Kajian Peningkatan Kualitas Citra Wajah Menggunakan Metode Deep Learning Vanilla Generative Adversarial Network (VGAN) dan Super Resolution Generative Adversarial Network (SRGAN)

1st Satriyo Sakti Wicaksono
School of Electrical Engeneering
Telkom University
Bandung, Indonesia
satriyosaktiww@student.telkomuniversity.ac.id

2nd Irma Safitri
School of Electrical Engineering
Telkom University
Bandung, Indonesia
irmasaf@student.telkomuniversity.ac.id

3rd Rustam
School of Electrical Engineering
Telkom University
Bandung, Indonesia
rustamtelu@telkomuniveristy.ac.id

A. ABSTRAK

Citra dengan resolusi rendah pada citra digital dapat membuat detail gambar kurang jelas. Hal ini dapat disebabkan adanya degradasi warna, *blur* (buram) atau pun noise sehingga secara visual citra menjadi tidak terlihat jelas. Selain itu, resolusi rendah dapat berpengaruh pada citra yang dipakai dalam face recognition yang menyebabkan kinerja deteksi kurang baik. Oleh karena itu, restorasi resolusi citra diperlukan untuk mengatasi masalah tersebut.

Pada tugas akhir ini, digunakan salah satu perbaikan citra yaitu metode VGAN (Vanilla Generative Adversarial Network) yang disisipkan downsampling dan upsampling pada layer strukturnya dan telah diuji dan dibandingkan dengan metode SRGAN (Super Resolution Generative Adversarial Network). VGAN dan SRGAN memiliki kapasitas untuk melakukan perbaikan citra resolusi rendah menjadi citra dengan resolusi tinggi.

Dataset yang digunakan adalah CelebA - HQ (Celeb Faces Attributes High Quality) terdiri dari 1000 citra wajah. Hasil akhir menunjukkan bahwa metode VGAN modifikasi mendapatkan nilai PSNR tertinggi sebesar 31.82 dB dan SSIM sebesar 0.91. Sementara itu, metode SRGAN modifikasi mendapatkan nilai PSNR tertingggi sebesar 33.9 dB dan SSIM sebesar 0.923. Berdasarkan pengujian pada dataset, dapat disimpulkan bahwa metode SRGAN lebih unggul dibandingkan VGAN dalam melakukan pengujian menggunakan dataset CelebA – HQ. Hasil tersebut menunjukkan bahwa SRGAN mampu melakukan perbaikan citra dengan baik.

Kata Kunci: Citra Wajah, Resolusi, SRGAN, VGAN

I. PENDAHULUAN

Isi Citra digital Citra digital dalam kehidupan sehari-hari sudah menjadi hal yang lumrah. Citra digital sering digunakan dalam teknologi sehari-hari, seperti *smartphone*, dll. Kebiasaan membagikan foto dan momen smartphone melalui media sosial merupakan salah satu contoh penggunaan gambar digital. Citra digital juga telah menjadi format umum untuk menyimpan foto dan gambar, sebuah teknologi yang sangat dekat dengan kehidupan masyarakat di era teknologi dan informasi. Di era teknologi dan informasi saat ini, penggunaan citra digital sudah menjadi hal yang sangat diperlukan. Faktanya, keberadaan bidang penelitian pemrosesan gambar khusus membuktikan kebutuhan akan gambar digital. Selain itu, bermunculan berbagai teknologi yang menggunakan gambar, seperti: Pengenalan wajah, pengenalan objek, dll. Beberapa teknik tersebut memerlukan gambar sebagai masukan untuk diproses lebih lanjut guna memperoleh informasi yang dibutuhkan. Misalnya, dalam pengenalan wajah, gambar masukan diproses untuk mengenali siapa pemilik wajah dalam gambar tersebut. Teknologi pengenalan wajah merupakan teknologi yang mengambil gambar wajah sebagai masukan. Teknologi ini menggunakan pendekatan biometrik untuk memverifikasi atau mengenali identitas orang yang hidup berdasarkan fitur fisiologis wajah [1]. Pengenalan wajah memudahkan Anda mengidentifikasi wajah orang dalam gambar yang diambil dengan kamera Anda.

Face recognition banyak digunakan di berbagai sistem seperti sistem pengawasan. Namun sistem pengawasan biasanya menggunakan kamera CCTV dengan bidang pandang yang luas. Saat menggunakan kamera seperti itu, gambar wajah sering kali buram atau tidak jelas. Masalah ini terjadi karena rendahnya resolusi gambar wajah, .Chen dkk. mengidentifikasi hal ini sebagai tantangan dalam penerapan pengenalan wajah di sistem pengawasan dalam studi tahun 2018 mereka. Salah satu cara untuk mengatasi masalah ini adalah dengan meningkatkan resolusi gambar, karena resolusi yang rendah dapat mempengaruhi kinerja pengenalan wajah [2].

Gambar resolusi tinggi (*high resolution image*) memberikan informasi lebih detail, sehingga analisis gambar menjadi

lebih akurat. Misalnya gambar medis resolusi tinggi akan membantu dokter untuk membuat yang tepat diagnosis. Gambar dapat diperoleh dengan menggunakan alat perekam. Outputnya bisa berupa foto, video dan digital yang dapat langsung disimpan di dalamnya pita magnetiknya. Gambar sering kali ada penurunan mutu, misalnya mengandung cacat atau berisik, warnanya terlalu kontras, kurang tajam, *blur* dan tingkat resolusi rendah [3].

Salah satu cara untuk meningkatkan resolusi suatu gambar adalah dengan menggunakan GAN (*Generative adversarial network*). GAN adalah pendekatan untuk pemodelan generatif menggunakan metode pembelajaran mendalam, seperti jaringan saraf convolutional. yang bertujuan untuk membentuk atau membuat suatu data yang benar-benar baru berdasarkan data latih atau data yang sudah dilihat sebelumnya, biasa di implementasikan pada data citra yang berupa pixel [4].

Pada tahun 2017 Jaemin Son, Sang Jun Park, dan Kyu-Hwan Jung meneliti dan menyempurnakan gambar *super resolution* berjudul *Retinal Vessel Segmentation in Fundoscopic* penelitian pada retina manusia menggunakan metode VGAN in Tensorflow. Metode ini menghasilkan hasil yang lebih akurat, detail dan memiliki *false positive* yang lebih sedikit dibandingkan dengan metode sebelumnya [5]. Hasil penelitian tersebut VGAN mendapatkan nilai ROC (*Receiving Operating Characteristic*) sebesar 0.9803 [5]. Metode ini merupakan metode yang menunjukkan hasil yang lebih baik daripada metode yang lain.

Selanjutnya, Bagus Hardiansyah dkk. Menggunakan metode interpolasi *bicubic* pada *dataset* citra *deep learning.* Hasil penelitian membuktikan metode interpolasi *bicubic* dapat memperolah nilai rata-rata PSNR yang lebih tinggi pada perbesaran 3 kali dibanding pada perbesaran 4 kali [6]. Akan tetapi, pada penelitian serupa yang menggunakan metode interpolasi [7][8] belum menghasilkan kinerja yang memuaskan untuk mengatasi masalah resolusi citra. Hal ini dibuktikan dengan nilai PSNR yang lebih rendah dibandingkan dengan metode *deep learning* seperti SRGAN [9] dan CNN [10]. Seperti pada penelitian Yudai Nagano dan Yohei Kikuta yang menerapkan metode SRGAN pada citra makanan. Hasil pada penelitian tersebut dapat memberikan Xception Score sebesar 0.34 untuk model yang dilatih menggunakan data roti [11]. Selain itu, hasil pengujian menunjukkan metode ini memiliki kinerja lebih baik dibanding metode interpolasi [12].

Berdasarkan uraian di atas, pada tahun 2017 telah dilakukan penelitian oleh Christian Ledig, Lucas Theis, Ferenc Huszar dkk. Berjudul *Photo Realistic Single Image Super Resolution* menggunakan *Generative Adversarial Network* yang memperbaiki citra digital dalam bentuk super resolusi menggunakan metode *Super Resolution Generative Adversarial Network* (SRGAN) yang mengevaluasi PSNR. Bahwa terdapat beberapa keterbatasan pada resolusi gambar *PSNR Focused image resolution* yang pada akhirnya menggunakan metode SRGAN yang menambahkan *content loss function* dengan pelatihan menggunakan *Generative Adversarial Network* (GAN) [13]. Saat menggunakan uji *mean opinion score* (MOS), ditemukan bahwa hasil rekonstruksi SRGAN untuk faktor skala besar meningkat 4 kali lipat, dengan margin yang cukup besar, dan lebih fotorealistik dibandingkan rekonstruksi yang diperoleh dengan *state of the art reference method*.

Oleh karena itu, tugas akhir ini akan dilakukan suatu modifikasi layer terhadap SRGAN dan VGAN. Layer pada dua metode tersebut akan di lakukan *downsampling* terlebih dahulu dan di lakukan *upsampling* dengan modifikasi ini, diharapkan model tersebut memiliki nilai PSNR dan SSIM yang lebih baik daripada model SRGAN dan VGAN yang asli.

