

Implementasi Neural Style Transfer Menggunakan VGG19 untuk Menggabungkan Konten Lukisan Mona Lisa dengan Gaya Lukisan Picasso

1st Nero Caesar Suprobo

Informatics Engineering
Universitas Darussalam Gontor
Tuban, Indonesia

nerocaesarsuprobo76@student.cs.unida.
gontor.ac.id

2nd Muhammad Mishbahul Muflihini

Informatics Engineering
Universitas Darussalam Gontor
Jombang, Indonesia

muhammadmishbahulmuslihini10@stud
ent.cs.unida.gontor.ac.id

3rd Mhd Nurdin Al Kahfi

Informatics Engineering
Universitas Darussalam Gontor
line 4: Riau, Indonesia

mhdalkahfi19@student.cs.unida.gontor.
ac.id

4th Muhammad Haekal

Informatics Engineering
Universitas Darussalam Gontor
Palembang, Indonesia

Muhammadhaekal74@student.cs.unida.
gontor.ac.id

5th Muhammad Akmal Najib Gunawan

Informatics Engineering
Universitas Darussalam Gontor
Pontianak, Indonesia

muhammadnajibgunawan29@student.c
s.unida.gontor.ac.id

I. LATAR BELAKANG

Neural Style Transfer (NST) adalah salah satu teknik paling menarik dalam bidang *computer vision* yang memungkinkan pembuatan karya seni artistik baru dengan menggabungkan dua gambar: gambar konten (misalnya, sebuah foto) dan gambar gaya (misalnya, sebuah lukisan terkenal). Teknik ini pertama kali diperkenalkan oleh Leon A. Gatys, Alexander S. Ecker, dan Matthias Bethge. NST memanfaatkan *deep neural networks*, khususnya Convolutional Neural Networks (CNN), untuk memisahkan representasi konten dan gaya dari sebuah gambar. Dengan kemampuan ini, kita dapat mentransfer gaya artistik dari satu gambar ke gambar lainnya, menciptakan hasil visual yang menakjubkan.

Aplikasi dari Neural Style Transfer sangat luas, mulai dari pembuatan seni digital, filter foto di media sosial, hingga penggunaannya dalam industri film dan desain. Kemampuannya untuk meniru gaya seniman terkenal seperti Van Gogh, Picasso, atau Monet secara otomatis membuka pintu bagi kreativitas tak terbatas. Model VGG19, yang telah dilatih sebelumnya pada dataset ImageNet, terbukti sangat efektif untuk tugas ini karena lapisan-lapisan konvolusinya mampu menangkap fitur-fitur kompleks yang merepresentasikan konten dan tekstur (gaya).

II. PERUMUSAN MASALAH

Bagaimana cara membangun dan menerapkan model Neural Style Transfer menggunakan arsitektur VGG19 untuk menggabungkan konten dari lukisan Mona Lisa dengan gaya artistik dari lukisan karya Picasso?

III. TUJUAN

Tujuan dari proyek mini ini adalah:

1. Mengimplementasikan algoritma Neural Style Transfer menggunakan framework TensorFlow/Keras.
2. Memanfaatkan model VGG19 yang telah dilatih sebelumnya (*pre-trained*) untuk mengekstraksi fitur konten dan gaya.
3. Melakukan optimisasi untuk menghasilkan gambar baru yang mempertahankan konten dari gambar pertama dan mengadopsi gaya dari gambar kedua.
4. Menganalisis hasil secara visual dan memahami bagaimana model merepresentasikan konten dan gaya.

IV. METODOLOGI

Arsitektur Model: VGG19 Dalam proyek ini, kami menggunakan arsitektur VGG19, yang merupakan Convolutional Neural Network (CNN) dengan kedalaman 19 lapisan. VGG19 dipilih karena kemampuannya yang telah terbukti dalam mengekstraksi fitur hierarkis dari gambar. Lapisan-lapisan yang lebih awal pada VGG19 cenderung menangkap fitur-fitur dasar seperti tepi dan warna, sedangkan lapisan yang lebih dalam menangkap fitur-fitur yang lebih kompleks dan abstrak seperti bentuk objek utuh.

