

# Neural Style Transfer Using VGG19

Departemen Ilmu Komputer dan Elektronika Universitas Gadjah Mada

1<sup>st</sup> Alleludo Panji Dewa Atmaja  
Departemen Ilmu Komputer dan  
Elektronika  
Universitas Gadjah Mada  
Yogyakarta, Indonesia  
alleludo.pan2003@mail.ugm.ac.id

2<sup>nd</sup> Fatha Marvin Yudhakusuma  
Departemen Ilmu Komputer dan  
Elektronika  
Universitas Gadjah Mada  
Yogyakarta, Indonesia  
fathamarvin@mail.ugm.ac.id

3<sup>rd</sup> Bagus Arif Arikusworo  
Departemen Ilmu Komputer dan  
Elektronika  
Universitas Gadjah Mada  
Yogyakarta, Indonesia  
bagus.arif.arikusworo@mail.ugm.ac.id

4<sup>th</sup> Muhammad Mahdi  
Departemen Ilmu Komputer dan  
Elektronika  
Universitas Gadjah Mada  
Yogyakarta, Indonesia  
muhammad.mahdi@mail.ugm.ac.id

**Abstrak**—Pembuatan seni melibatkan proses-proses rumit yang ditandai oleh abstraksi dan kebaruan. Untuk meningkatkan efisiensi produksi artistik, transfer gaya berbasis pembelajaran mesin telah menjadi terkenal dalam visi komputer. Namun, metode konvensional sering kali mengalami kesulitan dengan distorsi warna, kehilangan konten, dan durasi pemrosesan yang memanjang. Makalah ini memperkenalkan algoritma transfer gaya canggih yang memanfaatkan model jaringan saraf konvolusional VGG19. Jaringan komprehensif kami mengintegrasikan komponen transfer gaya dan jaringan kerugian. Melalui pelatihan, jaringan transfer gaya mendapatkan kemampuan untuk langsung memetakan gambar konten ke versi bergaya. Kerugian konten, kerugian gaya, dan Total kerugian, dihitung oleh jaringan kerugian, secara dinamis menyesuaikan bobot jaringan transfer gaya. Selain itu, kami memperkenalkan strategi pelatihan silang untuk dengan cermat menjaga detail gambar konten. Percobaan kami menunjukkan superioritas algoritma ini dibandingkan metode transfer gaya saraf klasik.

Melanjutkan karya Chang Zhou et al. (2019) dan Xin Wang et al. (2017), makalah ini memperkenalkan pendekatan baru untuk transfer gaya artistik dengan memanfaatkan jaringan saraf konvolusional VGG19. Dengan memanfaatkan fitur-fitur dalam VGG19, metode kami secara dinamis menggabungkan ulang gambar konten dan gaya. Melalui representasi berlapis, pendekatan kami mencapai kesetiaan visual yang luar biasa dalam mentransfer gaya gambar referensi ke gambar konten target. Hasil eksperimen menunjukkan efektivitas dan adaptabilitas metode kami di berbagai gaya seni dan jenis konten, menekankan potensinya untuk aplikasi waktu nyata dalam penciptaan seni digital dan peningkatan gambar.

**Kata Kunci**—*Transfer Gaya Artistik, VGG19, Pembelajaran Mesin Mendalam, Kesetiaan Visual*

## I. INTRODUCTION

Dalam beberapa tahun terakhir, Neural Style Transfer (NST) telah menjadi salah satu titik fokus utama dalam Penggabungan seni visual dan kecerdasan buatan. NST memungkinkan kita untuk mentransfer gaya seni dari satu gambar ke gambar lainnya, menghasilkan karya seni digital yang unik dan menggabungkan estetika visual dari berbagai sumber. Transfer gaya, yang melibatkan proses mengecat ulang foto yang ada dengan gaya lain, dianggap sebagai masalah yang menantang namun erat terkait dalam bidang

seni. Baru-baru ini, pendekatan ini telah menjadi topik yang aktif baik di dunia akademis maupun industri berkat beberapa penelitian terkini.

Salah satu terobosan penting dalam pengembangan Neural Style Transfer (NST) adalah penggunaan arsitektur jaringan saraf konvolusional (CNN) yang mendalam. Dalam konteks ini, VGG19, sebuah model CNN yang mendalam dengan 19 lapisan, telah menjadi pilihan populer karena kemampuannya untuk menangkap representasi tingkat tinggi dari gambar. Model ini dikembangkan oleh Visual Graphics Group (VGG) di Universitas Oxford, dan keunggulannya terletak pada struktur yang dalam dengan 19 lapisan, memungkinkannya mengekstraksi fitur-fitur kompleks dari gambar.