II. KAJIAN TEORI

A. Super Resolution

Citra super resolusi adalah salah satu teknik untuk mendapatkan citra yang beresolusi tinggi dari sekumpulan citra yang beresolusi rendah. Resolusi tinggi yang dihasilkan dapat berupa citra tunggal atau lebih. Citra resolusi tinggi didapat dari sekumpulan resolusi rendah yang diambil dari *scene (*adegan) yang sama. Karena dari *scene* yang sama akan menyediakan informasi yang mungkin dapat digunakan untuk merekonstruksi citra resolusi tinggi [14]. Berdasarkan *output* yang dihasilkan (*high resolution*), super resolusi dibedakan menjadi 2, yaitu super resolusi statis dan super resolusi dinamis. Super resolusi statis adalah metode super resolusi yang menghasilkan citra keluaran resolusi tinggi tunggal dan super resolusi dinamis adalah metode super resolusi yang menghasilkan citra keluaran resolusi tinggi yang lebih dari satu [15].

Super Resolution (SR) telah banyak dikembangkan seiring dengan meningkatnya kebutuhan akan citra digital berkualitas tinggi. Wang dkk.[16] menjelaskan bahwa secara prinsip, SR bertujuan merekonstruksi detail spasial yang hilang akibat keterbatasan sensor dan proses pengambilan sampel (*sampling*). Teknik ini awalnya berkembang pada pendekatan *Multi Image Super Resolution* (MISR) yang memanfaatkan informasi spasial dan temporal dari beberapa citra beresolusi rendah yang diambil dari adegan yang sama dengan pergeseran subpiksel. Informasi tambahan ini kemudian digabungkan melalui proses registrasi, interpolasi, dan rekonstruksi untuk menghasilkan satu citra resolusi tinggi.

Seiring perkembangan teknologi komputasi, pendekatan *Single Image Super Resolution* (SISR) semakin populer karena hanya memerlukan satu citra masukan. Pendekatan SISR memanfaatkan pola internal citra, seperti pengulangan tekstur atau struktur, yang dipelajari melalui metode interpolasi adaptif dan pembelajaran mendalam (*deep learning*). Penelitian terbaru banyak memfokuskan pada pemanfaatan jaringan saraf konvolusional (*CNN*) dan *Generative Adversarial Network* (GAN) untuk meningkatkan akurasi detail tekstur. Hal ini terbukti lebih unggul dibandingkan metode konvensional seperti interpolasi *bilinear* atau *bicubic* yang cenderung menghasilkan citra buram. Dengan demikian, SR menjadi salah satu topik penting dalam pengolahan citra digital modern karena aplikasinya sangat luas, mulai dari citra medis, satelit, pengawasan video (*surveillance*), hingga rekonstruksi citra wajah berkualitas tinggi [16].

B. Deep Learning

Deep learning adalah salah satu bidang *machine learning* yang memanfaatkan banyak *layer* pengolahan informasi nonlinier untuk melakukan ekstraksi fitur, pengenalan pola, dan klasifikasi [17]. Deep learning merupakan sebuah pendekatan dalam penyelesaian masalah

pada sistem pembelajaran komputer yang menggunakan konsep hierarki. Konsep hierarki membuat komputer mampu mempelajari konsep yang kompleks dengan menggabungkan dari konsep-konsep yang lebih sederhana. Jika digambarkan sebuah graf bagaimana konsep tersebut dibangun di atas konsep yang lain, graf ini akan dalam dengan banyak *layer*, hal tersebut menjadi alasan disebut sebagai *deep learning* [18].

Untuk memecahkan permasalahan yang kompleks, para peneliti mulai mengembangkan *deep learning.* Deep learning merupakan subbidang dari *machine learning* yang dirancang khusus untuk menangani data berukuran besar. Meskipun *deep learning* juga dapat diterapkan pada data berukuran kecil, pada kasus tersebut umumnya lebih efisien menggunakan algoritma *machine learning* tradisional. Yann LeCun dan peneliti lainnya dalam jurnal [19] menjelaskan bahwa *deep learning* sangat cocok digunakan untuk *big data*. Metode *machine learning* tradisional sering kali mengalami kesulitan dalam menangkap pola dan ketergantungan kompleks pada kumpulan *big data* karena umumnya masih bergantung pada rekayasa fitur manual dan arsitektur yang dangkal. Sebaliknya, *deep learning* menggunakan banyak lapisan jaringan saraf tiruan untuk mempelajari representasi fitur secara hierarkis langsung dari data, sehingga mampu mengekstraksi fitur penting secara otomatis dan menangani kerumitan *big data* dengan lebih efektif.

Secara mendasar, *deep learning* merupakan pendekatan pembelajaran mesin (*machine learning*) yang mengandalkan jaringan saraf tiruan (*Artificial Neural Network*) dengan banyak lapisan tersembunyi (*hidden layers*). Goodfellow dkk.[18] menjelaskan bahwa keunggulan utama *deep learning* terletak pada kemampuannya dalam mempelajari representasi data secara hierarkis, mulai dari fitur sederhana di lapisan awal hingga fitur kompleks di lapisan yang lebih dalam. Hierarki ini membuat *deep learning* mampu menangani data berukuran besar dengan pola yang sangat kompleks, seperti citra digital resolusi tinggi.

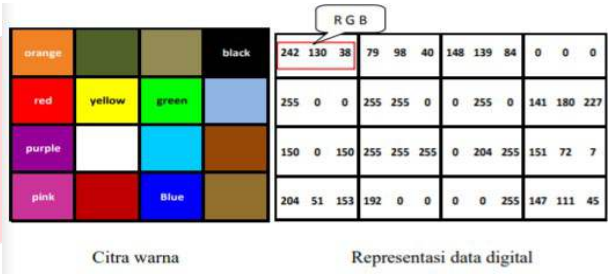
Dalam konteks pengolahan citra, *deep learning* banyak diimplementasikan pada arsitektur *Convolutional Neural Network* (CNN) untuk mendeteksi, mengenali, dan merekonstruksi pola visual. Salah satu terobosan penting adalah pengembangan *Generative Adversarial Network* (GAN), yang memanfaatkan prinsip persaingan dua jaringan: *generator* yang menghasilkan data tiruan, dan *discriminator* yang berfungsi membedakan data tiruan dengan data asli. Model ini terbukti efektif untuk menghasilkan citra realistis yang berkualitas tinggi. Pendekatan ini menjadi sangat relevan ketika diimplementasikan pada metode peningkatan resolusi citra, karena *deep learning* dapat secara otomatis mengekstraksi fitur - fitur penting dari data tanpa memerlukan rekayasa fitur (*feature engineering*) secara manual [18].

C. Citra *RGB*

Citra RGB merupakan citra yang nilai pikselnya disusun dengan warna merah, hijau, dan biru yang dijadikan patokan sebagai warna *universal*. Citra RGB mengandung matriks data berukuran M×N×3 mepresentasikan warna merah, hijau dan biru pada setiap pikselnya. Setiap warna mempunyai rentang nilai 0 sampai 255 dapat menghasilkan total warna sebanyak 16 juta warna.

GAMBAR 1
Citra RGB [20]

Kombinasi warna untuk citra dengan format bmp adalah 14-bit, atau lebih daari 16 juta warna yang dinamakan *true color.*



GAMBAR 2
Representasi citra warna data digital [21]

Menurut Gonzalez dan Woods [21], citra RGB adalah representasi citra digital berbasis warna dengan model warna aditif (*additive color model*). Warna pada citra RGB dihasilkan dari kombinasi intensitas tiga kanal warna dasar, yaitu merah (*Red*), hijau (*Green*), dan biru (*Blue*). Setiap piksel direpresentasikan oleh tiga nilai intensitas dengan rentang 0–255 pada format 8-bit per kanal, sehingga memungkinkan terciptanya variasi warna hingga lebih dari 16 juta warna (*True Color*).

Dalam sistem komputer dan perangkat digital seperti kamera, scanner, dan monitor, model RGB digunakan sebagai standar representasi warna karena kemampuannya meniru cara kerja penglihatan manusia yang sensitif terhadap spektrum merah, hijau, dan biru. Oleh karena itu, hampir seluruh data citra digital yang dihasilkan dalam kehidupan sehari-hari, termasuk citra wajah, direkam dalam format RGB. Informasi warna RGB sering diolah lebih lanjut dalam domain pengolahan citra digital untuk berbagai keperluan, seperti segmentasi warna, deteksi objek, hingga rekonstruksi resolusi tinggi [21].

D. Citra *Grayscale*

Citra *grayscale* merupakan nilai piksel derajat keabuan yang terdiri dari warna hitam, keabuan dan putih. Tingkat keabuan dimulai dari tingkatan hitam mendekati putih dengan nilai intensitas piksel bernilai tunggal. Citra *grayscale* mempunyai warna lebih banyak daripada citra biner. Citra *grayscale* disimpan pada format 8 bit pada setiap pikselnya mampu menghasilkan kemungkinan sebanyak 256 kali.