Untuk Neural Style Transfer, kita tidak melatih ulang model VGG19. Sebaliknya, kita menggunakannya sebagai *feature extractor* statis.:

1. Ekstraksi Fitur Konten: Representasi konten diekstraksi dari *feature maps* salah satu lapisan konvolusi di bagian tengah arsitektur (misalnya, `block5_conv2`). Lapisan ini dipilih karena mampu menangkap fitur tingkat tinggi dari objek utama tanpa terlalu detail pada tekstur.
2. Ekstraksi Fitur Gaya: Representasi gaya diekstraksi dari beberapa lapisan konvolusi, dari yang awal hingga yang lebih dalam (misalnya, `block1_conv1`, `block2_conv1`, `block3_conv1`, `block4_conv1`, `block5_conv1`). Untuk menangkap gaya, kita tidak menggunakan *feature maps* secara langsung, melainkan menghitung Matriks Gram dari *feature maps* tersebut. Matriks Gram mengukur korelasi antara fitur-fitur yang berbeda, yang secara efektif merepresentasikan tekstur, warna, dan pola visual tanpa memperhatikan tata letak spasial objeknya.

V. DATASET

Dataset untuk proyek ini sangat sederhana dan hanya terdiri dari dua gambar utama:.

1. Gambar Konten: Lukisan Mona Lisa karya Leonardo da Vinci. Gambar ini akan menjadi dasar struktur dan objek dari hasil akhir.
2. Gambar Gaya: Sebuah lukisan karya Pablo Picasso (misalnya, "The Weeping Woman" atau karya lain dalam gaya Kubisme). Gambar ini akan menjadi sumber tekstur, palet warna, dan gaya artistik.

VI. PREPROSESSING

Langkah-langkah preprocessing data yang dilakukan adalah:

- a. Memuat Gambar: Gambar konten dan gaya dimuat menggunakan Keras atau Pillow.
- b. Mengubah Ukuran (Resize): Kedua gambar diubah ukurannya agar memiliki dimensi yang seragam untuk memudahkan pemrosesan.
- c. Normalisasi: Nilai piksel gambar dinormalisasi sesuai dengan yang diharapkan oleh model VGG19. Ini biasanya melibatkan pengubahan rentang piksel [0, 255] menjadi rentang yang sesuai dan pemrosesan BGR (format yang digunakan VGG).
- d. Menambahkan Dimensi Batch: Gambar di-reshape agar memiliki dimensi batch, sesuai dengan format input Keras/TensorFlow (1, height, width, channels).

VII. PENDEKATAN PELATIHAN

Proses ini bukanlah "pelatihan" dalam artian tradisional karena kita tidak mengubah bobot dari model VGG19. Sebaliknya, kita melakukan optimisasi pada sebuah gambar input yang awalnya dibuat dari *noise* atau dari gambar konten itu sendiri. Tujuannya adalah untuk memodifikasi piksel-piksel gambar ini secara iteratif hingga *loss function*-nya minimal :

1. Loss Function: Total loss adalah gabungan dari dua komponen:
 - a. Content Loss: Dihitung sebagai *Mean Squared Error* (MSE) antara *feature maps* konten dari gambar yang dihasilkan dan *feature maps* konten dari gambar konten asli.
 - b. Style Loss: Dihitung sebagai MSE dari Matriks Gram pada setiap lapisan gaya yang dipilih antara gambar yang dihasilkan dan gambar gaya asli. *Style loss* dari setiap lapisan kemudian dijumlahkan (dengan bobot tertentu).
 - c. Total Loss = $\alpha * \text{content_loss} + \beta * \text{style_loss}$. Parameter alpha dan beta adalah bobot yang menentukan penekanan antara mempertahankan konten atau meniru gaya.
2. Optimizer: Optimizer seperti Adam atau L-BFGS digunakan untuk meminimalkan *total loss* dengan memperbarui piksel-piksel pada gambar yang dihasilkan. Proses ini diulang selama beberapa iterasi hingga hasilnya konvergen.

