

**IMPLEMENTASI REAL-ESRGAN UNTUK PENINGKATAN
RESOLUSI CITRA BANGUNAN BERSEJARAH DI INDONESIA**

LAPORAN TUGAS AKHIR

Laporan ini Disusun untuk Memenuhi Salah Satu Syarat Memperoleh Gelar
Sarjana Strata 1 (S1) pada Program Studi Teknik Informatika Fakultas Teknologi
Industri Universitas Islam Sultan Agung Semarang



DISUSUN OLEH:

ANANDA FERI SETYAWAN

NIM 32602100003

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNOLOGI INDUSTRI
UNIVERSITAS ISLAM SULTAN AGUNG
SEMARANG
2025**

***IMPLEMENTATION OF REAL-ESRGAN TO IMPROVE
THE IMAGE RESOLUTION OF HISTORIC BUILDINGS IN INDONESIA***

FINAL PROJECT

*Laporan Proposed to complete the requirement to obtain a bachelor's degree (S1)
at Informatics Engineering Departement of Industrial Technology Faculty
Sultan Agung Islamic University of Semarang*



ANANDA FERI SETYAWAN

NIM 32602100003

MAJORING OF INFORMATICS ENGINEERING

INDUSTRIAL TECHNOLOGY FACULTY

SULTAN AGUNG ISLAMIC UNIVERSITY

SEMARANG

2025

LEMBAR PENGESAHAN TUGAS AKHIR

IMPLEMENTASI REAL-ESRGAN UNTUK PENINGKATAN RESOLUSI CITRA BANGUNAN BERSEJARAH DI INDONESIA

**ANANDA FERI SETYAWAN
NIM 32602100003**

Telah dipertahankan di depan tim penguji ujian sarjana tugas akhir
Program Studi Teknik Informatika
Universitas Islam Sultan Agung
Pada tanggal : 16 Juli 2025

TIM PENGUJI UJIAN SARJANA:

Sam Farisa Chaerul Haviana, ST., M.Kom
NIK. 210615046
(Ketua Penguji)



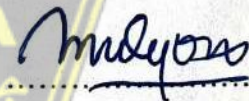
24-07-2025

Badieah, ST., M.Kom
NIK. 210615044
(Anggota Penguji)



23-07-2025

Ir. Sri Mulyono, M.Eng
NIK. 210616049
(Pembimbing)



30-07-2025

Semarang, 30 Juli 2025

Mengetahui,
Kaprodi Teknik Informatika
Universitas Islam Sultan Agung

Moch Taufik, S.T., M.IT
NIK. 210604034



SURAT PERNYATAAN KEASLIAN TUGAS AKHIR

Yang bertanda tangan dibawah ini :

Nama : Ananda Feri Setyawan

NIM : 32602100003

Judul Tugas : *IMPLEMENTASI REAL-ESRGAN UNTUK*

Akhir *PENINGKATAN RESOLUSI CITRA BANGUNAN
BERSEJARAH DI INDONESIA*

Dengan bahwa ini saya menyatakan bahwa judul dan isi Tugas Akhir yang saya buat dalam rangka menyelesaikan Pendidikan Strata Satu (S1) Teknik Informatika tersebut adalah asli dan belum pernah diangkat, ditulis ataupun dipublikasikan oleh siapapun baik keseluruhan maupun sebagian, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka, dan apabila di kemudian hari ternyata terbukti bahwa judul Tugas Akhir tersebut pernah diangkat, ditulis ataupun dipublikasikan, maka saya bersedia dikenakan sanksi akademis. Demikian surat pernyataan ini saya buat dengan sadar dan penuh tanggung jawab.

Semarang, 30 Juli 2025

Yang Menyatakan,



Ananda Feri Setyawan

PERNYATAAN PERSETUJUAN PUBLIKASI ILMIAH

Saya yang bertanda tangan dibawah ini :

Nama : Ananda Feri Setyawan

NIM : 32602100003

Program Studi : Teknik Informatika

Fakultas : Teknologi industri

Alamat Asal : Demak, Jawa Tengah

Dengan ini menyatakan Karya Ilmiah berupa Tugas akhir dengan Judul :
“Implementasi Real-ESRGAN Untuk Peningkatan Resolusi Citra Bangunan Bersejarah Di Indonesia” Menyetujui menjadi hak milik Universitas Islam Sultan Agung serta memberikan Hak bebas Royalti Non-Eksklusif untuk disimpan, dialihmediakan, dikelola dan pangkalan data dan dipublikasikan diinternet dan media lain untuk kepentingan akademis selama tetap menyantumkan nama penulis sebagai pemilik hak cipta. Pernyataan ini saya buat dengan sungguh-sungguh. Apabila dikemudian hari terbukti ada pelanggaran Hak Cipta/Plagiarisme dalam karya ilmiah ini, maka segala bentuk tuntutan hukum yang timbul akan saya tanggung secara pribadi tanpa melibatkan Universitas Islam Sultan agung.

Semarang, 30 Juli 2025

Yang menyatakan,



Ananda Feri Setyawan

KATA PENGANTAR

Segala puji dan syukur saya panjatkan ke hadirat Allah SWT atas limpahan rahmat dan karunia-Nya, sehingga saya dapat menyelesaikan Tugas Akhir ini dengan judul “*Implementasi Real-ESRGAN Untuk Peningkatan Resolusi Citra Bangunan Bersejarah Di Indonesia*”. Tugas Akhir ini disusun sebagai salah satu syarat kelulusan dalam menempuh studi serta untuk memperoleh gelar sarjana (S-1) pada Program Studi Teknik Informatika, Fakultas teknologi Industri, Universitas Islam Sultan Agung Semarang.

Dalam penyusunan Tugas Akhir ini, saya mendapatkan banyak bantuan, baik dalam aspek materi maupun teknis dari berbagai pihak. Oleh karena itu, dengan segala kerendahan hati, saya ingin mengucapkan terima kasih kepada :

1. Rektor UNISSULA Bapak Prof. Dr. H. Gunarto, SH., M.Hum yang telah memberikan kesempatan penulis menimba ilmu di kampus ini
2. Dekan Fakultas Teknologi Industri Ibu Dr. Ir. Hj. Novi Marlyana, S.T., M.T., IPU., ASEAN Eng
3. Dosen Pembimbing Bapak Ir. Sri Mulyono, M. Eng yang telah meluangkan waktu, membimbing, dan memberi ilmu.
4. Orang tua penulis yang telah mengizinkan untuk menyelesaikan laporan ini.
5. Dan kepada semua pihak yang tidak dapat saya sebutkan satu persatu.

Dengan segala kerendahan hati, penulis menyadari masih terdapat banyak kekurangan dari segi kualitas atau kuantitas maupun dari ilmu pengetahuan dalam penyusunan laporan, sehingga penulis mengharapkan adanya saran dan kritikan yang bersifat membangun demi kesempurnaan laporan ini dan masa mendatang.

Semarang, 30 Juli 2025



Ananda Feri Setyawan

DAFTAR ISI

COVER	i
LEMBAR PENGESAHAN TUGAS AKHIR	iii
SURAT PERNYATAAN KEASLIAN TUGAS AKHIR.....	iv
PERNYATAAN PERSETUJUAN PUBLIKASI ILMIAH	v
KATA PENGANTAR.....	vi
DAFTAR ISI.....	vii
DAFTAR GAMBAR	ix
DAFTAR TABEL	xii
ABSTRAK	xiii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Perumusan Masalah.....	5
1.3 Pembatasan Masalah.....	5
1.4 Tujuan	6
1.5 Manfaat	6
1.6 Sistematika Penulisan	6
BAB II TINJAUAN PUSTAKA DAN DASAR TEORI.....	8
2.1 Tinjauan Pustaka.....	8
2.2 Dasar Teori.....	10
2.2.1 Citra Bangunan Bersejarah Di Indonesia.....	10
2.2.2 Generative Advesarial Network (GAN).....	12
2.2.3 Super-Resolution Generative Adversarial Network (SRGAN)	14
2.2.4 Enhanced Super-Resolution Generative Adversarial Networks (ESRGAN).....	18
2.2.5 Real Enhanced Super Resolution Generative Adversarial Network (Real ESRGAN)	21
2.2.6 Python	24
2.2.7 Cross-Industry Standard Process for Data Mining (CRISP-DM)	25
2.2.8 Contrastive Language-Image Pretraining Image Quality Assessment CLIP-IQA	27
BAB III METODE PENELITIAN	29
3.1 Business Understanding.....	29
3.2 Data Understanding	31
3.3 Data Preparation	32
3.3.1 Blur.....	32

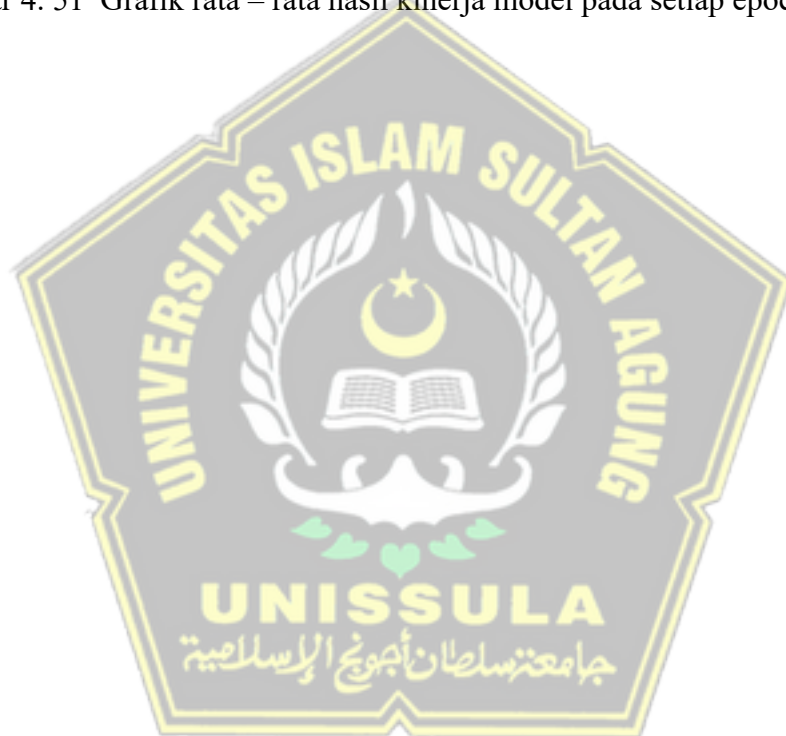
3.3.2	Resize (downsampling).....	33
3.3.3	Noise	33
3.3.4	JPEG Compression	34
3.4	Modeling.....	34
3.5	Evaluation	37
3.6	Deployment.....	38
BAB IV	HASIL DAN ANALISIS PENELITIAN.....	40
4.1	Hasil	40
4.1.1	Bussiness Understanding	40
4.1.2	Data Understanding.....	40
4.1.3	Data Preparation.....	44
4.1.3.1	Blur.....	45
4.1.3.2	Resize (downsampling).....	46
4.1.3.3	Noise	47
4.1.3.4	JPEG Compression	48
4.1.4	Modeling	49
4.1.5	Evaluation	52
4.1.6	Deployment.....	77
4.2	Analisis Penelitian	78
4.2.1	Impelementasi Metode Real-ESRGAN	78
4.2.2	Kinerja Metode Real-ESRGAN.....	80
BAB V	KESIMPULAN DAN SARAN	81
5.1	Kesimpulan	81
5.2	Saran.....	81
DAFTAR PUSTAKA.....		83

DAFTAR GAMBAR

Gambar 2. 1 Contoh citra bangunan bersejarah (Leiden University Libraries abad 19)	11
Gambar 2. 2 Arsitektur GAN (Goodfellow dkk., 2014)	12
Gambar 2.3 Arsitektur Generator SRGAN (Ledig dkk., 2017)	15
Gambar 2.4 Arsitektur Discriminator SRGAN (Ledig dkk., 2017)	16
Gambar 2. 5 Arsitektur Generator ESRGAN (Wang dkk., 2018)	18
Gambar 2. 6 Perbedaan antara generator SRGAN dan generator ESRGAN (Wang dkk., 2018)	19
Gambar 2. 7 Perbedaan antara standard discriminator dan relativistic discriminator dari (Wang dkk., 2018).....	19
Gambar 2.8 Arsitektur Generator Real ESRGAN (Wang dkk., 2021).....	22
Gambar 2. 9 Discriminator Real ESRGAN (Wang dkk., 2021)	22
Gambar 2. 10 Fase pada model CRISP-DM (Shearer, 2000)	25
Gambar 3. 1 Flowchart rancangan sistem	30
Gambar 3. 2 Flowchart high-order degradation model	32
Gambar 3. 3 Arsitektur dengan generator Real-ESRGAN (Wang dkk., 2021).....	35
Gambar 3. 4 Arsitektur dengan discriminator Real-ESRGAN (Wang dkk., 2021)	36
Gambar 3. 5 User interface halaman input.....	39
Gambar 3. 6 User interface halaman output.....	39
Gambar 4. 1 Contoh citra training dari Flickr2k.....	40
Gambar 4. 2 Citra Candi Borobudur (Leiden University Libraries abad 19)	41
Gambar 4. 3 Keterangan citra Candi Borobudur (Leiden University Libraries abad 19)	42
Gambar 4. 4 Citra Menara Kudus (Leiden University Libraries abad 19).....	43
Gambar 4. 5 Keterangan citra Menara Kudus (Leiden University Libraries abad 19)	44
Gambar 4. 6 Perbedaan anatara gambar asli dengan setelah proses high-order degradation model.....	45
Gambar 4. 7 Citra dengan degradasi blur.....	46
Gambar 4. 8 Citra dengan degradasi resize (downsampling)	47
Gambar 4. 9 Citra dengan degradasi noise.....	48
Gambar 4. 10 Degradasi dengan JPEG compression.....	49
Gambar 4. 11 Alur proses modeling pada metode Real-ESRGAN.....	49
Gambar 4. 12 Grafik hasil evaluasi citra Candi Borobudur original	53
Gambar 4. 13 Citra Candi Borobudur original (input).....	54
Gambar 4. 14 Citra Candi Borobudur original (output).....	54
Gambar 4. 15 Grafik hasil evaluasi citra Candi Borobudur dengan tambahan degradasi blur.....	55