GAMBAR 3
Citra *Grayscale* nilai piksel 0 – 255 [20]

Citra *grayscale* memiliki warna hitam (minimum) dan putih (maksimum). Sehingga untuk skala 4 bit, maka kemungkinan nilainya adalah 24 = 16 memiliki 16 warna dan nilai maksimum adalah 24 − 1 = 15, maka warna 0(min) sampai 15(maks). Sedangkan untuk skala 8 bit, kemungkinan nilainya adalah 28 = 256 memiliki 256 warna, nilai maksimumnya 28 − 1 = 255, maka warna 0 (min) sampai 255 (maks).

GAMBAR 4
Pallet Grayscale

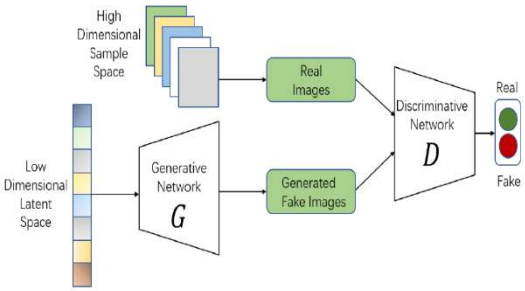
Citra *grayscale* banyak digunakan dalam tahap prapengolahan (*preprocessing*) citra digital karena strukturnya yang lebih sederhana dibandingkan citra berwarna, tetapi tetap memuat informasi penting seperti tepi objek, pola intensitas, dan tekstur. Dalam praktiknya, citra RGB sering dikonversi ke *format grayscale* untuk mengurangi beban komputasi, khususnya pada tahap segmentasi atau ekstraksi fitur. Selain itu, banyak arsitektur *deep learning* modern, termasuk *Generative Adversarial Networks* (GAN), memanfaatkan citra *grayscale* sebagai *input* karena detail spasialnya tetap memadai untuk proses pelatihan model[22].

E. Vanilla Generative Adversarial Network (VGAN)

Vanilla Generative Adversarial Networks, atau disingkat VGAN, adalah pendekatan untuk pemodelan generatif menggunakan metode pembelajaran mendalam, seperti jaringan saraf *convolutional.* yang bertujuan untuk membentuk atau membuat suatu data yang benar-benar baru berdasarkan data latih atau data yang sudah dilihat sebelumnya, biasa di implementasikan pada data citra yang berupa *pixel*. GAN memiliki kelebihan di bidang *image to image translation* dan *image generator* termasuk pewarnaan citra [20]. Dalam beberapa tahun terakhir GAN telah menjadi salah satu arsitektur yang paling menarik dalam sistem pembelajaran mesin. Sejak pertama kali diperkenalkan oleh Goodfellow dkk. pada tahun 2014 [23], berbagai varian GAN telah dikembangkan oleh para peneliti. Sebagian besar varian tersebut tetap mempertahankan kerangka dasar dari *vanilla* GAN, yang terdiri atas dua jaringan saraf: sebuah *generator* ***G*** dan sebuah *discriminator* ***D***. Kedua jaringan ini belajar secara adversarial (saling bersaing) selama proses pelatihan. Gambar 2.5 menggambarkan skema arsitektur *vanilla* GAN, di mana *generator* ***G*** menghasilkan citra palsu dari kode laten acak $z \sim p_z$, dan *discriminator* ***D*** belajar untuk membedakan antara sampel nyata dan palsu. Konsep utama dari GAN biasanya didefinisikan sebagai masalah permainan dengan tujuan *min–max*. Tujuannya adalah memperoleh *generator* yang optimal, yang mampu menghasilkan citra kualitas tinggi dan realistis seperti gambar alami, dengan melakukan penyetelan hiperparameter secara tepat.

Vanilla GAN merupakan arsitektur dasar *Generative Adversarial Network* yang pertama kali diperkenalkan oleh Goodfellow dkk. (2014) dan banyak dijadikan rujukan untuk pengembangan arsitektur generatif selanjutnya. Creswell dkk. [24] menjelaskan bahwa Vanilla GAN bekerja dengan dua komponen utama, yaitu *generator* dan *discriminator* yang dilatih secara bersamaan dengan prinsip *minimax game*. *Generator* bertugas menghasilkan data tiruan yang menyerupai data asli, sedangkan *discriminator* bertugas membedakan mana data asli dan mana data tiruan.

Pada tahap pelatihan, *generator* berupaya memperbaiki hasil buatannya agar semakin mirip dengan data asli, sementara *discriminator* terus belajar mendeteksi perbedaan. Melalui mekanisme persaingan ini, model dapat menghasilkan data sintetis berkualitas tinggi, termasuk citra wajah resolusi tinggi. Vanilla GAN menjadi landasan teori penting untuk metode SRGAN, yang merupakan pengembangan GAN khusus untuk peningkatan resolusi citra [24].



GAMBAR 5
Arsitektur VGAN

F. *Super Resolution Generative Adversarial Network (SRGAN)*

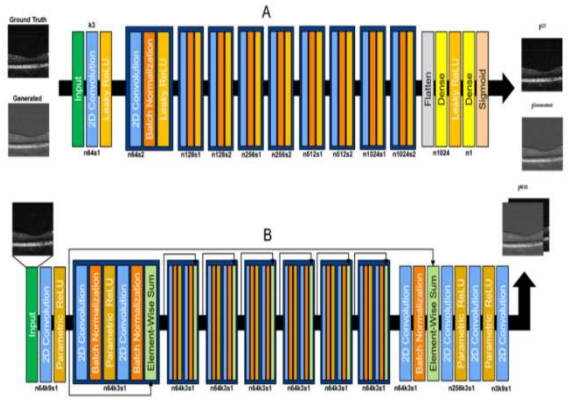
S*uper Resolution Generative Adversarial Network (SRGAN)* adalah metode untuk *single super resolution image*. SRGAN menggunakan *loss function perception* yang terdiri dari *an adversarial loss and a content loss*. Metode ini membuat solusi ke *manifold* agar gambar alami menggunakan jaringan *discriminator* yang dilatih untuk membedakan antara gambar super resolusi dan gambar foto realistis asli. Selain itu, menggunakan *content loss* yang didukung oleh kesamaan persepsi alih - alih kesamaan dalam ruang piksel. Jaringan yang sebenarnya digambarkan terutama terdiri dari blok sisa untuk ekstraksi fitur[13]. Secara formal kita dapat menulis *loss perception* sebagai

jumlah tertimbang dari *content loss* (VGG) dan *an adversarial loss component :*

$$L^{SR} = L_X^{SR} + 10^{-3} L_{GEN}^{SR} \quad (1)$$

Diadaptasi dari [25].
L^{SR} = total *perceptual loss*
L_X^{SR} = *Content loss* yang dihitung sebagai selisih fitur antara citra hasil (*generated image*) dan citra target (*ground truth*) dalam ruang fitur dari jaringan VGG.
L_{GEN}^{SR} = *Adversarial loss* yang diperoleh dari umpan balik *discriminator* untuk mendorong *generator* menghasilkan gambar yang tampak realistis.
(10^{-3}) = koefisien penimbang yang mengatur kontribusi relatif antara *content loss* dan *adversarial loss.*

Ledig dkk. [25] memperkenalkan *Super Resolution Generative Adversarial Network* (SRGAN) sebagai salah satu arsitektur GAN yang dirancang khusus untuk merekonstruksi citra resolusi tinggi dari citra resolusi rendah dengan hasil yang fotorealistik. SRGAN memadukan *generator* dengan jaringan *residual blocks* untuk meningkatkan kemampuan representasi fitur, serta *discriminator* yang menguji keaslian citra hasil rekonstruksi. Salah satu keunggulan SRGAN adalah penggunaan *perceptual loss* berbasis jaringan VGG, yang memungkinkan model meminimalkan perbedaan persepsi visual antara citra hasil rekonstruksi dengan citra asli. Hal ini sangat bermanfaat dalam konteks peningkatan resolusi citra wajah, karena detail tekstur halus, tepi objek, dan pola halus pada area sensitif seperti mata dan rambut dapat dipertahankan lebih baik dibandingkan metode interpolasi konvensional [25].



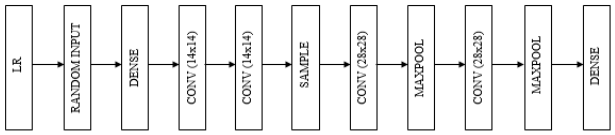
GAMBAR 6
Arsitektur SRGAN

III. METODE

Desain umum tugas akhir dengan melakukan modifikasi SRGAN dan VGAN berbasis tensorflow. Modifikasi dari SRGAN dan VGAN ini bertujuan menghasilkan output citra yang jelas dan lebih baik dibandingkan dengan SRGAN dan VGAN asli. Perbandingan nilai PSNR dan SSIM dapat mengetahui apakah sistem tersebut lebih baik dari sebelumnya.

A. VGAN

Penelitian ini menggunakan VGAN sebagai acuan dasar untuk dibandingkan dengan struktur VGAN yang telah dimodifikasi, sehingga dapat dievaluasi peningkatan performa yang diperoleh melalui penambahan *layer downsampling* dan *upsampling* pada jaringan.