VGG19 mampu menangkap representasi hierarkis dari sebuah gambar melalui konvolusi beberapa lapisan, mulai dari fitur-fitur rendah seperti garis dan bentuk hingga fitur tingkat tinggi seperti tekstur dan pola yang lebih kompleks. Dalam pengaplikasiannya dalam Neural Style Transfer (NST), penggunaan VGG19 memberikan keunggulan yang signifikan dalam menggabungkan konten dan gaya, menghasilkan hasil yang lebih kaya secara visual.

Penerapan VGG19 dalam NST membuka pintu untuk eksplorasi lebih lanjut dalam menciptakan karya seni visual yang unik serta dalam pengembangan aplikasi analisis citra yang lebih maju. Dengan keahlian VGG19 dalam mengekstraksi fitur-fitur gambar, penelitian ini bertujuan untuk mendalami proses transfer gaya saraf dengan fokus pada model ini, dengan harapan dapat meningkatkan pemahaman tentang cara kerja teknik ini dan menggali potensi aplikatifnya yang lebih luas.

Dalam konteks transfer gaya saraf, VGG19 menjadi sangat berguna karena kemampuannya dalam memahami struktur fitur-fitur kompleks dari gambar. Proses transfer gaya ini melibatkan dua gambar utama: gambar sumber (yang akan digunakan untuk mengekstraksi gaya) dan gambar target (yang akan dimodifikasi sesuai dengan gaya gambar sumber). Melalui representasi yang diperoleh dari VGG19, kita dapat mengidentifikasi korelasi antara fitur-fitur dari kedua gambar tersebut dan mengadaptasi gaya dari gambar sumber ke gambar target dengan presisi yang tinggi.

Penerapan transfer gaya saraf menggunakan model VGG19 memiliki implikasi yang luas. Di bidang seni digital, ini memungkinkan seniman untuk menciptakan karya-karya yang unik dengan menggabungkan gaya dari berbagai karya seni visual. Selain itu, dalam analisis citra medis, teknik ini dapat membantu dalam peningkatan visualisasi data medis yang kompleks atau dalam penyesuaian gambar radiologi untuk membantu diagnosis.

Penelitian yang mendalam tentang konsep transfer gaya saraf dengan menggunakan VGG19 terus berkembang. Tujuan dari penelitian ini adalah untuk lebih memahami representasi fitur-fitur yang diekstraksi oleh model VGG19 dan bagaimana kita dapat menggunakan informasi ini untuk meningkatkan keakuratan dan efisiensi dalam proses transfer gaya saraf.

Dengan demikian, pada penelitian kali ini bertujuan untuk mengeksplorasi konsep transfer gaya saraf dengan fokus pada model VGG19, menyoroti potensi aplikasi dalam berbagai bidang dan mendalaminya lebih lanjut untuk mengembangkan pemahaman yang lebih baik tentang proses ini. Diharapkan penelitian ini akan memberikan kontribusi yang signifikan dalam pengembangan teknik transfer gaya saraf yang lebih canggih dan aplikatif di masa depan.

## II. RELATED WORKS

Pada penelitian Neural Style Transfer kami kali ini mengacu kepada research paper berjudul Multimodal Transfer: A Hierarchical Deep Convolutional Neural Network for Fast Artistic Style Transfer oleh Xin Wang et al dan juga An Improved Style Transfer Algorithm Using Feedforward Neural Network for Real-Time Image Conversion oleh Chang Zhou et al.

Pada research paper Multimodal Transfer: A Hierarchical Deep Convolutional Neural Network for Fast Artistic Style Transfer mengusulkan jaringan saraf tiruan multimodal convolutional mempertimbangkan saluran warna dan pencahayaan untuk representasi yang lebih sesuai. Jaringan ini juga melakukan stylization secara hirarkis dengan beberapa losses dengan skala yang meningkat. Arsitektur ini akan terdiri dari 2 komponen yaitu Feed-forward Multimodal Network dan Loss Network. Lalu di dalam multimodal network, terbagi lagi menjadi 3 subnetwork yaitu style subnet, enhance subnet and refine subnet. Hasil yang didapat dari research tersebut adalah sebuah model lebih dapat menampilkan brushwork pada gambar yang sudah di stylized dan sekaligus runtime untuk testing lebih cepat dan penggunaan memori lebih sedikit. Selain itu, ternyata modelnya juga dapat melakukan styling pada style yang berbeda walaupun input image sama.