Gambar 4. 16 Citra Candi Borobudur dengan tambahan degradasi blur (input) ..	56
Gambar 4. 17 Citra Candi Borobudur dengan tambahan degradasi blur (output)	56
Gambar 4. 18 Grafik hasil evaluasi Citra Candi Borobudur dengan tambahan degradasi noise.....	57
Gambar 4. 19 Citra Candi Borobudur dengan tambahan degradasi noise (input)	58
Gambar 4. 20 Citra Candi Borobudur dengan tambahan degradasi noise (output)	58
Gambar 4. 21 Grafik hasil evaluasi Citra Candi Borobudur dengan tambahan degradasi resize	59
Gambar 4. 22 Citra Candi Borobudur dengan tambahan degradasi resize (input)	60
Gambar 4. 23 Citra Candi Borobudur dengan tambahan degradasi resize (output)	60
Gambar 4. 24 Grafik hasil evaluasi citra Candi Borobudur dengan tambahan degradasi kompresi.....	61
Gambar 4. 25 Citra Candi Borobudur dengan tambahan degradasi kompresi (input)	62
Gambar 4. 26 Citra Candi Borobudur dengan tambahan degradasi kompresi (output).....	62
Gambar 4. 27 Grafik hasil evaluasi Citra Candi Borobudur dengan tambahan degradasi blur, noise, resize dan kompresi.....	63
Gambar 4. 28 Citra Candi Borobudur dengan tambahan degradasi blur, noise, resize dan kompresi (input)	64
Gambar 4. 29 Citra Candi Borobudur dengan tambahan degradasi blur, noise, resize dan kompresi (output)	64
Gambar 4. 30 Grafik hasil evaluasi citra Menara Kudus Original.....	65
Gambar 4. 31 Citra Menara Kudus Original (input)	66
Gambar 4. 32 Citra Menara Kudus Original (output).....	66
Gambar 4. 33 Grafik hasil evaluasi citra Menara Kudus dengan tambahan degradasi blur	67
Gambar 4. 34 Citra Menara Kudus dengan tambahan degradasi blur (input)	68
Gambar 4. 35 Citra Menara Kudus dengan tambahan degradasi blur (output) ...	68
Gambar 4. 36 Grafik Hasil evaluasi citra Menara Kudus dengan tambahan degradasi noise	69
Gambar 4. 37 Citra Menara Kudus dengan tambahan degradasi noise (input) ...	70
Gambar 4. 38 Citra Menara Kudus dengan tambahan degradasi noise (output) ..	70
Gambar 4. 39 Grafik hasil evaluasi citra Menara Kudus dengan tambahan degradasi resize	71
Gambar 4. 40 Citra Menara Kudus dengan tambahan degradasi resize (input)....	72
Gambar 4. 41 Citra Menara Kudus dengan tambahan degradasi resize (output)..	72
Gambar 4. 42 Grafik hasil evaluasi citra Menara Kudus dengan tambahan degradasi kompresi.....	73
Gambar 4. 43 Citra Menara Kudus dengan tambahan degradasi kompresi (input)	74

Gambar 4. 44 Citra Menara Kudus dengan tambahan degradasi kompresi (output)	74
Gambar 4. 45 Grafik hasil evaluasi Citra Menara Kudus dengan tambahan degradasi blur, noise, resize dan kompresi.....	75
Gambar 4. 46 Citra Menara Kudus dengan tambahan degradasi blur, noise, resize dan kompresi (input)	76
Gambar 4. 47 Citra Menara Kudus dengan tambahan degradasi blur, noise, resize dan kompresi (output)	76
Gambar 4. 48 Halaman upload citra resolusi rendah	77
Gambar 4. 49 Halaman output citra resolusi tinggi	77
Gambar 4. 50 Flowchart Implementasi metode Real-ESRGAN	78
Gambar 4. 51 Grafik rata – rata hasil kinerja model pada setiap epoch	80



DAFTAR TABEL

Tabel 2. 1 State of The Art	8
Tabel 2. 2 State of The Art (Lanjutan)	9
Tabel 4. 1 Rincian arsitektur <i>generator</i>	51
Tabel 4. 2 Rincian arsitektur <i>discriminator</i>	51
Tabel 4. 3 Rincian arsitektur <i>discriminator</i> (lanjutan)	52
Tabel 4. 4 Hasil evaluasi citra Candi Borobudur original.....	53
Tabel 4. 5 Hasil evaluasi citra Candi Borobudur dengan tambahan degradasi blur	55
Tabel 4. 6 Hasil evaluasi Citra Candi Borobudur dengan tambahan degradasi noise	57
Tabel 4. 7 Hasil evaluasi Citra Candi Borobudur dengan tambahan degradasi resize	59
Tabel 4. 8 Hasil evaluasi Citra Candi Borobudur dengan tambahan degradasi kompresi.....	61
Tabel 4. 9 Hasil evaluasi Citra Candi Borobudur dengan tambahan degradasi blur, noise, resize dan kompresi	63
Tabel 4. 10 Hasil evaluasi citra Menara Kudus Original	65
Tabel 4. 11 Hasil evaluasi citra Menara Kudus dengan tambahan degradasi blur	67
Tabel 4. 12 Hasil evaluasi citra Menara Kudus dengan tambahan degradasi noise	69
Tabel 4. 13 Hasil evaluasi citra Menara Kudus dengan tambahan degradasi resize	71
Tabel 4. 14 Hasil evaluasi citra Menara Kudus dengan tambahan degradasi kompresi.....	73
Tabel 4. 15 Hasil evaluasi Citra Menara Kudus dengan tambahan degradasi blur, noise, resize dan kompresi	75
Tabel 4. 16 Hasil rata-rata evaluasi	80

ABSTRAK

Pelestarian citra bangunan bersejarah dalam bentuk digital di Indonesia menghadapi tantangan kualitas visual akibat degradasi seperti *blur*, *noise*, *resize* dan kompresi. Citra dari Leiden University Libraries mengalami degradasi visual akibat proses digitalisasi, usia dokumen, serta keterbatasan teknologi dokumentasi sehingga jenis degradasi pada citra tersebut tidak diketahui secara pasti. Oleh karena itu, dari permasalahan tersebut, penelitian ini mengimplementasikan metode Real-ESRGAN untuk meningkatkan resolusi citra bangunan bersejarah secara realistis tanpa informasi eksplisit jenis degradasinya. Real-ESRGAN merupakan pengembangan dari ESRGAN yang menggabungkan arsitektur RRDB pada *generator* dan U-Net pada *discriminator*. Model ini dilatih menggunakan dataset Flickr2K yang disimulasikan degradasinya dengan *high-order degradation model* agar mirip citra dunia nyata. Evaluasi menggunakan metrik *no-reference* CLIP-IQA menunjukkan model Real-ESRGAN mampu menghasilkan citra resolusi tinggi dengan kualitas perseptual yang baik tanpa memerlukan citra referensi. Skor rata-rata tertinggi CLIP-IQA yang diperoleh sebesar 0,70863 pada *epoch* ke-119. Nilai ini menandakan hasil citra yang dihasilkan memiliki kesesuaian dengan citra berkualitas baik secara visual dan menunjukkan potensi dalam meningkatkan resolusi citra bangunan bersejarah di Indonesia.

Kata Kunci: Real-ESRGAN, *High-Order Degradation Model* dan Citra Bangunan Bersejarah.

ABSTRACT

The preservation of digital images of historical buildings in Indonesia faces visual quality challenges due to degradation such as blur, noise, resizing, and compression. Images from Leiden University Libraries experience visual degradation due to the digitization process, the age of the documents, and the limitations of documentation technology, so the type of degradation in the images is not known with certainty. Therefore, to address these problems, this study implements the Real-ESRGAN method to realistically increase the resolution of historical building images without explicit information on the type of degradation. Real-ESRGAN is a development of ESRGAN that combines the RRDB architecture in the generator and U-Net in the discriminator. This model is trained using the Flickr2K dataset whose degradation is simulated with a high-order degradation model to resemble real-world images. Evaluation using the no-reference CLIP-IQA metric shows that the Real-ESRGAN model is capable of producing high-resolution images with good perceptual quality without the need for reference images. The highest average CLIP-IQA score obtained was 0,70863 at the 119th epoch. This value indicates that the resulting image results are in accordance with good quality images visually and show the potential to increase the resolution of images of historic buildings in Indonesia.

Keywords: Real-ESRGAN, *High-Order Degradation Model*, Historical Building Images.

BAB I

PENDAHULUAN

1.1 Latar Belakang

Pelestarian warisan budaya indonesia dalam bentuk digital merupakan upaya strategis bangsa indonesia untuk menjaga keberlanjutan nilai historis, seni, dan identitas suatu bangsa di tengah perkembangan digitalisasi (Safira dkk., 2020). Hal ini sejalan dengan Pasal 1 UU RI No. 11 Tahun 2010 tentang Cagar Budaya, bahwa cagar budaya adalah warisan budaya bersifat kebendaan berupa benda, bangunan, struktur, situs, dan kawasan cagar budaya yang perlu dilestarikan karena memiliki nilai penting bagi sejarah, ilmu pengetahuan, pendidikan, agama, dan/atau kebudayaan. Melalui pelestarian dalam bentuk digital, berbagai artefak budaya seperti manuskrip kuno, bangunan bersejarah dan karya seni dapat diarsipkan dalam bentuk citra atau video. Dari proses digitalisasi tersebut sehingga generasi selanjutnya dapat mengakses dan mempelajarinya tanpa bergantung pada kondisi fisik aslinya (Agustinova, 2022).

Contoh warisan budaya yang perlu dilestarikan dalam bentuk digital di indonesia adalah foto bangunan bersejarah, mengingat citra visualnya sering kali mengalami penurunan kualitas akibat keterbatasan resolusi, kerusakan fisik, dan proses penyimpanan yang kurang memadai (Agustinova, 2022). Bangunan bersejarah seperti candi, masjid tua, gereja kolonial, dan situs arkeologi lainnya memiliki nilai historis yang tinggi, sehingga dokumentasi visualnya memegang peranan penting dalam pelestarian budaya (Safira dkk., 2020).

Urgensi digitalisasi citra bangunan bersejarah juga diperkuat oleh beberapa penelitian di berbagai wilayah seperti (Darmarastri dkk., 2023) yang melakukan pengkajian terhadap digitalisasi katalog foto bangunan di wilayah Mangkunegaran. Studi tersebut menunjukkan bahwa dokumentasi visual beresolusi tinggi sangat dibutuhkan untuk menghindari kehilangan data arsitektural penting akibat kerusakan fisik pada media arsip konvensional.

Selain itu, (Sarwito dkk., 2024) menggarisbawahi bahwa kompleksitas kerusakan visual pada bangunan kolonial di Semarang membutuhkan solusi berbasis teknologi digital seperti pemodelan 3D dan peningkatan kualitas citra agar informasi spasial dan historis dapat tetap terjaga.

Resolusi rendah pada citra berdampak pada hilangnya kejelasan visual objek, yang umumnya disebabkan oleh degradasi warna, keburaman (*blur*), dan *noise*. Degradasi yang kompleks dan tak terduga seperti kerusakan fisik menjadi suatu permasalahan dalam dokumentasi bangunan bersejarah, karena detail arsitektural yang memiliki nilai historis tinggi dapat terdistorsi atau hilang (Sarwito dkk., 2024). Untuk menjaga representasi visual dalam pelestarian warisan budaya diperlukan peningkatan kualitas citra melalui pendekatan digital seperti *super-resolution* agar memastikan kelestarian informasi visual secara berkelanjutan.

Tantangan utama dalam peningkatan resolusi citra adalah keberagaman bentuk degradasi yang terjadi pada citra dunia nyata. Untuk mengatasi tantangan tersebut, (Wang dkk., 2021) mengembangkan pendekatan *High-Order Degradation Model* dalam penelitiannya. Model ini mensimulasikan proses degradasi secara bertingkat dengan kombinasi *blur*, *downsampling*, *noise*, dan *compression* yang disusun secara acak untuk meniru kondisi kerusakan nyata pada citra. Pendekatan ini memungkinkan model *super-resolution* dilatih dengan data yang lebih representatif sehingga mampu menghasilkan citra yang lebih sesuai terutama pada citra bangunan bersejarah yang sering kali mengalami degradasi yang kompleks.