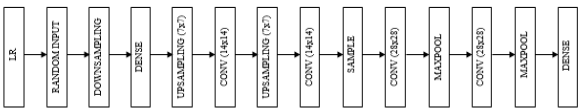


GAMBAR 7
Arsitektur VGAN Asli

Gambar 7 memperlihatkan arsitektur dasar VGAN Asli yang terdiri dari dua komponen utama, yaitu *generator* dan *discriminator*. Pada struktur VGAN asli, *generator* bertugas membangkitkan citra sintetis dari *input* berupa vektor acak (*random noise*), sedangkan *discriminator* berfungsi membedakan antara citra nyata dengan citra hasil *generator*.
Proses pembelajaran berlangsung melalui mekanisme kompetitif, di mana *generator* berusaha menghasilkan citra yang semakin menyerupai citra nyata agar dapat menipu *discriminator*. Sebaliknya, *discriminator* dilatih untuk semakin mahir mengenali citra palsu dari citra asli.
Struktur ini belum dilengkapi dengan proses *downsampling* maupun *upsampling* yang lebih kompleks, sehingga menjadi acuan dasar untuk dibandingkan dengan arsitektur VGAN yang telah dimodifikasi pada penelitian ini.

B. VGAN Modifikasi

Untuk meningkatkan performa dalam menghasilkan citra wajah yang jelas, diperlukan penyesuaian pada arsitektur VGAN agar mampu menangkap detail spasial dengan lebih baik. VGAN akan dimodifikasi pada struktur *layer* nya dengan menerapkan proses *downsampling* terlebih dahulu pada *dataset* wajah.



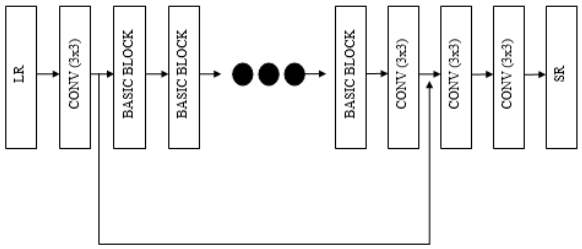
GAMBAR 8
Arsitektur VGAN Modifikasi

Gambar 8 menunjukkan arsitektur VGAN modifikasi. Modifikasi pada VGAN dilakukan dengan menambahkan *layer downsampling* dan *upsampling* ke dalam arsitektur aslinya. Pada dasarnya, VGAN standar hanya menggunakan *generator* untuk menghasilkan citra sintetis dari *input noise* (random *input*) dan *discriminator* untuk membedakan citra asli dan citra hasil generasi.
Dalam modifikasi ini, ditambahkan *layer downsampling* di tahap awal. Tujuannya adalah untuk memampatkan informasi dari citra masukan agar fitur global dapat diekstrak lebih baik.

Setelah proses *downsampling*, dilakukan *upsampling* agar citra yang dihasilkan jelas dan detail yang lebih tajam. Dengan modifikasi ini, diharapkan performa VGAN dapat meningkat dalam menghasilkan citra wajah yang jelas dibandingkan struktur VGAN asli.

C. *SRGAN*

Penelitian ini menggunakan SRGAN sebagai referensi dasar untuk mengevaluasi efektivitas modifikasi arsitektur melalui penambahan proses *downsampling* dan *upsampling* dalam meningkatkan kualitas hasil rekonstruksi citra



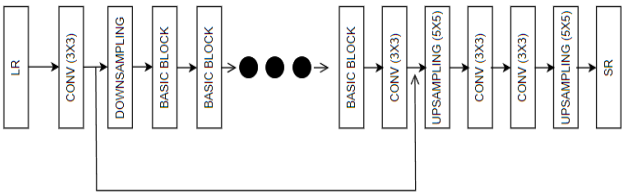
GAMBAR 9
Arsitektur SRGAN Asli

Gambar 9 menunjukkan bahwa struktur SRGAN asli. Pada arsitektur standarnya, SRGAN terdiri dari dua komponen utama, yaitu generator dan discriminator. Generator berfungsi membangkitkan citra resolusi tinggi dari input citra resolusi rendah, sedangkan discriminator bertugas membedakan antara citra hasil rekonstruksi dengan citra resolusi tinggi asli.

Keunggulan SRGAN terletak pada penggunaan residual blocks di dalam generator, yang memungkinkan jaringan mempelajari detail spasial yang lebih kompleks tanpa mengalami masalah degradasi gradien. Struktur residual block ini umumnya terdiri dari beberapa lapisan konvolusi berukuran kecil, batch normalization, serta fungsi aktivasi ReLU, dengan ciri khas adanya skip connection yang melewatkan input langsung ke output. Dengan kombinasi tersebut, SRGAN mampu menghasilkan citra resolusi tinggi yang lebih tajam dan mendetail dibandingkan metode konvensional.

D. SRGAN Modifikasi

SRGAN akan dilakukan modifikasi pada layernya dengan melakukan *downsampling* pada *dataset* wajah dan akan dilakukan *upsampling*. *Upsampling* bertujuan untuk meningkatkan kualitas gambar keluaran dengan pengaplikasian *super resolution*.



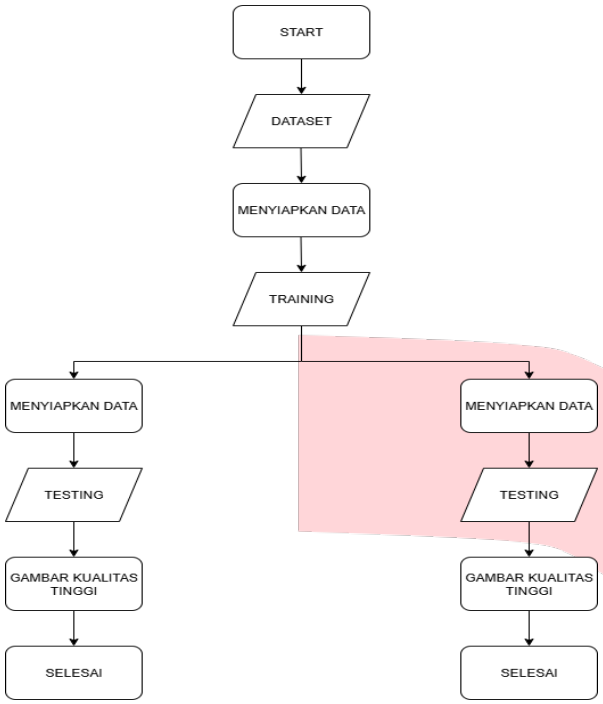
GAMBAR 10
Arsitektur SRGAN Modifikasi

Gambar 10 menunjukkan struktur SRGAN hasil modifikasi. Modifikasi dilakukan dengan menambahkan layer *downsampling* setelah *layer* konvolusi, kemudian diikuti proses *upsampling* sebanyak dua kali. Tujuan modifikasi ini adalah untuk membedakan hasil keluaran antara struktur asli dan struktur modifikasi, sehingga diharapkan citra yang dihasilkan memiliki kualitas resolusi lebih baik.

Arsitektur SRGAN sendiri dirancang untuk menghasilkan citra resolusi tinggi dari citra buram dengan memanfaatkan kemampuan GAN yang dilengkapi residual block. Komponen residual block terdiri dari beberapa lapisan konvolusi berukuran kecil, seperti 3×3, yang dipadukan dengan batch normalization dan fungsi aktivasi ReLU. Ciri khasnya adalah adanya skip connection yang langsung melewatkan *input* ke *output*, sehingga membantu mempertahankan informasi fitur penting dan menjaga propagasi gradien tetap stabil pada jaringan yang dalam.

Modifikasi pada Super Resolution Generative Adversarial Network (SRGAN) juga difokuskan pada penyesuaian struktur layer. SRGAN Asli sudah dirancang untuk meningkatkan resolusi citra rendah dengan memanfaatkan residual blocks dan sub pixel convolution untuk proses *upsampling*. Pada penelitian ini, modifikasi dilakukan dengan menambahkan layer *downsampling* di tahap awal *generator*, kemudian diikuti proses *upsampling* lebih dari satu tahap. Downsampling bertujuan untuk menekan informasi yang tidak relevan dan menyoroti fitur global, sedangkan *upsampling* bertugas membangun kembali detail spasial agar citra keluarannya tetap memiliki resolusi tinggi yang realistis. Dengan struktur ini, diharapkan SRGAN mampu menghasilkan *output* citra dengan kualitas tekstur dan ketajaman yang lebih baik dibandingkan SRGAN tanpa modifikasi.

E. Perancangan Sistem



GAMBAR 11
Arsitektur SRGAN Modifikasi

a. *Dataset*

Pada proses ini, akan dilakukan pemilihan *dataset*. Dataset yang akan digunakan adalah CelebA - HQ *dataset* yang untuk keperluan *testing*. Dataset yang digunakan untuk melatih dan menguji sistem adalah set CelebA - HQ yang mengandung objek wajah dengan resolusi 128x128 yang merupakan *dataset* publik. Proses penyamaan rasio akan dilakukan apabila tedapat citra yang belum memenuhi ukuran rasio yang akan dipakai untuk melakukan proses *training* model, dimana ukuran rasio yang akan digunakan adalah sebesar 1:1.

b. *Persiapan data*

Proses ini merupakan kegiatan untuk melakukan *mounting* di *Visual studio code*. Setelah itu, *dataset* yang telah diperoleh dari internet dipindahkan ke *folder* penyimpanan laptop, supaya pada sistem dapat memanggil data.Tahap persiapan data merupakan langkah awal yang sangat penting sebelum proses *training* model dilakukan. Pada tahap ini, *dataset* citra wajah diunduh dari sumber terpercaya di internet, salah satunya adalah *dataset* CelebA-HQ yang banyak digunakan pada penelitian peningkatan kualitas citra wajah. Dataset kemudian dipindahkan ke direktori lokal di laptop dan disusun dalam struktur *folder* yang rapi agar mudah diakses program. Folder *dataset* umumnya dibagi menjadi dua bagian, yaitu data *training* dan data *testing*, dengan proporsi 80% untuk *training* dan 20% untuk *testing*.