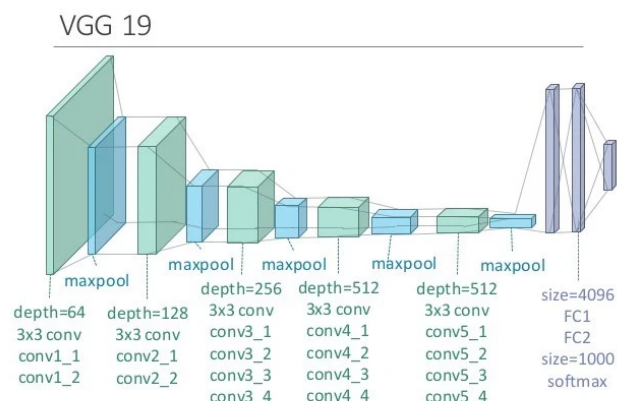
Lalu pada research paper An Improved Style Transfer Algorithm Using Feedforward Neural Network for Real-Time Image Conversion mengusulkan penggunaan jaringan syaraf feedforward dengan biaya rendah dan dapat mengatasi masalah anamorfosis warna dan kehilangan konten. Algoritma ini terdiri dari 2 bagian yaitu jaringan transfer gaya yang bertugas dalam memetakan gambar konten ke dalam gambar bergaya setelah pelatihan dan juga jaringan kerugian yang mencakup Kehilangan konten, kehilangan gaya, dan kehilangan Total Variasi (TV) yang akan dihitung oleh jaringan kerugian untuk memperbarui bobot jaringan transfer gaya. Dari penelitian ini didapatkan sebuah Convolution Neural Network memiliki hasil yang

bagus untuk ekstraksi fitur dan mampu menyediakan cara alternatif untuk perbandingan fitur antara gambar yang berbeda. Keuntungan dari penggunaan algoritma pada penelitian adalah waktunya yang singkat dalam proses stylization. Keunggulan yang lain adalah dapat memisahkan target yang penting dari background untuk menghindari hilangnya konten.

## III. METODOLOGI

Before you begin to format your paper, first write and save the content as a separate text file. Complete all content and organizational editing before formatting. Please note sections A-D below for more information on proofreading, spelling and grammar.

### A. Arsitektur



VGG19 merupakan model jaringan saraf konvolusional yang dikembangkan oleh Universitas Oxford, memiliki arsitektur yang relatif sederhana namun efektif. Ada total 16 lapisan konvolusional, 3 lapisan penggabungan, dan 3 lapisan terhubung penuh, dengan masing-masing lapisan konvolusional menggunakan filter 3x3 dengan ReLU diaktifkan. Lapisan pooling menggunakan operasi max-pooling dengan filter 2x2 untuk mengurangi dimensi gambar.

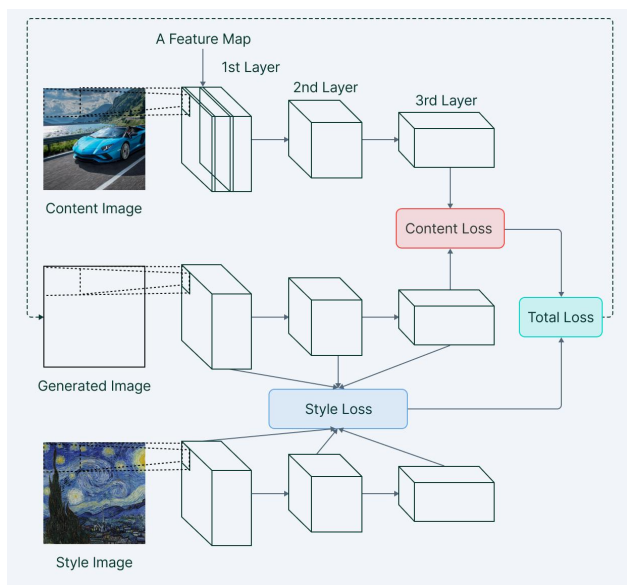
Lapisan yang terhubung sepenuhnya diikuti oleh lapisan keluaran yang berisi 1000 neuron, sesuai dengan jumlah kelas dalam kumpulan data ImageNet. Meskipun arsitektur ini mungkin tampak sederhana, penggunaan filter kecil dan aktivasi ReLU di setiap lapisan memungkinkan jaringan ini memahami dan mengekstrak fitur kompleks dari gambar. Operasi VGG19 biasanya melibatkan langkah konvolusi untuk mengekstrak fitur gambar di setiap lapisan. Proses ini diikuti dengan penggabungan untuk mengurangi dimensi dan lapisan yang terhubung sepenuhnya untuk klasifikasi akhir. Dengan menggunakan arsitektur terpadu, VGG19 dapat mempelajari representasi fitur yang baik dari data gambar. Meskipun arsitektur ini sederhana dibandingkan dengan beberapa arsitektur canggih, arsitektur ini telah terbukti berhasil dalam berbagai tugas pemrosesan gambar dan visi komputer.