Salah satu metode yang digunakan untuk meningkatkan resolusi citra adalah *super-resolution*, yaitu teknik yang menghasilkan citra beresolusi tinggi dari citra beresolusi rendah (Rohim dkk., 2024). Berbagai metode telah digunakan untuk implementasi *super-resolution*, seperti *interpolasi bicubic* dan metode berbasis *deep learning*. Penelitian oleh Hardiansyah dkk., 2019 menunjukkan bahwa *interpolasi bicubic* menghasilkan PSNR sebesar 30,40 dB untuk perbesaran 3 kali dan 28,43 dB untuk perbesaran 4 kali, namun kualitas yang dihasilkan masih belum menghasilkan kinerja yang memuaskan

untuk mengatasi masalah resolusi citra dibandingkan metode berbasis *deep learning*. Zhang & Yu, 2022 membandingkan *interpolasi bicubic* dan SRGAN pada citra topografi atau citra permukaan tanah dengan hasilnya menunjukkan nilai RMSE masing-masing sebesar 2,1 untuk *interpolasi bicubic*; 1,8 untuk SRGAN; Hasil pengujian mengindikasikan bahwa SRGAN memberikan performa yang lebih baik dari metode *interpolasi bicubic*. Sementara itu, Nandal dkk., 2024 menguji SRGAN, ESRGAN, dan Real-ESRGAN pada citra medis dan memperoleh PSNR rata-rata sebesar 35 dB untuk SRGAN, 37 dB untuk ESRGAN, dan 38 dB untuk Real-ESRGAN, hasil pengujian menunjukkan metode Real-ESRGAN memiliki kinerja lebih baik dibanding SRGAN dan ESRGAN.

Berdasarkan uraian sebelumnya, permasalahan resolusi citra yang rendah dapat diatasi dengan meningkatkan ukuran dan kualitas resolusi citra. Dalam penelitian ini, digunakan metode Real-ESRGAN (*Real-Enhanced Super-Resolution Generative Adversarial Network*) untuk mengatasi permasalahan tersebut. Real-ESRGAN merupakan pengembangan dari GAN (*Generative Adversarial Network*) yang dirancang untuk menghasilkan citra *super-resolution* dengan kualitas yang lebih realistis dan detail dari *input* gambar beresolusi rendah dengan degradasi yang kompleks dan tidak terduga (Wang dkk., 2021). Mengenai degradasi yang kompleks, dalam penelitian ini menggunakan *high-order degradation model* sebagai tahap *preprocessing data*, pendekatan ini sebagai simulasi degradasi bertingkat yang menggabungkan proses *blur*, *downsampling*, *noise*, dan *compression* secara acak agar mencerminkan kondisi nyata. Didukung pada Metode Real-ESRGAN dapat memperbesar citra hingga empat kali lipat dengan hasil visual yang mendekati citra resolusi tinggi asli sehingga relevan untuk mengatasi permasalahan pada citra bangunan bersejarah sebagai pelestarian warisan budaya dalam bentuk digital di Indonesia.

Dalam mendukung validitas penelitian, model Real-ESRGAN dievaluasi untuk mengukur kinerjanya secara objektif dalam meningkatkan resolusi citra dan agar tidak menghasilkan citra yang terlalu halus (*over-*

smooth) (Xu dkk., 2024), mengandung artefak buatan (*fake textures*) (Ward dkk., 2019), serta tidak terlalu fokus dengan data pelatihan (*overfitting*) (Feng dkk., 2019). Evaluasi model dalam penelitian ini menggunakan metrik objektif *Contrastive Language–Image Pretraining for Image Quality Assessment* (CLIP-IQA) yaitu metode evaluasi kualitas gambar tanpa referensi (*No-Reference*) yang berbasis model CLIP. CLIP-IQA mampu menilai kualitas perseptual gambar secara langsung tanpa data *ground truth* (data asli) karena menggunakan representasi semantik dari model CLIP yang telah dilatih sebelumnya pada pasangan citra dan teks (Wang dkk., 2022). Oleh karena itu, evaluasi dengan CLIP-IQA memberikan pendekatan yang lebih fleksibel dalam mengukur kualitas hasil peningkatan resolusi gambar terutama pada peningkatan citra seperti *super-resolution*.

Meskipun metode Real-ESRGAN telah digunakan dalam penelitian sebelumnya, penelitian ini memiliki perbedaan. Fokus utama dalam penelitian ini adalah peningkatan resolusi pada citra bangunan bersejarah di Indonesia dengan tahap *preprocessing* datanya menggunakan *high-order degradation model* sebagai simulasi degradasi yang kompleks. Selain itu terdapat perbedaan pada dataset yang digunakan dalam melatih data yaitu menggunakan dataset dari dataset modern Flickr2k sebagai pengganti citra bangunan bersejarah resolusi tinggi, mengingat sulitnya memperoleh data asli dengan kualitas tinggi secara konsisten di Indonesia. Dataset modern dipilih untuk melatih model agar tetap dapat menghasilkan citra *super-resolution* yang realistis dan mempertahankan detail penting. Penelitian ini bertujuan mengimplementasikan metode Real-ESRGAN dalam peningkatan resolusi citra bangunan di Indonesia dengan *output* berupa model *deep learning* yang sudah melalui tahap validitas sehingga dapat diterapkan secara langsung dalam *user interface* peningkatan resolusi citra. Penerapan model ini diharapkan memberikan kontribusi nyata dalam pelestarian digital warisan budaya, khususnya dalam meningkatkan kejelasan visual dokumentasi bangunan bersejarah di Indonesia.

Berdasarkan uraian latar belakang permasalahan sebelumnya, maka dirumuskan penelitian “*Implementasi Real-ESRGAN Untuk Peningkatan Resolusi Citra Bangunan Bersejarah Di Indonesia*”. Penelitian ini diharapkan juga dapat membantu memperbaiki kualitas citra menjadi lebih baik. Selain itu, hasil penelitian diharapkan dapat menjadi kontribusi dalam penelitian *super-resolution* sebuah citra.

1.2 Perumusan Masalah

Rumusan masalah dalam penelitian ini adalah sebagai berikut:

- a. Bagaimana menerapkan model Real-ESRGAN untuk meningkatkan resolusi citra bangunan bersejarah di Indonesia?
- b. Bagaimana kinerja model Real-ESRGAN untuk meningkatkan resolusi citra bangunan bersejarah di Indonesia?

1.3 Pembatasan Masalah

Batasan masalah dalam penelitian ini adalah sebagai berikut:

- a. Penelitian ini hanya meningkatkan resolusi pada citra bangunan bersejarah di Indonesia.
- b. Citra yang dipakai untuk *training* data dibatasi 2.000 data untuk pelatihan dan 2 data yang ditambahkan berbagai degradasi menjadi 12 data untuk *testing*.
- c. Penelitian ini dibatasi pada citra *testing* yaitu citra bangunan bersejarah dari arsip Leiden Libraries University yang memiliki degradasi visual, namun tanpa identifikasi spesifik terhadap jenis degradasi yang terjadi.
- d. Metode yang digunakan dalam penelitian ini ialah *super-resolution* hanya menggunakan Real-ESRGAN.
- e. Perbesaran citra yang dipakai hanya satu jenis yaitu *upscale* 4x atau 4 kali lipat dari resolusi sebelumnya.
- f. *Super-resolution* yang dihasilkan tidak merubah foto yang dari hitam putih (*Gray*) kedalam foto yang berwarna (*RGB*).
- g. Sistem dibuat menggunakan bahasa pemrograman python.

- h. Evaluasi kualitas hasil *super resolution* dilakukan menggunakan metrik objektif yaitu *Contrastive Language–Image Pretraining for Image Quality Assessment* (CLIP-IQA).
- i. Proses *super resolution* hanya dilakukan pada citra 2D dan tidak mencakup analisis video atau rekonstruksi 3D.

1.4 Tujuan

Penelitian ini bertujuan untuk:

- a. Merancang dan mengimplementasikan model Real-ESRGAN untuk meningkatkan resolusi pada citra bangunan bersejarah di Indonesia
- b. Mengetahui kinerja model Real-ESRGAN dalam meningkatkan resolusi pada citra bangunan bersejarah di Indonesia.

1.5 Manfaat

Adapun manfaat dari penelitian ini ialah:

- a. Membantu melestarikan warisan budaya dalam bentuk digital untuk mendukung digitalisasi arsip foto bangunan bersejarah di Indonesia.
- b. Membantu meningkatkan resolusi citra bangunan bersejarah di Indonesia
- c. Menjadi rujukan untuk penelitian selanjutnya khususnya terkait topik *super-resolution*, baik menggunakan metode Real-ESRGAN atau pun metode lain dalam meningkatkan resolusi citra.

1.6 Sistematika Penulisan

Untuk mempermudah penulisan tugas akhir ini, penulis menyusun sistematika penulisan yang terdiri dari beberapa bab sebagai berikut:

BAB I : PENDAHULUAN

Bab ini menjelaskan latar belakang pemilihan judul tugas akhir “Implementasi Real-ESRGAN Untuk Peningkatan Resolusi Citra Bangunan Bersejarah Di Indonesia”. Selain itu, bab ini juga memuat rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, dan sistematika penulisan yang disajikan.

BAB II: TINJAUAN PUSTAKA DAN DASAR TEORI

Tinjauan pustaka dan dasar teori membahas berbagai penelitian dan konsep sebelumnya, serta teori yang menjadi landasan dalam pelaksanaan penelitian dan pengembangan sistem. Selain itu, bagian ini juga menguraikan proses analisis kebutuhan yang akan digunakan sebagai acuan dalam pemodelan dan perancangan sistem yang direncanakan.

BAB III: METODE PENELITIAN

Bab ini membahas metode yang digunakan dalam penyusunan tugas akhir, serta proses analisis yang dilakukan. Metodologi penelitian disusun berdasarkan pendekatan *Cross-Industry Standard Process for Data Mining* (CRISP-DM), yang digunakan untuk menganalisis kebutuhan dan sebagai dasar dalam pelaksanaan penelitian.

BAB IV: HASIL DAN ANALISIS PENELITIAN

Bab ini menyajikan hasil dari penelitian yang telah dilakukan secara menyeluruh. Pembahasan dimulai dari proses pelatihan model (*training*), termasuk parameter yang digunakan, waktu pelatihan, dan evaluasi performa model selama pelatihan berlangsung.

BAB V: KESIMPULAN DAN SARAN

Bab ini berisi kesimpulan dari hasil penelitian yang telah dilakukan, isinya disusun atas rumusan masalah yang telah dirumuskan sebelumnya. Selain itu, disampaikan pula beberapa saran yang ditujukan untuk pengembangan dan penyempurnaan penelitian di masa mendatang agar dapat memberikan hasil yang lebih optimal dan relevan.

BAB II

TINJAUAN PUSTAKA DAN DASAR TEORI

2.1 Tinjauan Pustaka

Tinjauan pustaka dipakai untuk mempelajari penelitian terdahulu yang memiliki korelasi dengan penelitian yang akan dilakukan saat ini. Berdasarkan hasil tinjauan pustaka, topik penelitian yang sama ialah meningkatkan resolusi citra akan tetapi penelitian sebelumnya belum ada yang menerapkan metode Real-ESRGAN untuk citra bangunan bersejarah di Indonesia. Hasil dari tinjauan pustaka yang telah didapat adalah sebagai berikut.

Tabel 2. 1 *State of The Art*

No.	Peneliti	Metode	Data	Hasil
1	Bagus Hardiansyah, Aidil Primastya Armin, dan Anton Breva Yunanda. (2019)	<i>Interpolasi Bicubic</i>	<i>Deep Learning Dataset</i>	Hasil dari penelitian ini menunjukkan <i>upscale</i> 3 kali memiliki nilai rata-rata PSNR yang lebih tinggi dibandingkan pada <i>upscale</i> 4 kali
2	Yifan Zhang, Wenhao Yu (2022)	<i>Interpolasi Bicubic</i> , SRGAN dan CEDGAN	Citra Topografi atau Permukaan Tanah	Hasil penelitian menghasilkan model untuk peningkatan citra yang mendapatkan nilai RMSE masing-masing sebesar 2,1 untuk <i>interpolasi bicubic</i> dan 1,8 untuk SRGAN

Tabel 2. 2 *State of The Art* (Lanjutan)

3	Priyanka Nandal, Sudesh Pahal, Ashish Khanna, Placido Rogério Pinheiro (2024)	Real-ESRGAN	Citra Medis	Hasil penelitian memperoleh PSNR rata-rata sebesar 35 dB untuk SRGAN, 37 dB untuk ESRGAN, dan 38 dB untuk Real-ESRGAN
---	---	-------------	-------------	---

Tabel 2.1 dan 2.2 adalah tinjauan pustaka mengenai penelitian terdahulu. Penelitian berkaitan dengan masalah pada penelitian ini yaitu meningkatkan resolusi citra. Pada penelitian terdahulu, terdapat berbagai metode berbeda untuk meningkatkan resolusi citra.