Setelah *dataset* siap, dilakukan proses *mounting* direktori di *Visual Studio Code* agar *file* dapat dibaca oleh skrip *Python* secara langsung tanpa perlu memindahkan *file* secara manual. Selain itu, pada tahap ini juga dilakukan pengecekan format *file* gambar, perubahan ukuran citra jika diperlukan (misalnya *resize* ke ukuran *input* LR), dan normalisasi piksel agar data konsisten. Proses persiapan data yang tertata dengan baik akan memudahkan proses *training* dan *testing*, serta meminimalisir *error* saat program dijalankan.

c. *Training*

Dataset *training* pada penelitian kali ini ada 800 citra yang berarti 80% dari total keseluruhan *dataset*. Proses *training* bertujuan untuk melatih model jaringan saraf tiruan agar dapat menghasilkan citra resolusi tinggi yang mendekati citra referensi aslinya. Pada tahap ini digunakan 800 citra wajah (80% dari *dataset*) sebagai data latih. Citra *input* resolusi rendah akan melalui beberapa tahap pemrosesan di jaringan *generator* yang memiliki *layer* khusus seperti konvolusi, *downsampling*, *residual block*, dan *upsampling*. *Training* dilakukan secara iteratif dengan menggunakan hyperparameter epoch yang berfungsi menentukan berapa kali seluruh *dataset* dilatih secara penuh. *Epoch* adalah salah satu *hyperparameter* dari *deep learning* untuk menentukan berapa kali algoritma pembelajaran akan bekerja mengolah seluruh *dataset training*. Semakin banyak jumlah epoch, semakin besar peluang model untuk belajar pola detail spasial pada data.

Selama *training*, bobot pada setiap *layer* dioptimasi menggunakan algoritma optimasi seperti *Adam Optimizer*. Proses pelatihan ini memerlukan perhitungan *loss function* untuk mengukur seberapa jauh prediksi *output* dari citra *input* dibandingkan dengan citra referensi (*ground truth*). Dalam SRGAN, selain *content loss*, juga digunakan *perceptual loss* agar detail tekstur yang dihasilkan lebih halus dan mendekati persepsi manusia. Hasil dari tahap *training* adalah model *generator* yang telah dilatih untuk menghasilkan citra yang tidak jelas atau buram menjadi citra yang jelas.

d. *Modified VGAN* dan *SRGAN*

Pada penelitian ini, arsitektur SRGAN dan VGAN digunakan sebagai model generatif utama dengan beberapa modifikasi pada strukturnya. Modifikasi dilakukan dengan menambahkan *layer downsampling* di awal arsitektur untuk memampatkan informasi dari citra wajah resolusi rendah sehingga jaringan dapat mengekstrak fitur global lebih baik. Proses *downsampling* ini penting karena memungkinkan jaringan fokus pada representasi fitur utama dengan mengurangi detail redundan, mirip konsep *bottleneck* pada *autoencoder*, sehingga membantu mengurangi beban komputasi dan mempermudah proses pembelajaran pola penting dari data. Dengan kata lain, *downsampling* membuat model tidak sekadar mengingat piksel, tetapi benar-benar mempelajari fitur karakteristik wajah yang paling relevan.

Setelah informasi global berhasil diekstraksi, dilakukan proses *upsampling* secara bertahap pada *generator* untuk membangun kembali detail spasial citra hingga mencapai kualitas yang tinggi. Proses *upsampling* bertahap ini bertujuan menghasilkan citra dengan detail tekstur yang halus dan tajam, sekaligus meminimalkan artefak visual seperti checkerboard artifacts yang sering muncul jika pembesaran resolusi dilakukan secara mendadak dalam satu tahap. Dengan *upsampling* bertahap, citra diperbesar sedikit demi sedikit

sambil melewati lapisan konvolusi yang bertugas menambahkan detail tekstur secara progresif.

Pada SRGAN, modifikasi juga diterapkan dengan menggunakan beberapa blok residual yang berfungsi memperdalam jaringan agar dapat menangkap detail tekstur lebih kaya. Blok residual ini membantu propagasi gradien agar proses pelatihan tetap stabil meskipun jaringan lebih dalam, sehingga model dapat mempelajari detail halus pada area sensitif wajah seperti tepi rambut, mata, atau kulit. Sementara itu, VGAN tetap menggunakan pendekatan GAN dasar tanpa blok residual sebagai pembanding. Dengan demikian, perbandingan kinerja antara VGAN dan SRGAN dapat menunjukkan pengaruh penambahan blok residual, *perceptual loss*, serta strategi *downsampling* dan *upsampling* bertahap terhadap kualitas hasil super resolusi citra wajah.

e. *Testing*

Proses *testing* merupakan tahap evaluasi untuk mengukur performa model yang telah dilatih. Pada tahap ini digunakan 200 citra wajah atau 20% dari total *dataset* CelebA-HQ yang tidak pernah digunakan pada saat *training*. Dataset *testing* ini berfungsi untuk mengetahui kemampuan generalisasi model ketika diberikan data baru yang belum pernah dilihat sebelumnya. Setiap citra resolusi rendah pada *dataset testing* diproses melalui model VGAN dan SRGAN yang telah selesai dilatih.

Hasil *output* kemudian dibandingkan dengan citra referensi resolusi tinggi menggunakan dua parameter evaluasi, yaitu *Peak Signal to Noise Ratio* (PSNR) dan *Structural Similarity Index Measure* (SSIM). PSNR digunakan untuk menghitung kualitas numerik rekonstruksi piksel, sedangkan SSIM digunakan untuk menilai kesamaan struktur, tekstur, dan kontras visual antara citra hasil dan citra aslinya. Penggunaan dua parameter ini diharapkan dapat memberikan penilaian yang objektif terkait tingkat keberhasilan rekonstruksi resolusi citra. Hasil *testing* inilah yang akan dianalisis lebih lanjut untuk menarik kesimpulan keefektifan model yang digunakan.

f. Gambar Kualitas Tinggi

Tahap ini merupakan tahap akhir dari keseluruhan proses peningkatan kualitas citra wajah. Setelah melalui proses *testing*, sistem akan menghasilkan citra keluaran dengan kualitas yang lebih jelas dibandingkan *input* awal. Citra kualitas tinggi hasil rekonstruksi ini kemudian disimpan secara otomatis ke *folder output* dalam format *file* gambar (seperti *PNG* atau *JPEG*) dengan kualitas yang sesuai target. Hasil citra kualitas tinggi ini menjadi bukti visual keberhasilan model dalam melakukan super resolusi.

Selain disimpan, hasil citra kualitas tinggi juga dapat ditampilkan secara visual untuk keperluan analisis. Peneliti dapat membandingkan hasil rekonstruksi dengan citra referensi untuk menilai seberapa tajam detail tekstur yang dihasilkan, terutama pada area sensitif seperti tepi wajah, mata, dan rambut. Tahap ini sekaligus menunjukkan manfaat nyata dari penerapan model SRGAN dan VGAN yang dimodifikasi untuk meningkatkan kualitas citra wajah pada *dataset* CelebA-HQ.

g. Perangkat

Pada Tugas Akhir kali ini penulis menggunakan perangkat lunak (*software*) dan perangkat keras (*hardware*) dalam pembuatan sistem. Software yang digunakan *Visual Studio Code* , sementara perangkat keras yang digunakan adalah Laptop HP Omen 15. Spesifikasi *software* dan *hardware* yang digunakan dalam penelitian ini dijelaskan pada tabel 1.

TABEL 1
Spesifikasi Software

Sistem Operasi	Windows 11 64-bit
Software	*Visual Studio Code*
Hardware	Laptop HP Omen 15

IV. HASIL DAN PEMBAHASAN

Bab ini membahas mengenai hasil pengujian sistem yang telah dibuat dengan *dataset* yang telah diambil. Tujuan dari pengujian ini menganalisis perbandingan hasil performansi dari dua model VGAN dan dua model SRGAN melalui dua metode yang berbeda. Metode yang pertama adalah VGAN dan SRGAN original. Metode yang kedua adalah VGAN dan SRGAN akan disimulasi dengan dilakukan *upsampling* yang disisipkan pada *layer* - *layer*. Dengan menggunakan *dataset* yang sama pada ke-dua skenario pengujian, maka ke-dua model tersebut akan ditinjau dan dianalisa dengan dua parameter performansi, yaitu PSNR dan SSIM.

