleksitas motif Parang yang memiliki aturan geometris diagonal yang ketat (lereng). Tanpa bantuan pre-training atau ukuran batch yang memadai, model kesulitan menangkap konsistensi garis diagonal yang panjang dan presisi, yang sering kali terputus atau terdistorsi pada hasil sintesis.

d. Peran Arsitektur Lightweight GAN dan DiffAugment

Secara umum, kemampuan model untuk menghasilkan citra batik 256x256 yang koheren tanpa mengalami mode collapse meskipun data sangat terbatas memvalidasi keunggulan arsitektur Lightweight GAN. Modul Skip-Layer channel-wise Excitation (SLE) terbukti efektif memperkuat aliran gradien, memastikan detail tekstur pada resolusi tinggi tetap terjaga. Lebih lanjut, integrasi Differentiable Augmentation (DiffAugment) memainkan peran kunci dalam "memperkaya" dataset secara artifisial selama pelatihan, memaksa diskriminator untuk belajar fitur yang invariant terhadap transformasi, sehingga menjaga keseimbangan permainan min-max antara generator dan diskriminator.

6. Kesimpulan

Berdasarkan hasil perancangan, implementasi, dan pengujian sistem Batik Generator menggunakan arsitektur Lightweight GAN, dapat ditarik beberapa kesimpulan sebagai berikut:

a. Efektivitas Arsitektur Lightweight GAN pada Data Terbatas

Penelitian ini membuktikan bahwa arsitektur Lightweight GAN yang didukung oleh modul Skip-Layer channel-wise Excitation (SLE) dan teknik Differentiable Augmentation (DiffAugment) mampu menghasilkan citra batik sintetik resolusi tinggi (256x256 piksel) yang koheren meskipun dilatih dengan dataset yang sangat terbatas (kurang dari 150 citra per kelas). Mekanisme ini berhasil mencegah fenomena mode collapse dan menjaga stabilitas pelatihan GAN.

b. Dampak Signifikan Ukuran Batch (Batch Size)

Ukuran batch terbukti menjadi faktor krusial dalam kualitas akhir citra. Eksperimen pada Batik Megamendung dengan batch size 64 menghasilkan kualitas terbaik dengan skor FID terendah (150). Sebaliknya, penggunaan batch size kecil (16) pada Batik Parang menghasilkan skor FID tertinggi (293), karena diskriminator gagal mendapatkan estimasi gradien yang stabil untuk mempelajari pola geometris yang kompleks.

c. Keunggulan Metode Transfer Learning

Penerapan strategi Transfer Learning terbukti efektif dalam mempercepat konvergensi dan meningkatkan kualitas visual pada domain baru. Model Batik Kawung yang menggunakan inisialisasi bobot dari model Batik Parang berhasil mencapai skor FID 189, jauh lebih baik dibandingkan pelatihan dari nol pada Batik Parang, meskipun terdapat perbedaan karakteristik motif yang signifikan antara garis diagonal dan pola geometris bulat.

d. Implikasi Pelestarian Budaya Digital

Hasil penelitian ini menegaskan potensi besar sistem sebagai alat bantu kreatif dalam pelestarian batik digital. Fleksibilitas model dalam melakukan adaptasi lewat fine-tuning memungkinkannya menghasilkan ragam motif baru secara efisien, bahkan dengan sumber daya terbatas. Hal ini secara efektif menjawab tantangan klasik dalam implementasi AI

untuk budaya, yakni bagaimana mencapai performa optimal tanpa bergantung pada komputasi berat maupun dataset yang masif.

7. Referensi

- [1] H. Prasetyo and B. A. Putra Akardihas, "Batik image retrieval using convolutional neural network," *Telkomnika (Telecommunication Computing Electronics and Control)*, vol. 17, no. 6, pp. 3010–3018, Dec. 2019, doi: 10.12928/TELKOMNIKA.v17i6.12701.
- [2] A. Radford, L. Metz, and S. Chintala, "Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks," Jan. 2016, [Online]. Available: <http://arxiv.org/abs/1511.06434>
- [3] T. Karras NVIDIA and S. Laine NVIDIA, "A Style-Based Generator Architecture for Generative Adversarial Networks Timo Aila NVIDIA." [Online]. Available: <https://github.com/NVlabs/stylegan>
- [4] S. Zhao, Z. Liu, J. Lin, J.-Y. Zhu, and S. Han, "Differentiable Augmentation for Data-Efficient GAN Training," Dec. 2020, [Online]. Available: <http://arxiv.org/abs/2006.10738>
- [5] M. Heusel, H. Ramsauer, T. Unterthiner, B. Nessler, and S. Hochreiter, "GANs Trained by a Two Time-Scale Update Rule Converge to a Local Nash Equilibrium," Jan. 2018, [Online]. Available: <http://arxiv.org/abs/1706.08500>
- [6] B. Liu, Y. Zhu, K. Song, and A. Elgammal, "TOWARDS FASTER AND STABILIZED GAN TRAINING FOR HIGH-FIDELITY FEW-SHOT IMAGE SYNTHESIS." [Online]. Available: <https://github.com/odegeasslbc/FastGAN-pytorch>

8. Kontribusi & distribusi anggota kelompok

a. Wirawan Ariyoso

Mencari dataset, melatih model,membuat laporan & video

b. Yuafiq Alfin Al Khasan

Mencari dataset, melakukan EDA , melatih model

c. Adhitya Surya Handika

Melatih model, Membuat laporan , cek plagiatisme & AI