Penelitian pertama memuat penelitian tentang penerapan metode *interpolasi bicubic* untuk rekonstruksi citra pada *super-resolution*. Data yang dipakai menggunakan citra yang didapat dari dataset *deep learning*. Hasil penelitian mendapatkan citra dengan *upscale* 3 kali menghasilkan PSNR sebesar 30,40 dB dan *upscale* 4 kali menghasilkan PSNR sebesar 28,43 dB. sehingga perbesaran 3 kali memiliki nilai rata-rata PSNR yang lebih tinggi dibandingkan pada perbesaran 4 kali (Hardiansyah dkk., 2019).

Penelitian kedua memuat penelitian tentang peningkatan citra topografi atau permukaan tanah. Data yang digunakan untuk penelitian berasal dari *United States Geological Survey* (USGS). Hasil penelitian menghasilkan model untuk peningkatan citra yang mendapatkan nilai RMSE masing-masing sebesar 2,1 untuk *interpolasi bicubic* dan 1,8 untuk SRGAN. Sehingga penelitiannya menyimpulkan model SRGAN adalah model yang paling sesuai (Zhang & Yu, 2022).

Penelitian ketiga memuat penelitian untuk meningkatkan kualitas citra medis. Data yang digunakan dalam penelitian berasal dari brain MRI dataset (BraTS 2020). Hasil pada penelitian ini memperoleh PSNR rata-rata sebesar

35 dB untuk SRGAN, 37 dB untuk ESRGAN, dan 38 dB untuk Real-ESRGAN. Penelitian mendapatkan kesimpulan bahwa Real-ESRGAN merupakan model yang paling sesuai (Nandal dkk., 2024).

Berdasarkan tinjauan pustaka terhadap berbagai penelitian sebelumnya, permasalahan umum yang sering dihadapi dalam pengolahan citra adalah rendahnya resolusi yang berdampak pada hilangnya detail penting dalam gambar. Untuk mengatasi permasalahan tersebut, penelitian ini menerapkan metode Real-ESRGAN. Metode ini dapat meningkatkan resolusi citra hingga 4 (empat) kali lipat dari ukuran aslinya. Namun, berbeda dari penelitian sebelumnya yang umumnya berfokus pada citra umum seperti wajah manusia, lanskap, atau objek sehari-hari, penelitian ini berfokus pada citra bangunan bersejarah di Indonesia sebagai objek utama dan menggunakan *high-order degradation model* untuk mensimulasikan citra yang mengalami degradasi yang kompleks. Citra bangunan bersejarah ini memiliki nilai historis yang tinggi sehingga dokumentasi visualnya memegang peranan penting dalam pelestarian budaya tetapi citranya sering kali mengalami penurunan kualitas akibat keterbatasan resolusi, kerusakan fisik, dan proses penyimpanan yang kurang memadai sehingga membutuhkan pendekatan khusus agar detail tersebut dapat dipertahankan dan ditingkatkan dalam proses peningkatan resolusi.

2.2 Dasar Teori

2.2.1 Citra Bangunan Bersejarah Di Indonesia

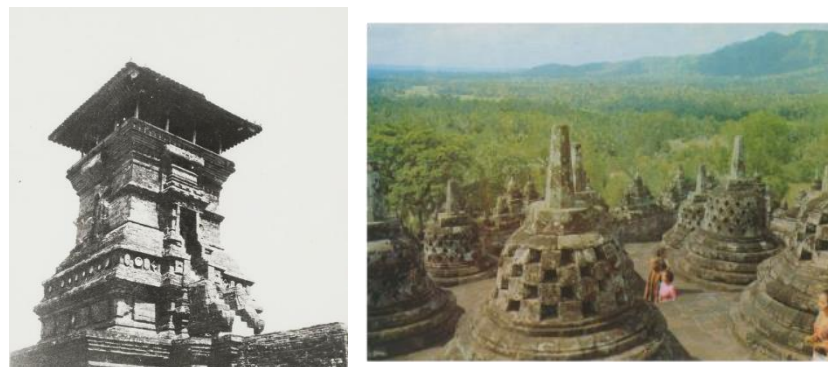
Bangunan bersejarah merupakan bagian penting dari warisan budaya yang mencerminkan nilai historis, arsitektural, sosial, dan identitas suatu bangsa (Safira dkk., 2020). Dalam konteks pelestarian budaya, dokumentasi visual bangunan bersejarah menjadi upaya strategis untuk menjaga keberlanjutan nilai-nilai tersebut di tengah perkembangan teknologi digital (Safira dkk., 2020).

Pelestarian bangunan bersejarah dalam bentuk citra digital tidak hanya bertujuan untuk menyimpan informasi visual, tetapi juga untuk mempermudah akses generasi mendatang terhadap warisan budaya tanpa

ketergantungan pada kondisi fisik aslinya. Digitalisasi artefak budaya seperti manuskrip kuno, karya seni, dan bangunan bersejarah memungkinkan warisan tersebut diarsipkan secara lebih aman dan efisien (Agustinova, 2022). Hal ini juga sejalan dengan Pasal 1 Undang-Undang Republik Indonesia Nomor 11 Tahun 2010 tentang Cagar Budaya, yang menyatakan bahwa cagar budaya adalah warisan budaya bersifat kebendaan berupa benda, bangunan, struktur, situs, dan kawasan cagar budaya yang memiliki nilai penting bagi sejarah, ilmu pengetahuan, pendidikan, agama, dan kebudayaan.

Namun, dalam dokumentasi citra bangunan bersejarah sering kali menghadapi tantangan kualitas visual seperti resolusi rendah, *blur*, *noise*, dan kerusakan fisik pada arsip lama (Agustinova, 2022). Permasalahan ini menyebabkan detail arsitektural yang memiliki nilai historis tinggi rentan hilang atau terdistorsi, sehingga menghambat proses pelestarian secara akurat dan berkelanjutan (Agustinova, 2022). Hal ini juga ditemukan pada citra bangunan bersejarah di Indonesia dari Leiden University Libraries telah mengalami degradasi visual akibat usia dokumen dan proses digitalisasi. Meskipun degradasi jelas secara visual, jenis degradasinya tidak dapat diidentifikasi secara pasti. Oleh karena itu, diperlukan pendekatan lanjutan seperti peningkatan kualitas citra berbasis kecerdasan buatan yang mampu menangani degradasi tanpa memerlukan identifikasi jenis degradasinya.

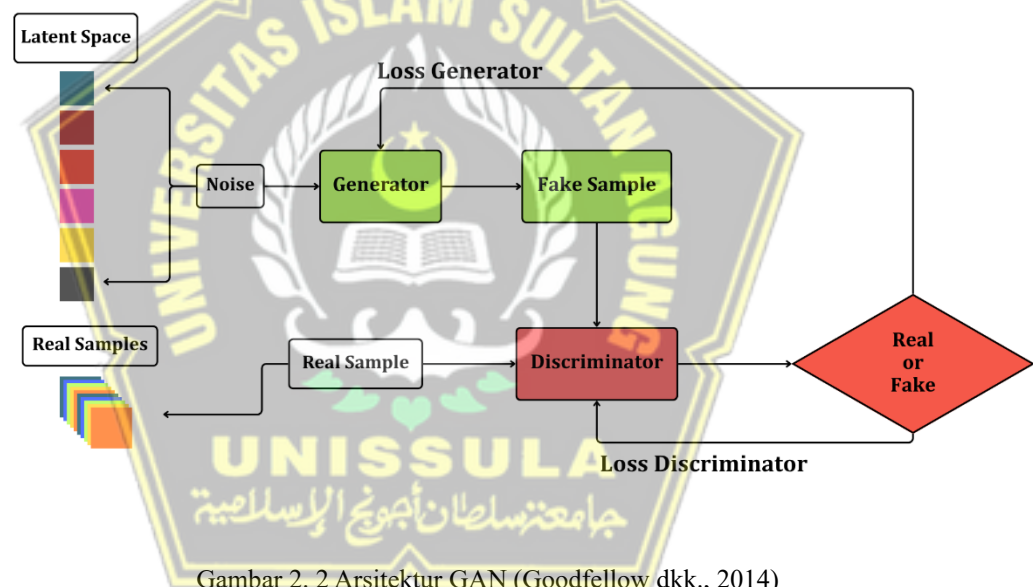
Berikut ini contoh citra bangunan bersejarah yang terdapat di Indonesia, diperoleh dari koleksi digital milik Leiden University Libraries sebagai salah satu sumber dokumentasi sejarah visual terkait Indonesia.



Gambar 2. 1 Contoh citra bangunan bersejarah (Leiden University Libraries abad 19)

2.2.2 Generative Adversarial Network (GAN)

Metode *Generative Adversarial Network* (GAN) pertama kali diperkenalkan oleh Ian Goodfellow dkk bersama rekannya pada tahun 2014 melalui penelitian berjudul "*Generative Adversarial Nets*". Dalam publikasi tersebut, penulis mengusulkan pendekatan baru dalam pembelajaran mesin generatif, di mana dua jaringan saraf buatan (*neural network*) saling bersaing dalam kerangka permainan dua pihak (*game theory*) (Goodfellow dkk., 2014). *Generative Adversarial Network* (GAN) adalah salah satu arsitektur jaringan saraf generatif yang diperkenalkan sebagai pendekatan inovatif untuk menghasilkan data baru yang menyerupai distribusi data asli.



Gambar 2. 2 Arsitektur GAN (Goodfellow dkk., 2014)

Pada gambar 2.2 merupakan arsitektur dari *Generative Adversarial Network* (GAN) di atas, *input* berupa *noise* acak dari ruang laten diubah oleh *generator* menjadi data sintetis. *Discriminator* kemudian menerima dua jenis data yaitu data asli dan data hasil *generator*, lalu memberikan probabilitas apakah data tersebut asli atau palsu. Proses ini berlangsung secara iteratif dan membentuk kompetisi yang saling meningkatkan performa kedua jaringan.

Adapun beberapa komponen dari *Generative Adversarial Networks* (GAN) yaitu *generator*, *discriminator*, dan beberapa komponen lainnya dapat dijelaskan sebagai berikut:

1. *Generator*

Generator adalah jaringan saraf yang memetakan *noise* acak dari ruang laten ke ruang data sintetis. Tujuannya adalah menghasilkan data *generator* yang secara statistik dan visual menyerupai data nyata. Arsitektur *generator* biasanya terdiri dari beberapa lapisan *fully connected* atau *convolutional transpose layer* tergantung pada jenis data yang digunakan.

2. *Discriminator*

Discriminator berperan sebagai klasifikator biner yang menerima *input* berupa data nyata maupun data sintetis dan memberikan keluaran probabilitas bahwa *input* tersebut asli. Arsitekturnya biasanya terdiri dari lapisan *convolutional* dan *fully connected*, tergantung pada kompleksitas data.

3. Komponen Tambahan

Selain dua jaringan utama, beberapa komponen tambahan penting dalam GAN meliputi:

- a. Ruang Laten (*Latent Space*): Merupakan vektor acak berdimensi rendah yang menjadi masukan bagi *generator*. Variasi pada ruang laten berpengaruh besar terhadap keanekaragaman dan kualitas data sintetis yang dihasilkan. *Latent space* biasanya mengikuti distribusi normal standar atau distribusi seragam.
- b. Fungsi Regularisasi dan Varian GAN: Dalam praktiknya, pelatihan GAN sering kali menghadapi masalah seperti ketidakstabilan gradien dan *mode collapse*. Oleh karena itu, varian GAN seperti Wasserstein GAN (WGAN), Least Squares GAN (LSGAN), dan WGAN-GP diperkenalkan dengan fungsi kerugian dan teknik regularisasi baru.

Secara keseluruhan, *Generative Adversarial Network* (GAN) merupakan pendekatan revolusioner dalam pembelajaran mesin yang

memungkinkan pembangkitan data sintetis berkualitas tinggi melalui proses pelatihan dua jaringan saraf yang saling berkompetisi. Dengan *generator* yang berfokus pada penciptaan data baru dan *discriminator* yang berperan sebagai pengevaluasi. GAN dapat menghasilkan data yang semakin realistis seiring proses pelatihan. Pendekatan ini telah banyak diadopsi dalam berbagai aplikasi seperti pemrosesan citra, peningkatan resolusi, dan restorasi citra, serta menjadi dasar bagi banyak arsitektur generatif modern.