PSNR (*Peak Signal to Noise Ratio*) digunakan untuk mengukur tingkat kerusakan gambar asli dengan *noise* sebagai perserpsi manuusia terhadap gambar. PSNR dihasilkan dari MSE (*Mean Square Error*) suatu citra dimana semakin tinggi nilai PSNR pada gambar maka, semakin baik gambar yang dihasilkan. MSE merupakan ukuran yang digunakan untuk menilai seberapa baik sebuah metode dalam melakukan restorasi citra relatif terhadap citra aslinya. MSE didefinisikan sebagai berikut:

$$\text{MSE} = \frac{1}{M \times N} \sum_{i=1}^{M} \sum_{j=1}^{N} [I(i,j) - K(i,j)]^2 \qquad (2)$$

Diadaptasi dari [21].

Dimana :

M : adalah jumlah baris piksel gambar (tinggi gambar)

N : adalah jumlah kolom piksel gambar (lebar gambar)

I(i,j) : adalah nilai intensitas piksel pada koordinat (i,j) pada gambar asli (ground truth)

K(i,j): adalah nilai intensitas piksel pada koordinat (i,j) pada gambar hasil rekonstruksi

Untuk gambar berwarna, perhitungan MSE dilakukan pada masing-masing kanal, lalu dirata-ratakan.Nilai MSE ini digunakan untuk menghitung nilai PSNR. Semakin kecil nilai MSE, ini menunjukkan bahwa citra yang dihasilkan semakin

baik atau mirip dengan citra aslinya. PSNR dalam desibel didefinisikan sebagai berikut:

$$\text{PSNR} = 10 \cdot \log_{10} \frac{L^2}{MSE} \quad (3)$$

Dimana :

L^2 : adalah piksel maksimum dari citra

MSE : adalah rata – rata nilai *error* pada citra asli dengan modifikasi

SSIM (*Structural Similarity Index*) adalah alat ukur yang terdiri dari tiga faktor yaitu pencahayaan, kontras, dan struktur. Metode ini biasanya digunakan untuk membandingkan dua gambar dan mengukur kualitas sebuah gambar dapat ditulis sebagai berikut:

$$\text{SSIM}(x,y) = \frac{(2\mu_x\mu_y+C_1)+(2\sigma_{xy}+C_2)}{(\mu_{x^2}+\mu_{y^2}+C_1)(\sigma_{x^2}+\sigma_{y^2}+C_2)} \quad (4)$$

Dimana:

μ_x ,μ_y, σ_x, σ_y, dan σ_{xy} adalah *local means*, *standart deviations*, dan *cross covariance* untuk citra x,y.

A. Hasil Simulasi

Pada penelitian ini metode yang digunakan adalah VGAN dan SRGAN yang kemudian dimodifikasi dengan meyisipkan *downsampling* lalu dilakukan *upsampling* pada *layer* - *layer*. Dataset yang digunakan untuk seluruh pengujian ini yaitu sebanyak 1000 data yang akan dibagi 80% gambar *training* dan 20% untuk data tes. Pada *dataset* CelebA - HQ, pengujian akan di lakukan dengan lima gambar dari *dataset* CelebA - HQ yaitu wajah1.png, wajah2.png, wajah3.png, wajah4.png, dan wajah5.png. Sistem pengujian dengan model VGAN dan SRGAN. Model VGAN asli dan VGAN modifikasi menggunakan *dataset* yang akan di*training* dengan *epoch* 50 sehingga dapat diambil hasil nilai dari gambar *output* dengan meninjau parameter PSNR dan SSIM. Model SRGAN dan SRGAN modifikasi *dataset* akan di di*training* dengan epoch 50 sama seperti VGAN dan diambil hasil dari gambar *output* dengan meninjau parameter PSNR dan SSIM. Pada *dataset* CelebA - HQ terdapat 1000 gambar yang berisi beragam gambar wajah manusia yang berbeda asal negara , gender , dan usia. Faktor yang banyaknya perbedaan variasi gambar inilah yang menghasilkan *training* yang baik sehingga dapat berpengaruh terhadap nilai PSNR dan SSIM. Pengujian akan dilakukan dengan skenario program VGAN dan SRGAN yang menggunakan *epoch* 50 dan melakukan *testing* dengan memperhatikan nilai parameter PSNR dan SSIM . Berikut pada tabel dibawah ini akan menampilkan data hasil nilai PSNR dan SSIM dari simulasi VGAN dan SRGAN.

TABEL 2
Hasil PSNR VGAN and SRGAN (Asli vs Modifikasi)

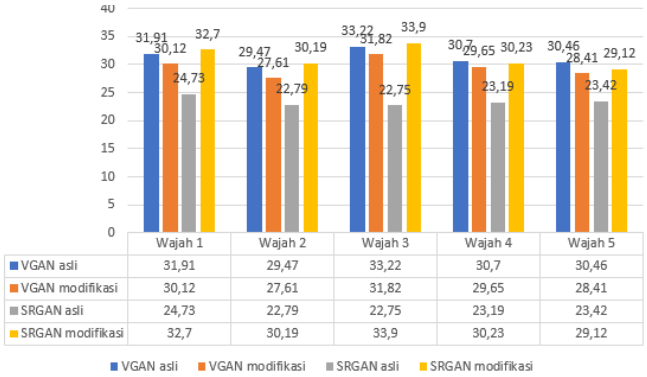
Image	PSNR (VGAN original)	PSNR (SRGAN original)	PSNR (VGAN Mod)	PSNR (SRGAN Mod)
wajah1.jpg	31.91 dB	24.73 dB	30.12 dB	32.7 dB
wajah2.jpg	29.47 dB	22.79 dB	27.61 dB	30.19 dB
wajah3.jpg	33.22 dB	22.75 dB	31.82 dB	33.9 dB
wajah4.jpg	30.70 dB	23.19 dB	29.65 dB	30.23 dB
wajah5.jpg	30.46 dB	23.42 dB	28.41 dB	30.82 dB

TABEL 3
Hasil SSIM VGAN and SRGAN (Asli vs Modifikasi)

Image	SSIM (VGAN original)	SSIM (SRGAN original)	SSIM (VGAN Mod)	SSIM (SRGAN Mod)
wajah1.jpg	0.936	0.74	0.913	0.923
wajah2.jpg	0.871	0.637	0.828	0.843
wajah3.jpg	0.932	0.639	0.91	0.84
wajah4.jpg	0.897	0.681	0.885	0.89
wajah5.jpg	0.9	0.684	0.868	0.82

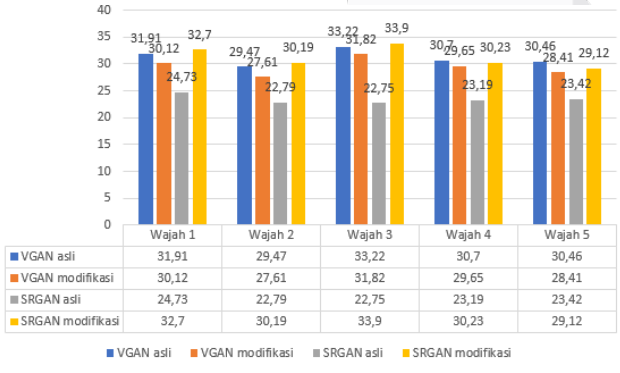
Dapat dilihat dari Tabel 2 dan Tabel 3 dari hasil pengujian didapatkan pada *training* dengan *dataset* wajah 1 hingga wajah 5. Proses pengajian *dataset* menggunakan SRGAN mendapatkan peningkatan hasil yang signifikan jika ditinjau dari model VGAN, dan dapat dilihat dari nilai rata-rata PSNR model SRGAN modifikasi dapat mengungguli VGAN modifikasi saat *testing* menggunakan citra wajah1.png hingga wajah5.png dengan nilai PNSR berturut – turut sebesar 30,12 dB, 27,61 dB, 31.82 dB, 29,65 dB, dan 28,41 dB tersebut tidak ada satupun yang mengungguli SRGAN modifikasi.

Berdasarkan hasil pengukuran PSNR dan SSIM pada lima gambar wajah, dapat dilihat perbedaan performa antara model VGAN dan SRGAN dalam hal peningkatan kualitas gambar. Pada model VGAN, kualitas gambar cenderung menurun setelah proses modifikasi. Hal ini terlihat dari rata-rata PSNR yang turun dari 31.15 menjadi 29.52 dan rata-rata SSIM yang menurun dari 0.9072 menjadi 0.8808. Sebaliknya, model SRGAN menunjukkan peningkatan yang signifikan setelah modifikasi, dengan rata-rata PSNR naik dari 23.78 menjadi 31.23 dan SSIM dari 0.6762 menjadi 0.8632. Jika dibandingkan antar model pada hasil gambar modifikasi, SRGAN menghasilkan nilai PSNR dan SSIM yang lebih tinggi dibandingkan VGAN. Ini menunjukkan bahwa SRGAN lebih mampu meningkatkan ketajaman gambar (PSNR) sekaligus mempertahankan struktur asli gambar (SSIM). Secara individual, peningkatan paling mencolok terlihat pada gambar wajah1.png, di mana nilai PSNR SRGAN naik dari 24.73 menjadi 32.7 dan SSIM dari 0.74 menjadi 0.923, sementara VGAN justru mengalami penurunan. Berdasarkan temuan ini, dapat disimpulkan bahwa SRGAN secara konsisten memberikan peningkatan kualitas visual yang lebih baik dibandingkan VGAN, baik dari aspek kejernihan maupun kemiripan struktural terhadap gambar asli.