VIII. IMPLEMENTASI

1) Penjabaran Proses Training

Proses optimisasi dimulai dengan sebuah gambar input (misalnya, salinan dari gambar konten). Pada setiap iterasi:

- a. Gambar input dilewatkan melalui model VGG19 untuk mendapatkan *feature maps* pada lapisan konten dan gaya.
- b. *Content loss* dihitung dengan membandingkan fitur konten gambar input dengan fitur konten dari gambar Mona Lisa.
- c. *Style loss* dihitung dengan membandingkan Matriks Gram dari fitur gaya gambar input dengan Matriks Gram dari fitur gaya lukisan Picasso.
- d. *Total loss* dihitung sebagai kombinasi keduanya.
- e. Gradien dari *loss* terhadap piksel gambar input dihitung.
- f. Optimizer menggunakan gradien ini untuk memperbarui piksel-piksel gambar input. Langkah-langkah ini diulang untuk sejumlah iterasi yang telah ditentukan (misalnya, 1000-2000 iterasi) hingga gambar yang dihasilkan secara visual memuaskan.

2) Library yang Digunakan:

- TensorFlow & Keras: Untuk membangun dan menjalankan model, serta proses optimisasi. `tensorflow.keras.applications.VGG19` digunakan untuk memuat model *pre-trained*.
- NumPy: Untuk operasi numerik dan manipulasi array piksel.
- Pillow (PIL) atau OpenCV: Untuk memuat, memanipulasi, dan menyimpan gambar.
- Matplotlib: Untuk visualisasi gambar konten, gaya, dan hasil akhir.

IX. EKSPERIMEN

Eksperimen utama berfokus pada penyesuaian bobot α dan β untuk *loss function*.

- Bobot Konten (α) lebih tinggi: Hasilnya akan lebih mirip dengan lukisan Mona Lisa, dengan sedikit pengaruh gaya Picasso.
- Bobot Gaya (β) lebih tinggi: Hasilnya akan sangat abstrak dan penuh tekstur seperti lukisan Picasso, namun struktur wajah Mona Lisa mungkin hilang. Eksperimen juga dapat dilakukan dengan memilih lapisan yang berbeda untuk ekstraksi konten dan gaya, yang akan menghasilkan variasi pada hasil akhir.

X. TANTANGAN TEKNIS

- Kebutuhan Komputasi: Proses optimisasi sangat intensif secara komputasi dan memakan waktu lama, terutama untuk gambar beresolusi tinggi. Penggunaan GPU sangat disarankan.
- Tuning Hyperparameter: Menemukan keseimbangan yang tepat antara α dan β memerlukan banyak percobaan. Demikian pula, pemilihan lapisan gaya yang tepat dapat secara signifikan memengaruhi kualitas hasil.
- Konvergensi Loss: Memastikan *loss* konvergen ke hasil yang baik bisa menjadi tantangan. Terkadang, hasilnya bisa menjadi terlalu "berisik" atau artefak visual lainnya muncul.

XI. HASIL DAN EVALUASI

- Gambar Konten Asli: Lukisan Mona Lisa.
- Gambar Gaya Asli: Lukisan Picasso.

- Gambar Hasil: Gambar yang dihasilkan oleh model, yang menunjukkan wajah Mona Lisa dengan gaya Kubisme Picasso.



Gambar hasil akhir berhasil menggabungkan struktur wajah dan postur dari Mona Lisa dengan palet warna yang berani, goresan kuas yang patah-patah, dan distorsi geometris yang menjadi ciri khas gaya Picasso. Kontur wajah Mona Lisa masih dapat dikenali, namun tekstur kulit dan latar belakang telah sepenuhnya diubah menjadi gaya artistik yang baru.

Analisis Performa Model

Performa dalam konteks Neural Style Transfer bersifat kualitatif dan subjektif, tidak diukur dengan metrik akurasi seperti pada tugas klasifikasi. Evaluasi dilakukan secara visual:

- Seberapa baik konten dipertahankan? Apakah objek utama (Mona Lisa) masih jelas dan tidak terdistorsi secara berlebihan?
- Seberapa baik gaya ditransfer? Apakah tekstur, warna, dan pola dari lukisan Picasso berhasil diterapkan secara konsisten di seluruh gambar?
- Apakah ada artefak yang tidak diinginkan? Apakah hasilnya terlihat bersih atau terdapat *noise* dan area yang aneh?