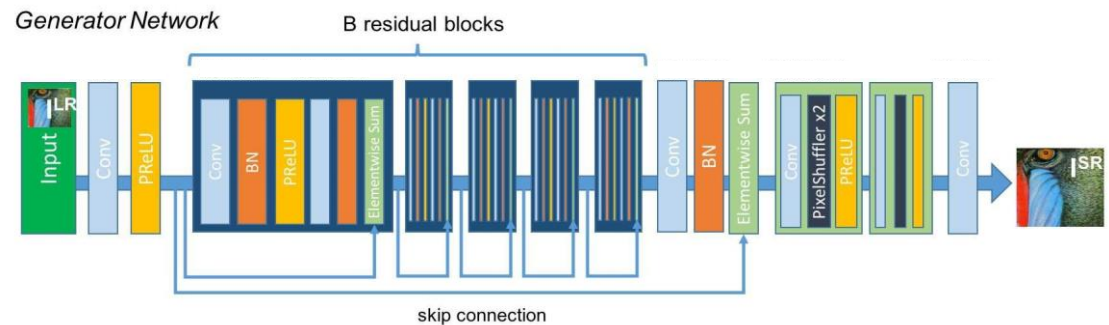
2.2.3 Super-Resolution Generative Adversarial Network (SRGAN)

Metode *Super-Resolution Generative Adversarial Network* (SRGAN) pertama kali diperkenalkan oleh Christian Ledig dan rekan-rekannya pada tahun 2017. SRGAN merupakan pengembangan dari konsep *Generative Adversarial Network* (GAN) yang sebelumnya diperkenalkan oleh Ian Goodfellow dkk. pada tahun 2014. *Super-Resolution Generative Adversarial Network* (SRGAN) merupakan metode yang dapat menghasilkan gambar realistis secara alami untuk faktor perbesaran 4 kali (Ledig dkk., 2017). Metode ini didasarkan pada metode *Generative Adversarial Network* (GAN).

Metode SRGAN seperti halnya pada metode GAN memiliki dua jenis model yaitu model *generator* dan *discriminator*. Model *generator* menangkap distribusi data untuk membuat sampel baru, sedangkan model *discriminator* untuk membandingkan sampel data dengan sampel baru dari *generator*. Pada SRGAN model *generator* dapat membuat gambar baru dari gambar resolusi rendah. Kemudian, *discriminator* dapat membanding gambar resolusi tinggi dengan gambar yang dihasilkan dari *generator*. Tujuan dari SRGAN ini membuat *discrimintor* tidak dapat membedakan gambar asli dengan gambar yang dihasilkan *generator*. Sehingga, metode ini dapat menghasilkan gambar resolusi tinggi yang realistis.

Adapun beberapa komponen dari *Super-Resolution Generative Adversarial Networks* (SRGAN) yaitu *generator*, *discriminator*, dan *perceptual loss function* dapat dijelaskan sebagai berikut:

1. Generator



Gambar 2.3 Arsitektur *Generator* SRGAN (Ledig dkk., 2017)

Gambar 2.3 merupakan arsitektur dari *generator* SRGAN. *Generator* pada SRGAN berfungsi untuk mengubah citra beresolusi rendah menjadi citra beresolusi tinggi yang tampak realistis. Proses ini diawali dengan ekstraksi fitur dari citra *input* menggunakan lapisan konvolusi dan aktivasi PReLU yaitu fungsi yang membantu jaringan mengenali pola kompleks. Di bagian tengah, terdapat beberapa *residual block*, yaitu rangkaian lapisan konvolusi yang memiliki jalur pintas (*skip connection*) agar informasi penting dari input tidak hilang saat diproses berulang kali. Setelah itu, hasil dari blok-blok tersebut digabungkan kembali dengan fitur awal melalui penjumlahan elemen demi elemen (*elementwise sum*) untuk menjaga detail citra asli. Untuk memperbesar ukuran citra, digunakan teknik *PixelShuffle* yaitu metode yang secara efisien menyusun ulang fitur dari dimensi saluran (*channel*) menjadi dimensi spasial (tinggi dan lebar) sehingga citra menjadi lebih besar tanpa merusak detail. Proses aliran data pada *generator* dilakukan secara *feed-forward*, yaitu data mengalir maju dari input hingga output tanpa perulangan, dan bobot jaringan diperbarui berdasarkan kesalahan dari hasil keluaran melalui algoritma *backpropagation*. Terakhir, lapisan konvolusi menghasilkan citra beresolusi tinggi sebagai *output* akhir.

2. Discriminator

Discriminator Network



Gambar 2.4 Arsitektur *Discriminator* SRGAN (Ledig dkk., 2017)

Pada Gambar 2.4 ialah model *discriminator* untuk melatih *generator* secara *adversarial*. *Discriminator* pada SRGAN berfungsi sebagai jaringan yang membedakan antara citra beresolusi tinggi hasil rekonstruksi *generator* (SR) dan citra asli resolusi tinggi (HR). Proses dimulai dengan memasukkan citra ke dalam beberapa lapisan konvolusi yang bertugas mengekstraksi fitur visual dari gambar. Setiap lapisan konvolusi diikuti oleh fungsi aktivasi *Leaky ReLU*, yaitu varian dari fungsi *ReLU* yang masih mengizinkan aliran informasi meskipun nilai *input* negatif sehingga mengurangi risiko hilangnya informasi selama proses pelatihan. Beberapa lapisan juga dilengkapi dengan *batch normalization*, yaitu teknik yang menstabilkan dan mempercepat pelatihan dengan cara menormalkan nilai *output* dari lapisan sebelumnya. Seiring berjalannya proses, ukuran spasial dari citra diperkecil secara bertahap (melalui *stride 2*), sementara jumlah fitur yang dipelajari meningkat, memungkinkan jaringan mengenali pola-pola kompleks pada citra. Setelah melewati seluruh lapisan konvolusi, data diproses oleh lapisan *dense* atau *fully connected* yang berfungsi menggabungkan semua informasi fitur menjadi keputusan akhir. Di tahap akhir, aktivasi *sigmoid* digunakan untuk menghasilkan nilai antara 0 dan 1 yang merepresentasikan probabilitas apakah citra yang diberikan merupakan citra asli (HR) atau hasil buatan *generator* (SR). Dengan arsitektur ini, *discriminator* melatih *generator* secara tidak langsung melalui *adversarial training* yaitu dengan terus mencoba membedakan

mana gambar yang asli dan mana yang palsu, sehingga mendorong *generator* untuk menghasilkan citra yang semakin realistis.

3. *Perceptual Loss Function*

Perceptual loss function adalah metode untuk mengukur seberapa mirip hasil citra dengan citra asli berdasarkan nilai piksel dan kemiripan visual. Pada SRGAN, fungsi ini sangat penting karena membantu *generator* menghasilkan citra yang akurat secara struktur melalui *content loss* dan tampak realistis melalui *adversarial loss*, sehingga citra yang dihasilkan lebih tajam dan menyerupai gambar asli.

a. *Content Loss*

Content loss pada SRGAN digunakan untuk memastikan bahwa hasil citra yang dihasilkan oleh *generator* memiliki struktur dan konten yang mirip dengan citra asli resolusi tinggi. SRGAN menggunakan dua pendekatan dalam menghitung *content loss*, yaitu *Mean Squared Error* (MSE) dan *VGG loss*. MSE menghitung perbedaan nilai piksel antara citra hasil dan citra asli, namun sering kali menghasilkan gambar yang terlalu halus dan kurang detail. Oleh karena itu, digunakan juga *VGG loss*, yaitu metode yang membandingkan fitur visual dari dua gambar berdasarkan ekstraksi fitur dari jaringan VGG. Dengan memanfaatkan *VGG loss*, SRGAN mampu menilai kemiripan secara *perceptual* sehingga hasil citra menjadi lebih tajam dan menyerupai gambar nyata di mata manusia.

b. *Adversarial Loss*

Adversarial loss pada SRGAN digunakan untuk mendorong *generator* menghasilkan citra *super-resolution* yang lebih realistis dan sulit dibedakan dari citra asli oleh *discriminator*. Fungsi ini bekerja dalam kerangka kerja GAN (*Generative Adversarial Network*), di mana *generator* dan *discriminator* saling berkompetisi. *Adversarial loss* ditambahkan ke dalam total *loss* untuk memperbaiki kualitas visual hasil dari *generator*. *adversarial loss* didefinisikan sebagai tingkat keberhasilan *discriminator* dalam

membedakan citra hasil *generator* dengan citra asli. Semakin baik *generator* menipu *discriminator* maka semakin kecil nilai *loss*-nya. Dengan cara ini, *adversarial loss* berperan penting dalam meningkatkan hasil citra yang dihasilkan oleh SRGAN.

2.2.4 Enhanced Super-Resolution Generative Adversarial Networks (ESRGAN)

Enhanced Super-Resolution Generative Adversarial Networks (ESRGAN) merupakan metode peningkatan resolusi citra berbasis *deep learning* yang dikembangkan oleh Xintao Wang dan rekan-rekannya pada tahun 2018. ESRGAN merupakan pengembangan dari SRGAN dengan peningkatan arsitektur dan fungsi *loss* agar menghasilkan citra *super-resolusi* yang lebih tajam dan realistis (Wang dkk., 2018). Secara umum, ESRGAN terdiri dari dua jaringan utama yaitu *generator* dan *discriminator*, yang dilatih secara *adversarial* menggunakan *Generative Adversarial Networks* (GAN).

Adapun beberapa komponen dari *Enhanced Super-Resolution Generative Adversarial Networks* (ESRGAN) yaitu *generator*, *discriminator*, dan *perceptual loss function* dapat dijelaskan sebagai berikut:

1. *Generator*



Gambar 2. 5 Arsitektur *Generator* ESRGAN (Wang dkk., 2018)

Pada gambar 2.5 ialah arsitektur *generator* pada ESRGAN. *Generator* pada *Enhanced Super-Resolution Generative Adversarial Networks* (ESRGAN) merupakan bagian dari jaringan yang bertugas mengubah citra resolusi rendah menjadi citra resolusi tinggi. ESRGAN menggunakan arsitektur dasar dari SRResNet, SRResNet atau *Super-Resolution Residual Network* merupakan metode peningkatan resolusi citra berbasis jaringan konvolusional dalam *deep convolutional neural network* yang dikembangkan oleh Christian Ledig dan rekan-rekannya pada tahun 2017 (Ledig dkk., 2017). Dalam pengembangannya,

Average Discriminator yang tidak hanya menilai keaslian citra secara terpisah, melainkan membandingkan seberapa realistis citra asli dibanding citra hasil *generator*. Dengan kata lain, *discriminator* dalam ESRGAN berusaha memprediksi apakah citra asli lebih *realistic* dibanding citra buatan, bukan hanya apakah citra tersebut *real* atau *fake*.

3. *Perceptual Loss*

Perceptual loss function pada ESRGAN adalah metode untuk mengukur kualitas hasil citra berdasarkan persepsi visual manusia, bukan hanya kesamaan nilai piksel. ESRGAN mengembangkan pendekatan ini dari SRGAN agar hasil citra *super-resolution* tidak hanya akurat secara konten tetapi juga tampak lebih tajam, natural, dan detail. Fungsi *loss* ini terdiri dari beberapa komponen utama, yaitu *content loss* dan *adversarial loss* yang dimodifikasi untuk menghasilkan detail visual yang lebih kaya dan realistis.

a. *Content Loss*

Berbeda dari SRGAN yang menggunakan *Mean Squared Error* (MSE) sebagai dasar *content loss*, ESRGAN menggantinya dengan *L1 loss*. *L1 loss* dinilai lebih efektif dalam menjaga ketajaman detail visual karena tidak terlalu memberi nilai kesalahan pada perbedaan kecil antar piksel. Penggunaan *L1 loss* membantu jaringan mempertahankan struktur citra secara umum tanpa menghasilkan tampilan yang terlalu halus. Dengan demikian, citra hasil ESRGAN tampak lebih tajam dan alami dibandingkan jika menggunakan MSE.

b. *Adversarial Loss*

ESRGAN memperkenalkan *Relativistic Adversarial Loss*, yaitu pengembangan dari *adversarial loss* pada SRGAN. Pada metode ini, *discriminator* tidak hanya menilai apakah gambar palsu itu realistis, tetapi juga melakukan pengecekan gambar asli lebih realistis tidak dibanding gambar palsu. Dengan pendekatan *relativistic discriminator*, *generator* terdorong untuk menghasilkan citra yang

kualitasnya tinggi dan secara relatif lebih realistis dibanding citra palsu sebelumnya. Hal ini membuat *generator* ESRGAN mampu menghasilkan gambar *super-resolution* dengan tekstur yang lebih meyakinkan.

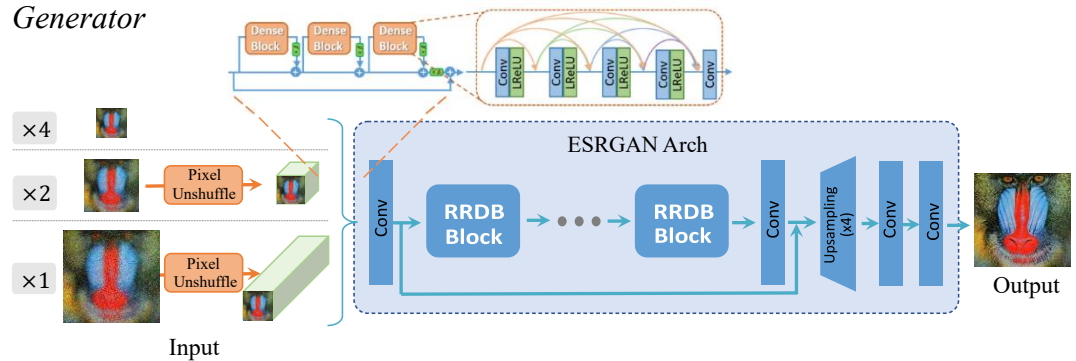
2.2.5 Real Enhanced Super Resolution Generative Adversarial Network (Real ESRGAN)

Real ESRGAN (*Real Enhanced Super Resolution Generative Adversarial Network*) adalah pengembangan dari metode ESRGAN yang ditujukan untuk mengatasi masalah *blind super-resolution*, yaitu peningkatan resolusi gambar dari citra resolusi rendah (*low resolution*) ke citra resolusi tinggi (*high resolution*) tanpa mengetahui bentuk degradasi yang terjadi. Gambar resolusi rendah dalam kasus nyata tidak hanya dihasilkan oleh proses *downsampling* biasa, namun juga kombinasi kompleks dari *blur*, *noise*, dan kompresi JPEG.