GAMBAR 12
Grafik Hasil PSNR Simulasi 2 Model

Diagram batang pada gambar 12 menunjukkan perbandingan nilai PSNR dari lima wajah menggunakan empat metode berbeda, yaitu VGAN asli, VGAN modifikasi, SRGAN asli, dan SRGAN modifikasi. Dapat dilihat bahwa modifikasi pada kedua model mendapatkan hasil PSNR yang baik. SRGAN memberikan peningkatan kualitas hasil yang signifikan dibandingkan versi aslinya. VGAN modifikasi menghasilkan nilai PSNR yang lebih rendah dibandingkan VGAN asli pada setiap wajah, menandakan penurunan kualitas citra. Pada SRGAN versi modifikasi secara konsisten menunjukkan nilai PSNR yang jauh lebih tinggi dibandingkan versi aslinya. Menariknya, meskipun SRGAN asli memiliki nilai PSNR paling rendah di antara semua metode, setelah dimodifikasi, SRGAN justru menjadi metode dengan performa terbaik dan bahkan mengungguli VGAN asli di semua wajah. Hal ini menunjukkan bahwa modifikasi yang dilakukan pada SRGAN memberikan dampak yang sangat positif terhadap kualitas citra yang dihasilkan. Secara keseluruhan, dapat disimpulkan bahwa modifikasi pada SRGAN sangat efektif dalam meningkatkan kualitas hasil, dengan SRGAN modifikasi menjadi metode terbaik berdasarkan nilai PSNR.



GAMBAR 13
Grafik Hasil SSIM Simulasi 2 Model

Diagram batang pada gambar 13 diatas menunjukkan perbandingan nilai SSIM (*Structural Similarity Index*) dari lima wajah menggunakan empat metode, yaitu VGAN asli, VGAN modifikasi, SRGAN asli, dan SRGAN modifikasi. Secara umum, VGAN asli menghasilkan nilai SSIM tertinggi pada sebagian besar wajah, menandakan bahwa metode ini memiliki kemampuan terbaik dalam mempertahankan kesamaan struktur citra dibanding citra referensi. VGAN modifikasi menunjukkan performa yang sedikit lebih rendah dari versi aslinya, tetapi tetap berada pada kisaran nilai yang tinggi dan stabil, sehingga modifikasi tidak secara signifikan mengurangi kualitas struktural citra. Sebaliknya, SRGAN asli menunjukkan performa terendah dalam semua wajah, dengan nilai SSIM yang rendah dan tidak stabil. Namun, setelah dilakukan modifikasi, SRGAN mengalami peningkatan yang sangat signifikan. SRGAN modifikasi berhasil mendekati, dan dalam beberapa kasus hampir menyamai, performa VGAN modifikasi. Hal ini menunjukkan bahwa modifikasi pada SRGAN sangat efektif dalam meningkatkan kesamaan struktural citra hasil super-resolusi. Dengan demikian, dapat disimpulkan bahwa VGAN asli tetap menjadi metode terbaik dalam hal struktur citra, tetapi SRGAN modifikasi menunjukkan potensi besar sebagai alternatif yang kompetitif setelah dilakukan perbaikan.

B. Perbandingan Hasil Simulasi Testing Dataset

Perbandingan hasil simulasi akan dilakukan pada model VGAN dan SRGAN dengan menggunakan parameter yang sama, yaitu variasi epoch sebanyak 50 dan lima *dataset* berbeda, di mana masing-masing *dataset* menyimpan gambar dengan format yang bervariasi. Parameter ini digunakan untuk mengamati sejauh mana kemampuan model dalam memperbaiki kualitas citra setelah melalui tahap pelatihan (*training*) dan pengujian (*testing*) dengan memanfaatkan *dataset* CelebA-HQ.

Untuk memberikan gambaran yang lebih komprehensif mengenai performa masing-masing metode, berikut disajikan penjelasan mendetail mengenai hasil perbandingan kualitas citra pada lima gambar wajah, termasuk wajah1.png. Penjelasan ini tidak hanya memaparkan hasil pengamatan visual, tetapi juga dilengkapi dengan analisis metrik kuantitatif berupa nilai PSNR (*Peak Signal to Noise Ratio*) dan SSIM (*Structural Similarity Index*) yang diperoleh dari keempat skenario pengujian, yakni VGAN asli, SRGAN asli, VGAN modifikasi, dan SRGAN modifikasi.

Dengan adanya evaluasi ini, diharapkan dapat terlihat secara jelas sejauh mana tingkat kemiripan hasil rekonstruksi citra dengan gambar asli, serta identifikasi area-area perbaikan yang masih perlu dilakukan. Selain itu, melalui perbandingan ini pula, dapat diidentifikasi keunggulan relatif masing-masing model, baik dalam mempertahankan detail struktur wajah, mengurangi artefak, maupun menjaga akurasi warna. Dengan demikian, analisis yang disajikan di bagian ini diharapkan dapat mendukung simpulan mengenai model mana yang paling optimal dalam meningkatkan kualitas citra wajah sesuai dengan tujuan penelitian ini. Evaluasi dilakukan pada lima gambar uji wajah1.png, wajah2.png, wajah3.png, wajah4.png, dan wajah5.png untuk memberikan gambaran komprehensif tentang bagaimana setiap metode mempertahankan detail struktur wajah dan meningkatkan kualitas citra. Berikut ini merupakan hasil evaluasi terhadap lima gambar:

GAMBAR 14
Perbandingan wajah1.png

Dapat dilihat dari gambar 14 adalah perbandingan wajah1.png dari ke-empat skenario pengujian gambar hasil dari VGAN asli tampak mampu mempertahankan struktur wajah dengan cukup baik, meskipun terdapat sedikit terlihat *blur* pada area mata dan rambut. Sebaliknya, hasil dari SRGAN asli terlihat paling buruk, dengan distorsi warna yang signifikan dan detail wajah yang kabur, sesuai dengan nilai PSNR 24.73 dB dan SSIM 0.74 yang dihasilkan menunjukkan nilai yang rendah. VGAN asli dengan nilai PSNR 31.91 dB dan SSIM 0.936 memperlihatkan penurunan kualitas dibandingkan dengan VGAN modifikasi dengan nilai PSNR 30.12 dB dan SSIM 0.916. Hasil terbaik ditunjukkan oleh SRGAN modifikasi, yang secara visual sangat mendekati gambar asli, dengan detail wajah yang jelas dan alami, serta minim artefak. Temuan ini sejalan dengan nilai PSNR 32.7 dan SSIM 0,923 yang lebih tinggi dibandingkan SRGAN asli, menjadikan SRGAN modifikasi sebagai metode terbaik dalam menghasilkan kualitas citra yang menyerupai gambar asli.

GAMBAR 15
Perbandingan wajah2.png

Dapat dilihat dari gambar 15 adalah perbandingan wajah2.png dari ke-empat skenario pengujian gambar hasil dari VGAN asli tampak mampu mempertahankan struktur wajah dengan cukup baik, meskipun terdapat sedikit terlihat *blur* pada area mata dan rambut. Sebaliknya, hasil dari SRGAN asli terlihat paling buruk, dengan distorsi warna yang signifikan dan detail wajah yang kabur, sesuai dengan nilai PSNR 22.79 dB dan SSIM 0.637 yang dihasilkan menunjukkan nilai yang rendah. VGAN asli dengan nilai PSNR 29.47 dB dan SSIM 0.871 memperlihatkan penurunan kualitas dibandingkan dengan VGAN modifikasi dengan nilai PSNR 27.61 dB dan SSIM 0.828. Hasil terbaik ditunjukkan oleh SRGAN modifikasi, yang secara visual sangat mendekati gambar asli, dengan detail wajah yang jelas dan alami, serta minim artefak. Temuan ini sejalan dengan nilai PSNR 30.19 dan SSIM 0,843 yang lebih tinggi dibandingkan SRGAN asli, menjadikan SRGAN modifikasi sebagai metode terbaik dalam menghasilkan kualitas citra yang menyerupai gambar asli.

GAMBAR 16
Perbandingan wajah3.png

Dapat dilihat dari gambar 16 adalah perbandingan wajah3.png dari ke-empat skenario pengujian gambar hasil dari VGAN asli tampak mampu mempertahankan struktur wajah dengan cukup baik, meskipun terdapat sedikit terlihat *blur* pada area mata dan rambut. Sebaliknya, hasil dari SRGAN asli terlihat paling buruk, dengan distorsi warna yang signifikan dan detail wajah yang kabur, sesuai dengan nilai PSNR 22.75 dB dan SSIM 0.639 yang dihasilkan menunjukkan nilai yang rendah. VGAN asli dengan nilai PSNR 33.22 dB dan SSIM 0.932 memperlihatkan penurunan kualitas dibandingkan dengan VGAN modifikasi dengan nilai PSNR 31.82 dB dan SSIM 0.91. Hasil terbaik ditunjukkan oleh SRGAN modifikasi, yang secara visual sangat mendekati gambar asli, dengan detail wajah yang jelas dan alami, serta minim artefak. Temuan ini sejalan dengan nilai PSNR 33.9 dan SSIM 0,84 yang lebih tinggi dibandingkan SRGAN asli, menjadikan SRGAN modifikasi sebagai metode terbaik dalam menghasilkan kualitas citra yang menyerupai gambar asli.