### B. Flow Chart

Model arsitektur VGG19 meliputi penyediaan gambar berupa input yang kemudian akan dilakukan normalisasi gambar dengan menyesuaikan ukuran dari gambar tersebut sesuai kebutuhan. Gambar yang sudah dinormalisasi akan dilakukan ekstraksi fitur pada setiap layernya, Langkah ini

dilakukan sesuai dengan berapa layer yang ditentukan untuk mendapatkan informasi gambar yang selanjutnya akan dilakukan perhitungan content loss.

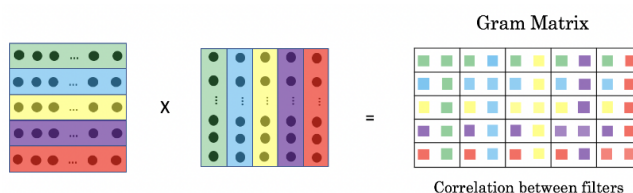
Hal serupa juga dilakukan pada input style untuk mendapatkan informasi gambar yang selanjutnya akan dilakukan perhitungan style loss. Hasil content loss dan style loss akan ditambahkan bersama untuk membentuk total loss, total loss ini kemudian digunakan sebagai fungsi objektif yang akan dioptimasi selama proses pelatihan menggunakan adam optimizer untuk meminimalkan total loss, sehingga model output yang diberikan dapat lebih baik dengan referensi style gambar yang diinginkan.



### C. Gram Matrix

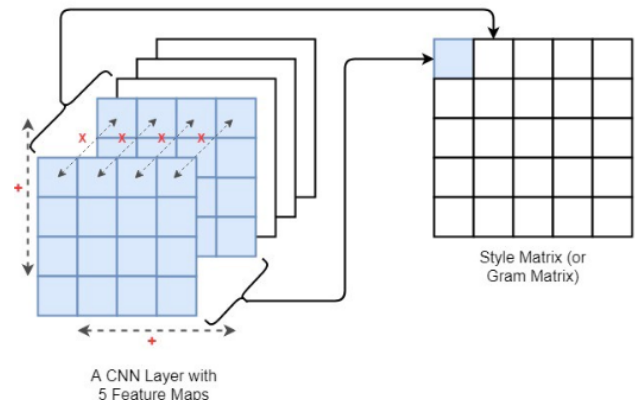
Gram matrix adalah representasi korelasi antarfitur dalam gambar, penting dalam Neural Style Transfer (NST). Dihasilkan dari produk dot fitur-fitur gambar dengan transposenya, matrix ini menangkap pola tekstur, hubungan spasial antarfitur, dan gaya visual dari gambar. Dalam NST, gram matrix memisahkan informasi gaya dan konten dari gambar sumber, memungkinkan transfer gaya yang presisi ke gambar target. Penggunaannya memungkinkan penciptaan karya seni visual baru yang memadukan estetika satu gambar dengan konten gambar lainnya, memanfaatkan representasi statistik untuk menciptakan hasil yang menarik secara visual.

Jadi untuk gaya komputasi biaya kita tidak boleh membandingkan aktivasi secara langsung. sebaliknya kita akan mendapatkan tensor baru yang berisi gaya aktivasi dan kita dapat melakukan penghitungan biaya yang serupa dengan biaya konten. Sehingga tensor baru tersebut tidak lain hanyalah matriks gram.



### D. Style Loss

Disini Style loss menggunakan matrix gram. Matriks Gram mengukur distribusi fitur pada peta fitur dalam sebuah lapisan, meminimalkan hilangnya gaya dengan mempertahankan tingkat distribusi fitur yang sama pada gaya dan gambar yang dihasilkan.