Metode Real-ESRGAN, seperti halnya metode GAN pada umumnya, terdiri dari dua komponen utama yaitu *generator* dan *discriminator*. *Generator* berfungsi untuk mempelajari distribusi data dan menghasilkan sampel baru, sedangkan *discriminator* bertugas membedakan antara sampel asli dan hasil buatan *generator*. Dalam Real-ESRGAN, *generator* digunakan untuk merekonstruksi gambar resolusi tinggi dari gambar beresolusi rendah. Selanjutnya, *discriminator* akan mengevaluasi sejauh mana gambar hasil rekonstruksi menyerupai gambar resolusi tinggi yang asli. Tujuan akhirnya adalah membuat gambar hasil *generator* sedemikian realistis hingga tidak dapat dibedakan dari gambar asli oleh *discriminator*. Dengan demikian, Real-ESRGAN mampu menghasilkan gambar beresolusi tinggi yang tampak alami.

Adapun beberapa komponen dari *Real Enhanced Super-Resolution Generative Adversarial Networks* (Real-ESRGAN) yaitu *Generator*, *Discriminator*, dan *Loss Function* dapat dijelaskan sebagai berikut:

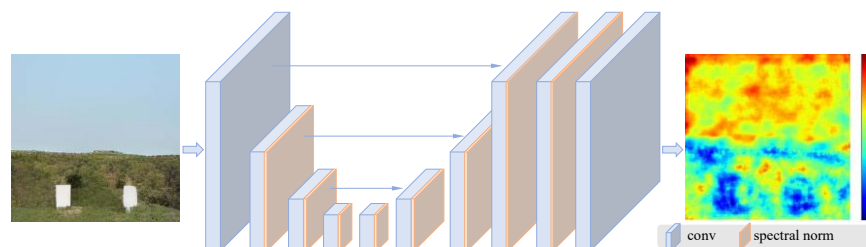
1. Generator



Gambar 2.8 Arsitektur *Generator* Real ESRGAN (Wang dkk., 2021)

Gambar 2.8 merupakan arsitektur dari *generator* pada metode Real-ESRGAN. *Generator* dilatih untuk dapat menghasilkan citra resolusi tinggi berdasarkan masukan citra resolusi rendah (Sulaeman, 2022). Pada Gambar tersebut. Pertama, citra resolusi rendah diproses oleh fungsi *pixel unshuffle* untuk menguraikan citra resolusi rendah menjadi format yang lebih padat dan detail secara *channel*. Kemudian dilanjutkan fungsi konvolusi awal untuk mengubah data *input* menjadi representasi fitur awal agar bisa diproses lebih lanjut oleh jaringan. Setelah itu, data dikirim ke dalam serangkaian *block Residual-in-Residual Dense Block* (RRDB). Hasil dari semua *block* RRDB dijumlahkan bersama dengan konvolusi awal untuk menjaga kestabilan informasi awal. Selanjutnya, data diperbesar ukurannya sesuai dengan resolusi yang diinginkan. Terakhir, data dikirimkan ke konvolusi akhir untuk menyempurnakan hasil akhir sebelum menjadi *output* akhir yaitu citra resolusi tinggi dengan kualitas visual yang lebih baik dan realistis.

2. Discriminator



Gambar 2.9 *Discriminator* Real ESRGAN (Wang dkk., 2021)

Pada gambar 2.9 merupakan model *discriminator*, Real-ESRGAN menggantikan *VGG-style discriminator* dengan arsitektur *U-Net* yang dilengkapi dengan *skip connections*. *Discriminator* ini bertugas membedakan citra hasil *generator* dengan citra asli, dan memberikan sinyal gradien kepada *generator* untuk meningkatkan kualitas hasil.

Gambar tersebut tersebut dimulai dari *input* yang masuk ke lapisan konvolusi pertama lalu diproses menggunakan *spectral normalization* yaitu teknik untuk menjaga stabilitas pelatihan dengan membatasi nilai spektral tertinggi dari bobot konvolusi. Hasil dari lapisan awal ini kemudian dimasukkan ke dalam fungsi *residual* atau *skip connection*, yang berfungsi menghubungkan hasil dari lapisan awal ke lapisan yang lebih dalam untuk mempertahankan informasi penting. Proses ini berlanjut hingga ke lapisan konvolusi terakhir yang seluruhnya tetap menggunakan *spectral normalization* untuk menjaga konsistensi dan kestabilan. Akhirnya, keluaran dari seluruh proses ini masuk ke fungsi akhir untuk menghasilkan nilai probabilitas atau skor yang menunjukkan seberapa realistis citra tersebut.

3. *Loss Function*

Real-ESRGAN menggunakan beberapa jenis fungsi *loss* untuk melatih model agar mampu menghasilkan citra *super-resolution* yang tajam, detail, dan terlihat alami. Fungsi-fungsi *loss* ini bekerja sama untuk mengarahkan model dalam memperbaiki gambar. Ada empat jenis *loss* utama yang digunakan, yaitu *pixel loss*, *perceptual loss*, *adversarial loss*, dan *GAN loss*.

a. *Pixel loss*

Pixel loss adalah fungsi yang membandingkan hasil citra dengan citra aslinya berdasarkan nilai piksel satu per satu. Real-ESRGAN menggunakan *L1 loss*, yaitu menghitung selisih langsung antar piksel. *L1 loss* lebih baik dalam menjaga ketajaman gambar karena tidak terlalu sensitif terhadap perbedaan kecil seperti *MSE*, sehingga gambar yang dihasilkan tidak tampak terlalu halus.

b. *Perceptual Loss*

Perceptual loss digunakan untuk menilai kemiripan antara citra hasil dan citra asli berdasarkan tampilan visual, bukan hanya nilai piksel. Dalam Real-ESRGAN, fitur visual ini diambil dari jaringan VGG, yaitu model komputer yang dilatih untuk mengenali gambar. Fitur diambil sebelum proses aktivasi agar lebih peka terhadap detail seperti tepi, tekstur, dan bentuk.

c. *Adversarial Loss*

Adversarial loss digunakan agar citra hasil tampak lebih realistis. Ini dilakukan dengan cara membuat *generator* dan *discriminator* saling bersaing seperti dalam sistem GAN. *Generator* berusaha menghasilkan gambar yang sulit dibedakan dari gambar asli, sementara *discriminator* mencoba membedakannya. Pada Real-ESRGAN, digunakan pendekatan *relativistic*, yaitu *discriminator* tidak hanya menilai gambar satu per satu, tapi membandingkan mana yang lebih realistis. Ini membuat hasil gambar lebih alami.

d. *GAN Loss*

Loss ini membantu menjaga stabilitas pelatihan GAN. Alih-alih hanya melihat hasil akhir dari *discriminator*, *loss* ini membandingkan fitur internal (fitur menengah) yang dihasilkan oleh *discriminator* saat memproses gambar asli dan gambar hasil. Dengan cara ini, *generator* belajar meniru struktur internal gambar asli, sehingga hasil akhirnya lebih halus dan konsisten.

2.2.6 Python

Bahasa pemrograman Python dirilis pertama kali pada tahun 1991 oleh Guido van Rossum di Scitching Mathematisch Centrum Belanda. Nama python sendiri diambil dari grup komedi Inggris yang bernama Monty Python karena Guido merupakan penggemar grup komedi tersebut (Rossum & Boer, 1991).

Python merupakan bahasa pemrograman yang bersifat *open-source* dengan sebagian besar versinya menggunakan lisensi *GFL-compatible*.

Model proses CRISP-DM memberikan kerangka kerja umum mengenai siklus hidup dalam proyek data mining. Siklus hidup ini terdiri dari enam tahap utama sebagaimana ditunjukkan pada Gambar 2.10. Urutan dari tahap-tahap tersebut bersifat fleksibel, yang berarti alur proses dapat bergerak maju maupun mundur antar fase. Hal ini dikarenakan hasil dari satu tahap dapat memengaruhi pemilihan langkah berikutnya. Panah dalam gambar menunjukkan hubungan dan ketergantungan paling signifikan antar fase. Sementara itu, lingkaran luar dalam gambar mengindikasikan bahwa proses data mining, atau dalam konteks yang lebih luas *data science*, tidak berhenti pada tahap *deployment*. Sebaliknya, proses ini bersifat berkelanjutan pada hasil dari tahap sebelumnya dapat dipelajari kembali untuk menghasilkan solusi yang lebih optimal.

Adapun penjelasan dari keenam tahap dalam model CRISP-DM adalah sebagai berikut:

- a. *Business Understanding*
Tahap awal ini berfokus pada pemahaman terhadap tujuan dan kebutuhan dari perspektif bisnis. Tujuannya adalah untuk merumuskan permasalahan data mining atau *data science* secara jelas serta menyusun rencana awal untuk mencapai target bisnis yang telah ditetapkan.
- b. *Data Understanding*
Tahap ini dimulai dengan pengumpulan data awal, lalu dilanjutkan dengan eksplorasi data untuk memahami struktur dan kualitas data. Proses ini juga bertujuan untuk menemukan wawasan awal dan mengidentifikasi data yang berpotensi untuk pengembangan hipotesis.
- c. *Data Preparation*
Fase ini mencakup serangkaian aktivitas yang diperlukan untuk membentuk dataset akhir yang akan digunakan dalam tahap pemodelan. Prosesnya meliputi pemilihan data, pembersihan, transformasi, serta pemilihan fitur. Proses ini sering dilakukan secara iteratif dan tidak harus dalam urutan tertentu.

d. *Modeling*

Pada tahap ini, berbagai teknik pemodelan dipilih dan diterapkan pada data. Parameter model akan disesuaikan hingga diperoleh hasil yang optimal. Karena beberapa metode memiliki persyaratan khusus terhadap struktur data, sering kali perlu dilakukan revisi kembali pada tahap persiapan data.

e. *Evaluation*

Evaluasi menyeluruh dilakukan untuk memastikan bahwa hasil pemodelan sudah selaras dengan tujuan bisnis. Langkah untuk meninjau kembali keseluruhan proses dan memastikan tidak ada aspek penting yang terlewat.

f. *Deployment*

Bergantung pada kebutuhan organisasi, fase *deployment* bisa sesederhana menghasilkan laporan, atau cukup kompleks seperti menerapkan model dalam sistem operasional perusahaan secara berkelanjutan. Pada tahap ini, model mulai diintegrasikan dalam aplikasi nyata untuk memberikan solusi atas permasalahan bisnis yang ada.

2.2.8 Contrastive Language-Image Pretraining Image Quality Assessment

CLIP-IQA

Contrastive Language-Image Pretraining Image Quality Assessment (CLIP-IQA) merupakan pendekatan dalam penilaian kualitas citra tanpa referensi (*no-reference*) yang memanfaatkan representasi gambar dari model CLIP. Model CLIP sendiri merupakan hasil pelatihan berskala besar yang menghubungkan pasangan gambar dan teks sehingga mampu merepresentasikan aspek visual dan semantik secara bersamaan (Wang dkk., 2022). CLIP-IQA mengevaluasi kualitas citra bukan hanya dari aspek numerik seperti piksel melainkan berdasarkan kesan visual dan estetika yang mendekati persepsi manusia (Wang dkk., 2022). Dengan demikian CLIP-IQA menghadirkan pendekatan baru dalam menilai kualitas citra secara fleksibel tanpa bergantung pada citra referensi.

Dalam pendekatan CLIP-IQA, penilaian dilakukan dengan membandingkan fitur visual dari gambar yang diuji terhadap fitur representatif dari kumpulan gambar dengan kualitas tinggi. Pendekatan melakukan *distance computation* antara fitur gambar dan *anchor representations* dari gambar berkualitas tinggi sehingga model mampu mengkuantifikasi kualitas gambar secara efektif. Proses ini dilakukan pada ruang laten CLIP berformat array sehingga evaluasi mempertimbangkan struktur visual yang ditangkap.

$$s = \frac{x \odot t}{||x|| \cdot ||t||} \quad (1)$$

Pendekatan dasar dalam CLIP-IQA untuk menilai kualitas visual gambar adalah memanfaatkan kemiripan antara representasi fitur gambar dan teks dalam ruang laten. Penilaian ini dihitung menggunakan rumus *cosine similarity* yaitu ukuran kesamaan arah antara dua vektor. Dalam konteks ini, gambar yang telah di *encode* oleh model CLIP direpresentasikan sebagai vektor fitur x , sementara prompt teks (misalnya "*a good photo*") direpresentasikan sebagai vektor t . Nilai kesamaan antara keduanya dihitung menggunakan rumus sehingga nilai skor s ini berada dalam rentang 0 hingga 1 dan merefleksikan seberapa besar gambar tersebut "cocok" atau "relevan" terhadap deskripsi tekstual yang diberikan. Pendekatan ini memungkinkan model CLIP untuk memberikan penilaian kualitas berdasarkan pemahaman semantik terhadap visual dan bahasa tanpa memerlukan referensi gambar asli.