Interpretasi

Hasil ini menunjukkan bahwa model VGG19 mampu secara efektif memisahkan representasi konten dari gaya. Dengan mendefinisikan *loss function* yang menargetkan kedua aspek ini secara terpisah, kita dapat "memaksa" sebuah gambar untuk memiliki konten dari satu sumber dan gaya dari sumber lain. Ini mengkonfirmasi hipotesis bahwa jaringan saraf dalam dapat mempelajari fitur-fitur gambar yang dapat dipisahkan dan dimanipulasi.

XII. KESIMPULAN DAN SARAN

Kesimpulan

Proyek ini berhasil mengimplementasikan Neural Style Transfer menggunakan model VGG19 untuk menggabungkan konten dari lukisan Mona Lisa dengan gaya

lukisan Picasso. Melalui proses optimisasi yang meminimalkan gabungan *content loss* dan *style loss*, model mampu menghasilkan karya seni baru yang secara visual koheren dan menarik. Proyek ini mendemonstrasikan kekuatan model *deep learning pre-trained* sebagai alat untuk tugas-tugas kreatif di luar tujuan pelatihan asli mereka.:

1. Kelebihan:

- Mampu menghasilkan gambar artistik berkualitas tinggi.
- Implementasi relatif mudah dipahami dengan *framework* modern seperti TensorFlow.
- Sangat fleksibel, dapat digunakan dengan kombinasi gambar konten dan gaya apa pun..

2. Kekurangan:

- Prosesnya lambat dan membutuhkan sumber daya komputasi yang besar.
- Hasil sangat bergantung pada pemilihan hyperparameter (α , β , lapisan gaya) yang memerlukan banyak *tuning* manual.
- Kualitas hasil bersifat subjektif.

3. Saran :

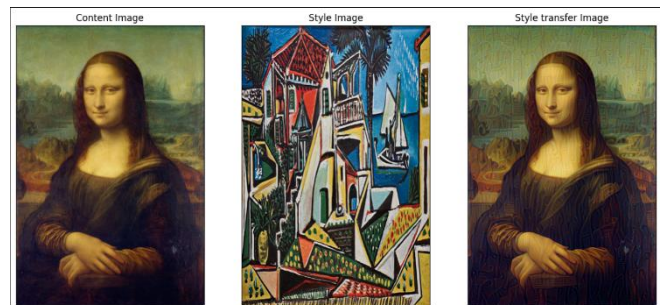
- Optimisasi Kecepatan: Mengimplementasikan metode yang lebih cepat seperti *Fast Neural Style Transfer* yang melatih *feed-forward network* terpisah untuk setiap gaya, memungkinkan transfer gaya secara *real-time* setelah pelatihan.
- Kontrol yang Lebih Baik: Mengeksplorasi teknik yang memungkinkan kontrol spasial atas transfer gaya (misalnya, menerapkan gaya hanya pada area latar belakang).
- Ekstensi ke Video: Menerapkan algoritma Neural Style Transfer pada setiap frame video untuk menciptakan video dengan gaya artistik, sambil memastikan konsistensi temporal antar-frame.
- Eksplorasi Arsitektur Lain: Mencoba model lain selain VGG19, seperti Inception atau ResNet, untuk melihat bagaimana arsitektur yang berbeda memengaruhi ekstraksi fitur dan hasil akhir.

XIII. LAMPIRAN

1. Link Github :

- <https://github.com/NEROCAESAR76/proyek-neural-style-transfer-monalisa-picasso.git>

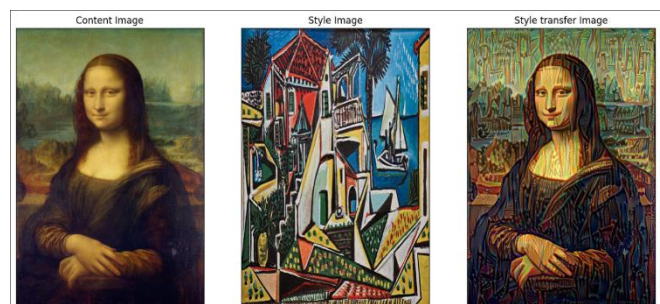
2. Hasil Training :



Epoch 1000



Epoch 5000



Epoch 10000