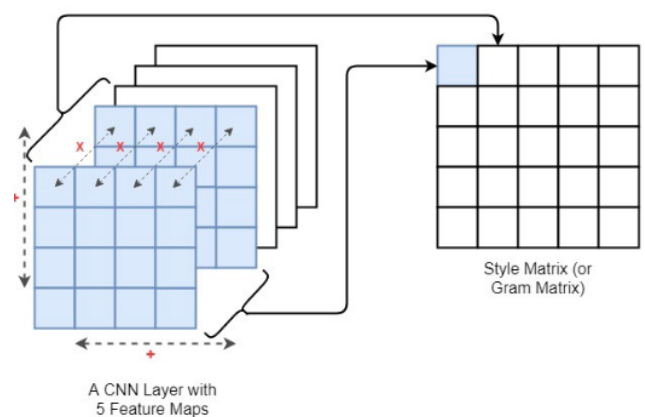
### E. Total Cost

$$L = \alpha L_{content} + \beta L_{style}$$

Loss gaya komputasi sedikit berbeda. Loss konten memberikan hasil matematis yang menunjukkan betapa miripnya kedua gambar ini. Semakin mirip berarti semakin menuju nol

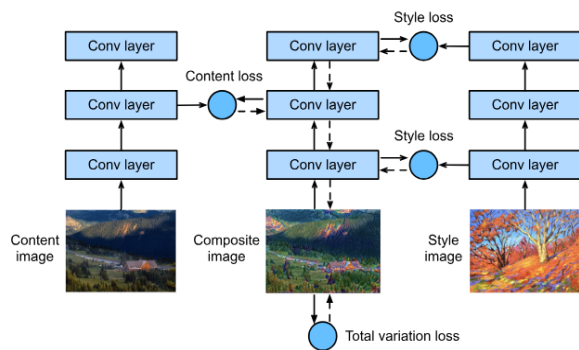
Namun gaya komputasi memberikan hasil matematis yang menunjukkan betapa miripnya fitur atau gaya dibandingkan kontennya.

Itu sebabnya total loss memberikan kesalahan dengan menggunakan kedua loss tersebut. Loss konten menghitung untuk menentukan seberapa mirip gambar-gambar itu, apa pun warnanya, gaya Loss gaya menghitung untuk menentukan seberapa mirip gaya gambar satu sama lain daripada konten. Misalnya. Gaya font a dan b --> loss gaya huruf a dan b --> loss konten



### E. Optimasi





Optimasi Adam merupakan suatu metode optimisasi yang digunakan untuk mengubah citra input sedemikian rupa sehingga mencerminkan gaya (style) suatu citra referensi, sambil mempertahankan kontennya. Adam adalah sebuah algoritma optimisasi stokastik yang menggabungkan konsep momentum dari metode gradient stokastik (SGD) dengan perhitungan adaptif dari learning rate untuk setiap parameter. Dengan memanfaatkan momentum, Adam dapat menangani permukaan fungsi loss yang berlekuk-lekuk dengan lebih baik, sementara penggunaan learning rate yang adaptif membantu konvergensi yang lebih cepat dan stabil. Optimasi ini digunakan untuk mengoptimalkan citra input dengan menyesuaikan piksel-piksennya sedemikian rupa sehingga loss style dan loss konten mencapai nilai minimum, sehingga citra tersebut mendekati gaya citra referensi dengan mempertahankan informasi kontennya.

#### IV. RESULT AND DISCUSSION

Gambar yang diuji sebanyak 10 gambar berbeda dengan 1 style gambar, mendapatkan hasil yang cukup baik untuk setiap gambar memperoleh waktu yang relatif cukup cepat dalam 1000 epoch. Berdasarkan hasil tersebut output yang diberikan cukup memuaskan, berikut salah satu contoh input dan output Neural Transfer Style.



Gambar input



Gambar style-input



Gambar output

#### V. CONCLUSION

Secara kesimpulan, model dapat berjalan dengan lancar setelah dilakukan training. Diketahui bahwa model masih dapat diperbaiki secara lebih dengan menambahkan jumlah training. Selain itu, untuk menambahkan iterasi training serta menambahkan optimizer untuk mendapatkan hasil yang lebih bagus.

- [1] Wang, Xin, et al. "Multimodal Transfer: A Hierarchical Deep Convolutional Neural Network for Fast Artistic Style Transfer." 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, doi:10.1109/cvpr.2017.759.
- [2] Zhou, et al. "An Improved Style Transfer Algorithm Using Feedforward Neural Network for Real-Time Image Conversion." Sustainability, vol. 11, no. 20, 2019, p. 5673, doi:10.3390/su11205673.