BAB III

METODE PENELITIAN

3.1 *Business Understanding*

Kualitas citra salah satunya dipengaruhi oleh tingkat resolusi. Pada citra resolusi tinggi, citra akan semakin detail. Sebaliknya, pada resolusi rendah detailnya akan semakin kecil. Resolusi rendah menyebabkan citra tidak terlihat dengan jelas. Hal ini disebabkan adanya degradasi pada citra seperti degradasi warna, *blur* (buram) atau pun *noise*. Oleh karena itu, informasi pada citra tidak dapat diekstraksi dengan baik oleh algoritma pemrosesan citra. Berdasarkan faktor tersebut, untuk meningkatkan kualitas citra perlu memperbesar resolusi citra.

Resolusi citra yang rendah dapat diatasi dengan memperbesar ukuran resolusi citra. Permasalahan tersebut dengan metode Real-ESRGAN. Metode ini dapat memperbesar resolusi citra hanya dengan satu *input* citra resolusi rendah. Berdasarkan penelitian sebelumnya, metode ini cukup baik dalam meningkatkan resolusi citra dibandingkan metode tradisional seperti *interpolasi bicubic*.

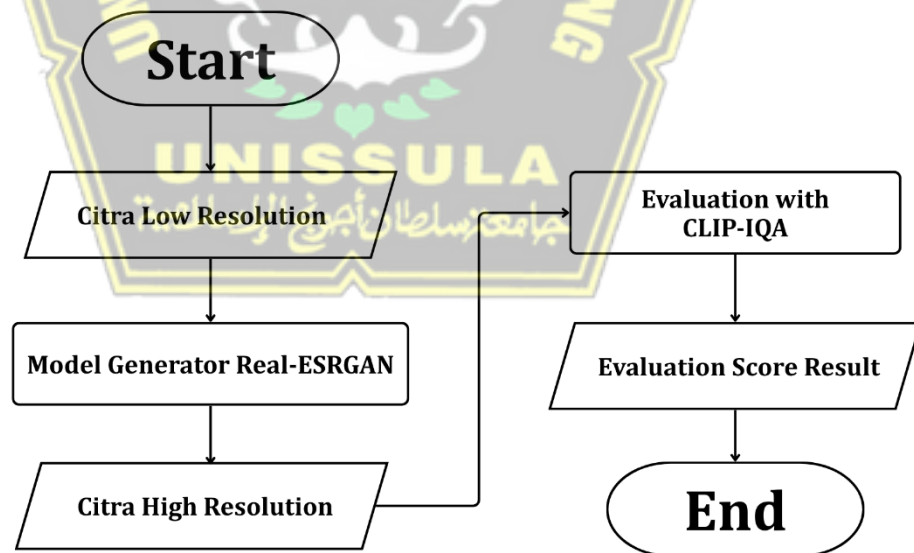
Real Enhanced Super-Resolution Generative Adversarial Network (Real-ESRGAN) adalah pengembangan dari metode ESRGAN yang ditujukan untuk mengatasi masalah *blind super-resolution*, yaitu peningkatan resolusi gambar dari citra resolusi rendah (LR) ke citra resolusi tinggi (HR) tanpa mengetahui bentuk degradasi yang terjadi. Gambar resolusi rendah dalam kasus nyata tidak hanya dihasilkan oleh proses *downsampling* biasa, namun juga kombinasi kompleks dari *blur*, *noise*, dan kompresi *JPEG*.

Untuk mengatasi degradasi kompleks ini, Real-ESRGAN mengandalkan jaringan *generator* yang identik dengan ESRGAN, yaitu sebuah jaringan konvolusional dalam dengan *block Residual-in-Residual Dense Blocks* (RRDB). Block ini menggabungkan kekuatan *residual learning* dan *dense connection* untuk mempertahankan detail fitur meskipun jaringan sangat dalam. Untuk efisiensi komputasi, diterapkan teknik *pixel-unshuffle*,

yaitu kebalikan dari *pixel-shuffle* yang berfungsi mengurangi dimensi spasial input sambil meningkatkan jumlah saluran (*channels*). Dengan demikian, sebagian besar komputasi terjadi pada resolusi yang lebih rendah.

Pada sisi *discriminator*, Real-ESRGAN menggantikan *VGG-style discriminator* dengan arsitektur *U-Net* yang dilengkapi dengan *skip connections*. Hal ini dilakukan agar jaringan dapat memberikan *gradient* pada level piksel terhadap *generator*, yang penting untuk mempertajam detail tekstur lokal. *Discriminator* ini bertugas membedakan citra hasil *generator* dengan citra nyata, dan memberikan sinyal *gradient* kepada *generator* untuk meningkatkan kualitas hasil.

Berdasarkan tujuan penelitian yang telah dijelaskan pada bab sebelumnya, penelitian ini bertujuan untuk mengimplementasikan dan menguji kinerja model Real ESRGAN dalam meningkatkan resolusi citra bangunan bersejarah. Tujuan tersebut dapat dicapai apabila penelitian dapat menghasilkan model yang optimal dalam meningkatkan resolusi citra. Sehingga, perlu perancangan yang baik untuk mencapai tujuan tersebut.



Gambar 3. 1 *Flowchart* rancangan sistem

Gambar 3.1 menunjukkan alur kerja sistem peningkatan kualitas citra menggunakan metode Real-ESRGAN dan evaluasi dengan CLIP-IQA. Proses diawali dengan citra beresolusi rendah (*low resolution image*) yang

menjadi input awal sistem. Citra ini kemudian diproses oleh model generator Real-ESRGAN yaitu suatu model *deep learning* yang dirancang khusus untuk meningkatkan resolusi dan kualitas visual gambar. Hasil dari proses ini adalah citra beresolusi tinggi (*high resolution image*). Setelah citra ditingkatkan, dilakukan proses evaluasi kualitas visual menggunakan metrik CLIP-IQA. CLIP-IQA memanfaatkan model *Contrastive Language-Image Pretraining* (CLIP) untuk menilai kesesuaian semantik antara gambar dan deskripsi tekstual tanpa memerlukan *ground truth* sebagai referensi. Tahapan ini memberikan skor penilaian yang menunjukkan seberapa baik kualitas citra hasil peningkatan dibandingkan dengan persepsi kualitas visual manusia.

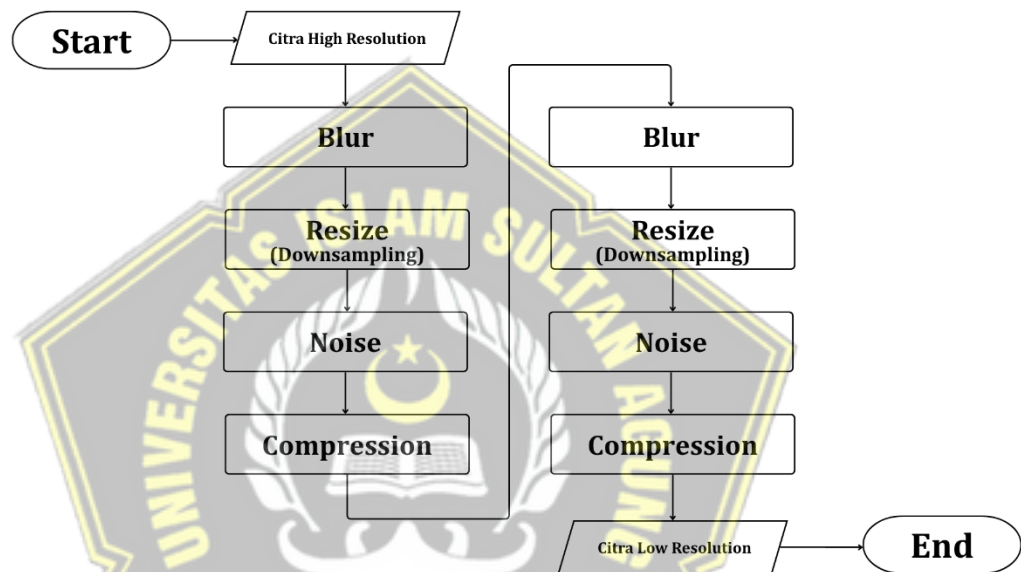
Skor hasil evaluasi tersebut kemudian dicatat dan ditampilkan sebagai *evaluation score result* yang menjadi output akhir dari sistem. Seluruh proses ini membantu dalam mengukur kinerja model Real-ESRGAN dalam meningkatkan kualitas citra secara objektif dan mendekati persepsi visual manusia.

3.2 Data Understanding

Tahapan selanjutnya dalam proses penelitian adalah *data understanding* atau pemahaman data. Pada penelitian ini, data pelatihan yang digunakan berasal dari dataset Flickr2k, yaitu kumpulan citra beresolusi tinggi yang diambil dari situs *Flickr.com*. Dataset ini memiliki berbagai variasi, seperti citra bangunan, hewan, tumbuhan, dan manusia, serta bersifat *open source* sehingga dapat dimanfaatkan secara bebas untuk keperluan penelitian. Citra dalam dataset ini memiliki format *Portable Network Graphics* (PNG). Jumlah data pelatihan (*training*) yang digunakan dibatasi sebanyak 2.000 citra untuk menyesuaikan dengan keterbatasan sumber daya komputasi yang tersedia. Sementara itu, untuk data pengujian (*testing*), digunakan 2 citra beresolusi rendah yang diambil dari koleksi Leiden University Libraries yang diberikan berbagai jenis degradasi sehingga menghasilkan total 12 citra pengujian. Strategi ini dilakukan untuk mengevaluasi performa model dalam menangani berbagai kondisi degradasi nyata dan menilai efektivitas peningkatan kualitas citra menggunakan metode Real-ESRGAN.

3.3 Data Preparation

Tahap selanjutnya *data preparation* atau persiapan data. Pada tahap ini, data dipersiapkan untuk dapat dilatih pada model menggunakan *High-order Degradation Model*. Model ini mensimulasikan proses degradasi citra resolusi rendah secara bertahap dan kompleks agar menyerupai kondisi nyata. Tahap ini dilakukan melalui beberapa proses berurutan, yaitu *blur*, *resize* (*downsampling*), *noise*, dan *JPEG compression*.



Gambar 3. 2 Flowchart high-oder degradation model

Gambar 3.2 menunjukkan skema alur *High-order Degradation Model* yang terdiri dari dua tahap utama, yaitu *first-order degradation* dan *second-order degradation*, yang masing-masing mencakup proses *blur*, *downsampling*, penambahan *noise*, dan kompresi JPEG dengan berbagai variasi parameter.

3.3.1 Blur

Blur merupakan teknik yang digunakan untuk mengurangi ketajaman suatu citra dengan cara memperhalus transisi warna antar piksel. Dalam penelitian ini, proses *blurring* dilakukan menggunakan *Generalized Gaussian filter* untuk mensimulasikan degradasi citra resolusi rendah secara lebih realistis. Filter ini memungkinkan pengaturan bentuk dan arah sebaran *blur*, baik secara *isotropic* (merata ke segala arah) maupun *anisotropic* (menyebar

dominan ke satu arah). *Isotropic* digunakan untuk meniru efek *blur* dari lensa kamera biasa, sementara *anisotropic* mensimulasikan efek *blur* akibat gerakan atau getaran saat pengambilan gambar. Setelah proses *Gaussian*, citra juga diproses dengan *2D sinc filter*, yang berfungsi sebagai *low-pass filter* untuk menyaring frekuensi tinggi dan memperhalus citra secara lebih presisi. Penambahan degradasi ini bertujuan untuk merepresentasikan kondisi nyata dari citra resolusi rendah akibat keterbatasan perangkat atau lingkungan.

3.3.2 *Resize (downsampling)*

Resize (downsampling) adalah proses untuk menurunkan resolusi citra dengan cara memperkecil ukuran dimensinya. Tahapan ini dilakukan dalam penelitian guna menghasilkan citra beresolusi rendah yang digunakan sebagai data uji atau pelatihan. Pada proses ini, Untuk melakukan *resize (downsampling)*, digunakan tiga metode *interpolasi* yang umum dalam pengolahan citra digital yaitu *bicubic*, *bilinear*, dan *area*. Metode *bicubic* menggunakan perhitungan piksel yang lebih kompleks sehingga menghasilkan citra yang lebih halus dan detail. Sementara itu, metode *bilinear* melakukan interpolasi berdasarkan empat piksel terdekat sehingga menghasilkan citra yang cukup halus namun dengan perhitungan lebih ringan. Adapun metode *area* bekerja dengan menghitung rata-rata nilai piksel dalam suatu wilayah dan lebih cocok digunakan untuk memperkecil gambar karena mampu mempertahankan informasi keseluruhan citra.