GAMBAR 17
Perbandingan wajah4.png

Dapat dilihat dari gambar 17 adalah perbandingan wajah4.png dari ke-empat skenario pengujian gambar hasil dari VGAN asli tampak mampu mempertahankan struktur wajah dengan cukup baik, meskipun terdapat sedikit terlihat *blur* pada area mata dan rambut. Sebaliknya, hasil dari SRGAN asli terlihat paling buruk, dengan distorsi warna yang signifikan dan detail wajah yang kabur, sesuai dengan nilai PSNR 23.19 dB dan SSIM 0.681 yang dihasilkan menunjukkan nilai yang rendah. VGAN asli dengan nilai PSNR 30.1 dB dan SSIM 0.897 memperlihatkan penurunan kualitas dibandingkan dengan VGAN modifikasi dengan nilai PSNR 29.75 dB dan SSIM 0.885. Hasil terbaik ditunjukkan oleh SRGAN modifikasi, yang secara visual sangat mendekati gambar asli, dengan detail wajah yang jelas dan alami, serta minim artefak. Temuan ini sejalan dengan nilai PSNR 30.23 dan SSIM 0,89 yang lebih tinggi dibandingkan SRGAN asli, menjadikan SRGAN modifikasi sebagai metode terbaik dalam menghasilkan kualitas citra yang menyerupai gambar asli.

GAMBAR 18
Perbandingan wajah5.png

Dapat dilihat dari gambar 18 adalah perbandingan wajah5.png dari ke-empat skenario pengujian gambar hasil dari VGAN asli tampak mampu mempertahankan struktur wajah dengan cukup baik, meskipun terdapat sedikit terlihat *blur* pada area mata dan rambut. Sebaliknya, hasil dari SRGAN asli terlihat paling buruk, dengan distorsi warna yang signifikan dan detail wajah yang kabur, sesuai dengan nilai PSNR 23.42 dB dan SSIM 0.684 yang dihasilkan menunjukkan nilai yang rendah. VGAN asli degan nilai PSNR 30.46 dB dan SSIM 0.9 memperlihatkan penurunan kualitas dibandingkan dengan VGAN modifikasi dengan nilai PSNR 28.841 dB dan SSIM 0.868. Hasil terbaik ditunjukkan oleh SRGAN modifikasi, yang secara visual sangat mendekati gambar asli, dengan detail wajah yang jelas dan alami, serta minim artefak. Temuan ini sejalan dengan nilai PSNR 30.82 dan SSIM 0,82 yang lebih tinggi dibandingkan SRGAN asli, menjadikan SRGAN modifikasi sebagai metode terbaik dalam menghasilkan kualitas citra yang menyerupai gambar asli.

V. KESIMPULAN

Berdasarkan hasil analisis simulasi sistem perbaikan citra menggunakan model SRGAN (Super Resolution Generative Adversarial Network) dan VGAN (Vanilla Generative Adversarial Network), dapat disimpulkan bahwa metode SRGAN modifikasi merupakan yang paling unggul karena mampu menghasilkan citra high resolution dengan peningkatan signifikan dibandingkan metode lain. Penelitian ini menitikberatkan pada evaluasi performa melalui nilai PSNR dan SSIM, di mana SRGAN modifikasi menunjukkan performa lebih baik dengan nilai PSNR 1,3 kali dan SSIM 1,2 kali lebih tinggi dibandingkan model SRGAN asli. Sementara itu, model VGAN modifikasi justru mengalami penurunan performa sebesar 0,1 atau 10% dibandingkan model VGAN asli.

REFERENSI

[1] M. P. Beham and S. M. M. Roomi, “A review of face recognition methods,” Int. J. Pattern Recognit. Artif. Intell., vol. 27, no. 4, pp. 1–35, 2013.

[2] Z. Cheng, X. Zhu, and S. Gong, “Surveillance face recognition challenge,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW), Salt Lake City, UT, USA, pp. 705–714, Jun. 2018.

[3] R. Yadlapati, P. J. Kahrilas, M. R. Fox, A. J. Bredenoord, C. P. Gyawali, S. Roman, et al., “Esophageal motility disorders on high-resolution manometry: Chicago classification version 4.0©,” Neurogastroenterol. Motil., vol. 33, no. 1, p. e14058, 2021, doi: 10.1111/nmo.14053

[4] W. W. W. Zou and P. C. Yuen, “Very low resolution face recognition problem,” IEEE Trans. Image Process., vol. 21, no. 1, pp. 327–340, Jan. 2012.

[5] J. Son, S. J. Park, and K.-H. Jung, “Retinal vessel segmentation in fundoscopic images with generative adversarial networks,” arXiv preprint arXiv:1706.09318, Jun. 2017. [Online]. Available: http://arxiv.org/abs/1706.09318

[6] B. Hardiansyah, A. P. Armin, and A. B. Yunanda, “Rekonstruksi citra pada super resolusi menggunakan interpolasi bicubic,” INTEGER J. Inf. Technol., vol. 4, no. 2, pp. 1–12, 2019.

[7] N. M. Abdi and S. Aisyah, “Peningkatan kualitas citra digital menggunakan metode super resolusi pada domain spasial,” J. Rekayasa Elektr., vol. 9, no. 3, pp. 137–142, 2011.

[8] W. Astuti, “Implementasi metode super resolusi untuk meningkatkan kualitas citra hasil screenshot,” JURIKOM J. Ris. Komputer, vol. 7, no. 3, p. 432, 2020.

[9] Y. Xiong et al., “Improved SRGAN for remote sensing image super-resolution across locations and sensors,” Remote Sens., vol. 12, no. 8, pp. 1–21, 2020.

[10] M. E. Abdulfattah, L. Novamizanti, and S. Rizal, “Super resolution pada citra udara menggunakan convolutional neural network,” ELKOMIKA J. Tek. Energi Elektr. Tek. Telekomun. Tek. Elektron., vol. 9, no. 1, pp. 71–78, 2021.

[11] G. A. Anarki, K. Auliasari, and M. Orisa, “Penerapan metode Haar cascade pada aplikasi deteksi masker,” JATI J. Mahasiswa Tek. Inform., vol. 5, no. 1, pp. 179–186, 2021.

[12] Y. Nagano and Y. Kikuta, “SRGAN for super-resolving low-resolution food images,” ACM Int. Conf. Proceeding Ser., pp. 33–37, 2018.

[13] C. Ledig et al., “Photo-realistic single image super-resolution using a generative adversarial network,” arXiv preprint arXiv:1609.04802, Sep. 2016. [Online]. Available: http://arxiv.org/abs/1609.04802.

[14] S. Chauduri, N. P. Galatsanos, and B. C. Tom, “Super Resolution Imaging in Reconstruction of a High Resolution Image from Low Resolution Image”. London, U.K.: Kluwer Academic, 2001.

[15] R. A. Sandi, “Super-resolusi berdasar pada fast registrasi dan rekontruksi maximum posteriori,” Skripsi, Institut Teknologi Sepuluh Nopember, 2009.

[16] Z. Wang, J. Chen, and S. C. H. Hoi, “Deep learning for image super-resolution: A survey,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 43, no. 10, pp. 3365–3387, Oct. 2020.

[17] L. Deng and D. Yu, “Deep learning: Methods and applications,” Found. Trends Signal Process., vol. 7, no. 3–4, pp. 197–387, 2014, doi: 10.1561/2000000039.

[18] I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning. Cambridge, MA, USA: MIT Press, 2016.

[19] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” Nature, vol. 521, no. 7553, pp. 436–444, May 2015, doi: 10.1038/nature14539.

[20] P. N. Andono, T. Sutojo, et al., Pengolahan Citra Digital. Yogyakarta, Indonesia: Penerbit Andi, 2017.

[21] R. C. Gonzalez and R. E. Woods, “Digital Image Processing”, 4th ed. New York, NY, USA: Pearson Education, 2018, p. 1022.

[22] R. Kaur and S. Kaur, “Comparative analysis of grayscale and RGB image classification using deep learning convolutional neural networks,” Procedia Comput. Sci., vol. 173, pp. 146–153, 2020, doi: 10.1016/j.procs.2020.06.017.

[23] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, “Generative adversarial nets,” in Advances in Neural Information Processing Systems (NeurIPS), Montreal, QC, Canada, pp. 2672–2680, Dec. 2014.

[24] A. Creswell, T. White, V. Dumoulin, K. Arulkumaran, B. Sengupta, and A. A. Bharath, “Generative adversarial networks: An overview,” IEEE Signal Process. Mag., vol. 35, no. 1, pp. 53–65, Jan. 2018.

[25] C. Ledig, L. Theis, F. Huszár, J. Caballero, A. Cunningham, A. Acosta, et al., “Photo-realistic single image super-resolution using a generative adversarial network,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Honolulu, HI, USA, pp. 4681–4690, Jul. 2017.