3.3.3 *Noise*

Noise merupakan gangguan acak pada citra digital yang memengaruhi kualitas visualnya. Dalam konteks pemrosesan citra, *noise* dapat berasal dari berbagai sumber seperti sensor kamera, kondisi pencahayaan yang buruk dan proses transmisi data. Beberapa jenis *noise* yang umum digunakan dalam simulasi gangguan pada citra antara lain *gaussian noise*, *poisson noise*, *color noise*, dan *gray noise*. *Gaussian noise* ditandai dengan distribusi nilai *pixel* acak yang mengikuti kurva distribusi normal dan sering digunakan untuk mensimulasikan gangguan elektronik. *Poisson noise* biasanya muncul akibat *fluktuasi kuantum* dalam pengambilan gambar, yang erat kaitannya dengan

jumlah foton yang ditangkap oleh sensor. *Color noise* memengaruhi saluran warna dalam citra RGB secara acak, menghasilkan bintik-bintik berwarna yang mengganggu keutuhan warna. Sementara itu, *gray noise* lebih merujuk pada gangguan acak pada tingkat intensitas keabuan dalam citra *grayscale*. Penerapan berbagai jenis noise ini bertujuan untuk mensimulasikan kondisi nyata di mana kamera mengalami keterbatasan dalam menangkap gambar secara optimal.

3.3.4 JPEG Compression

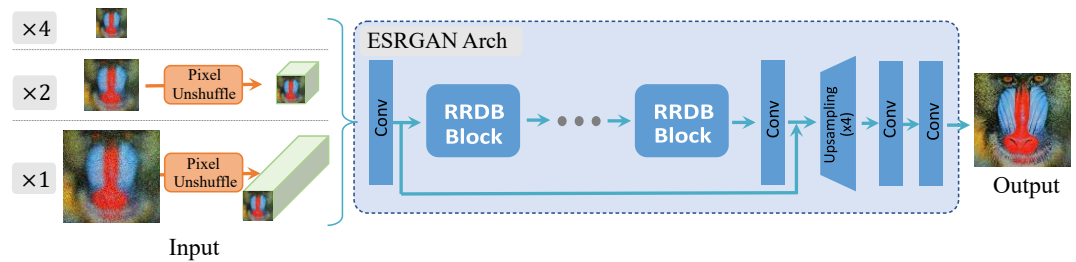
Kompresi JPEG merupakan metode pengurangan ukuran *file* gambar dengan menghilangkan beberapa detail yang dianggap kurang penting, namun hal ini dapat menyebabkan munculnya artefak kompresi seperti blok-blok yang terlihat dan distorsi pada warna maupun kecerahan. Gangguan ini biasanya terjadi karena keterbatasan algoritma dalam merekonstruksi gambar asli setelah kompresi, terutama pada tingkat kompresi yang tinggi. Penerapan kompresi *JPEG* pada citra dapat memberikan efek penurunan kualitas yang menyerupai *noise*, membuat gambar tampak kurang tajam dan detailnya hilang.

3.4 Modeling

Tahap berikutnya adalah tahapan pemodelan, di mana data *training* yang telah melalui proses *preprocessing* dimasukkan ke dalam metode Real-ESRGAN. Pada tahap ini, data citra diproses oleh Real-ESRGAN untuk menghasilkan citra dengan resolusi tinggi yang memiliki kualitas visual lebih nyata dan detail yang tajam.

Real Enhanced Super-Resolution Generative Adversarial Network (Real-ESRGAN) merupakan metode peningkatan resolusi citra yang dikembangkan berdasarkan arsitektur *Generative Adversarial Networks* (GAN), dengan tujuan utama memulihkan detail pada citra hasil pengambilan gambar dunia nyata. Metode ini terdiri dari dua komponen utama, yaitu model *generator* dan model *discriminator*, yang berperan secara kompetitif untuk menghasilkan gambar super resolusi dengan kualitas yang optimal. Kedua model tersebut dijelaskan secara rinci sebagai berikut:

1. Generator



Gambar 3. 3 Arsitektur dengan generator Real-ESRGAN (Wang dkk., 2021)

Generator pada Real-ESRGAN berfungsi untuk meningkatkan kualitas dan ukuran citra dari resolusi rendah ke resolusi tinggi. Proses dimulai dengan tahap *Pixel Unshuffle*, yaitu metode yang mengubah citra resolusi tinggi menjadi format resolusi rendah dengan jumlah saluran (*channel*) yang lebih banyak. Tujuannya adalah untuk menyederhanakan struktur citra sehingga lebih mudah diproses oleh jaringan.

Setelah itu, data masuk ke dalam lapisan *Convolutional* pertama untuk mengambil fitur awal dari citra. Fitur ini kemudian diproses oleh beberapa blok *Residual-in-Residual Dense Block* (RRDB). RRDB adalah blok utama yang terdiri dari beberapa lapisan yang saling terhubung dan dilengkapi jalur *shortcut*. Struktur ini membantu jaringan agar bisa belajar lebih baik dan stabil meskipun jumlah lapisannya banyak.

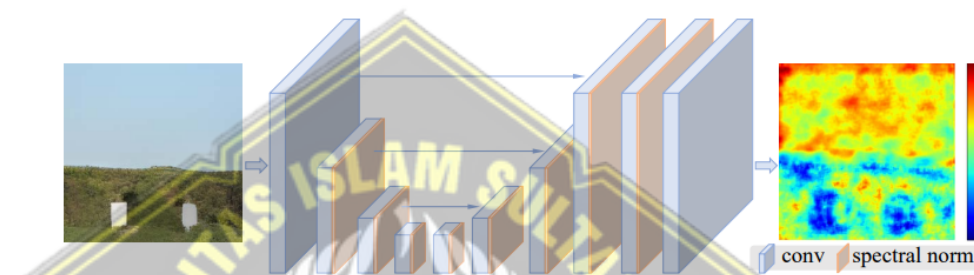
Setelah semua *block* RRDB selesai memproses data, hasilnya dilanjutkan ke beberapa lapisan *Convolutional* lainnya. Lalu, dilakukan proses *upsampling* atau peningkatan ukuran gambar menggunakan metode *Nearest-Neighbor Interpolation* yang disusul oleh lapisan *Convolutional* dan fungsi aktivasi. Proses ini dilakukan bertahap sampai citra mencapai resolusi akhir.

Pada bagian akhir, ada satu lapisan *Convolutional* lagi untuk menghasilkan citra akhir dengan 3 saluran (RGB). Tidak digunakan fungsi aktivasi di lapisan ini agar hasil citra tidak mengalami perubahan nilai yang tidak perlu.

Hasil akhir dari *generator* ini memiliki ukuran 4 kali lebih besar dari ukuran citra *input*. Contoh ukuran *input* dan *output* adalah sebagai berikut:

$$\begin{aligned}
 \text{input} &= (64, 64) \\
 \text{output} &= (x * 4, x * 4) \\
 \text{output} &= (64 * 4, 64 * 4) \\
 \text{output} &= (256, 256)
 \end{aligned}
 \tag{2}$$

2. Discriminator



Gambar 3. 4 Arsitektur dengan *discriminator* Real-ESRGAN (Wang dkk., 2021)

Discriminator pada Real-ESRGAN menggunakan arsitektur *U-Net*, yang umum digunakan untuk segmentasi citra karena kemampuannya dalam mempertahankan informasi lokal dan global secara bersamaan. Pada arsitektur ini, struktur jaringan dibagi menjadi dua bagian utama, yaitu *encoder* dan *decoder*.

Bagian *encoder* bertugas mengekstraksi fitur dari citra *input* melalui beberapa lapisan *Convolutional* yang disertai fungsi aktivasi. Setiap tahap *encoder* akan memperkecil ukuran fitur (*downsampling*), tetapi meningkatkan jumlah *channel* untuk menangkap informasi kontekstual. Selanjutnya, bagian *decoder* bertugas untuk mengembalikan ukuran fitur ke ukuran semula melalui proses *upsampling*. Di setiap tahap *decoder*, hasil dari *encoder* sebelumnya disambungkan *skip connection* ke layer yang sesuai di *decoder*. Ini membantu jaringan mempertahankan detail spasial dari citra asli.

Untuk meningkatkan stabilitas pelatihan dan menghindari *overfitting*, digunakan teknik *Spectral Normalization* (SN) pada setiap

lapisan *Convolutional*. Teknik ini bertujuan untuk mengontrol nilai bobot pada jaringan agar tetap stabil, sehingga hasil diskriminasi menjadi lebih akurat dan pelatihan lebih konsisten.

Discriminator ini tidak hanya memutuskan apakah gambar hasil *generator* realistis atau tidak, tetapi juga memberikan peta keyakinan (*confidence map*) yang menunjukkan bagian mana dari gambar yang dianggap tidak realistis. Hal ini membantu *generator* untuk fokus memperbaiki detail lokal secara lebih efektif.

3.5 Evaluation

Tahapan selanjutnya ialah tahap pengujian atau evaluasi. Tahap ini dilakukan untuk mengukur kinerja model Real-ESRGAN dalam menghasilkan citra resolusi tinggi yang sesuai dengan persepsi visual manusia. Pengujian dilakukan menggunakan metode *Contrastive Language-Image Pretraining for Image Quality Assessment* (CLIP-IQA). CLIP-IQA merupakan pendekatan *no-reference image quality assessment* (NR-IQA) atau tidak memerlukan citra referensi (*ground truth*) untuk melakukan evaluasi kualitas. Hal ini menjadikannya fleksibel dalam situasi di mana citra asli tidak tersedia atau sulit diakses.

CLIP-IQA memanfaatkan model CLIP yang telah dilatih untuk memahami hubungan semantik antara teks dan gambar dan menilai kualitas citra berdasarkan tingkat kesesuaian dengan deskripsi atau prompt tertentu seperti "*Good photo*". Evaluasi dilakukan dengan menghitung *cosine similarity* antara representasi fitur dari citra dengan representasi fitur dari prompt, sehingga menghasilkan skor kualitas yang berada dalam rentang [0, 1]. Semakin tinggi nilai tersebut, maka semakin baik kualitas perseptual dari citra. Pendekatan ini memberikan keunggulan karena mempertimbangkan aspek estetika dan semantik serta lebih sesuai dengan persepsi visual manusia.

Skenario pengujian yang digunakan dalam penelitian ini hanya menggunakan satu jenis skenario, yaitu skenario pengujian *epoch*. Skenario ini dilakukan dengan cara menguji model berdasarkan jumlah *epoch* yang

digunakan selama tahap pelatihan dengan 2 (dua) citra yang ditambahkan berbagai degradasi menjadi 12 citra untuk *testing*. Tujuan dari skenario ini adalah untuk mengetahui jumlah *epoch* yang paling optimal dalam melatih model sehingga dapat menghasilkan performa yang terbaik. Dengan melakukan variasi jumlah *epoch*, diharapkan dapat diperoleh informasi yang tepat mengenai pengaruh jumlah *epoch* terhadap kinerja model.

3.6 Deployment

Selanjutnya tahap *deployment*, untuk mengimplementasikan model Real-ESRGAN ke dalam sebuah *user interface*. Tahap ini merupakan tahap akhir yang dilakukan pada metodologi CRISP-DM. *Deployment* dilakukan setelah tahapan evaluasi agar mengetahui model yang memiliki kinerja terbaik. Setelah didapatkan nilai pengujian PSNR dan SSIM, maka model yang optimal kemudian disimpan lalu dimasukkan ke dalam aplikasi.

Pada penelitian ini, media yang dipakai berbasis *website* yang dibuat menggunakan *framework Flask*. Flask adalah *framework web* sederhana dan ringan yang dibuat dengan bahasa pemrograman Python (Kho & Braganca, 2023). Penggunaan flask dikarenakan mudah integrasi ke dalam kote python dan mudah dihubungkan dengan HTML serta *javascript*. *User interface* ini memiliki fitur untuk memperbesar resolusi citra. Cara kerjanya dengan memasukkan gambar atau citra resolusi rendah kemudian sistem akan mengubah resolusi citra 4x dari ukuran semula menggunakan model Real-ESRGAN. Pengguna dapat mengunduh citra hasil dari pembesaran resolusi pada aplikasi tersebut. Adapun rancangan tampilan aplikasi dapat dilihat pada Gambar 3.5 dan Gambar 3.6.



Gambar 3. 5 *User interface* halaman *input*



Gambar 3. 6 *User interface* halaman *output*

Gambar 3.5 di atas merupakan rancangan tampilan aplikasi untuk *input* gambar atau citra. *Input* citra dapat berupa gambar dengan resolusi rendah. Lalu, pada Gambar 3.6 ialah rancangan tampilan hasil *output* gambar setelah dilakukan proses pembesaran resolusi menggunakan model Real-ESRGAN. Pada halaman ini, pengguna dapat mengunduh gambar output dengan menekan tombol *download*.

BAB IV

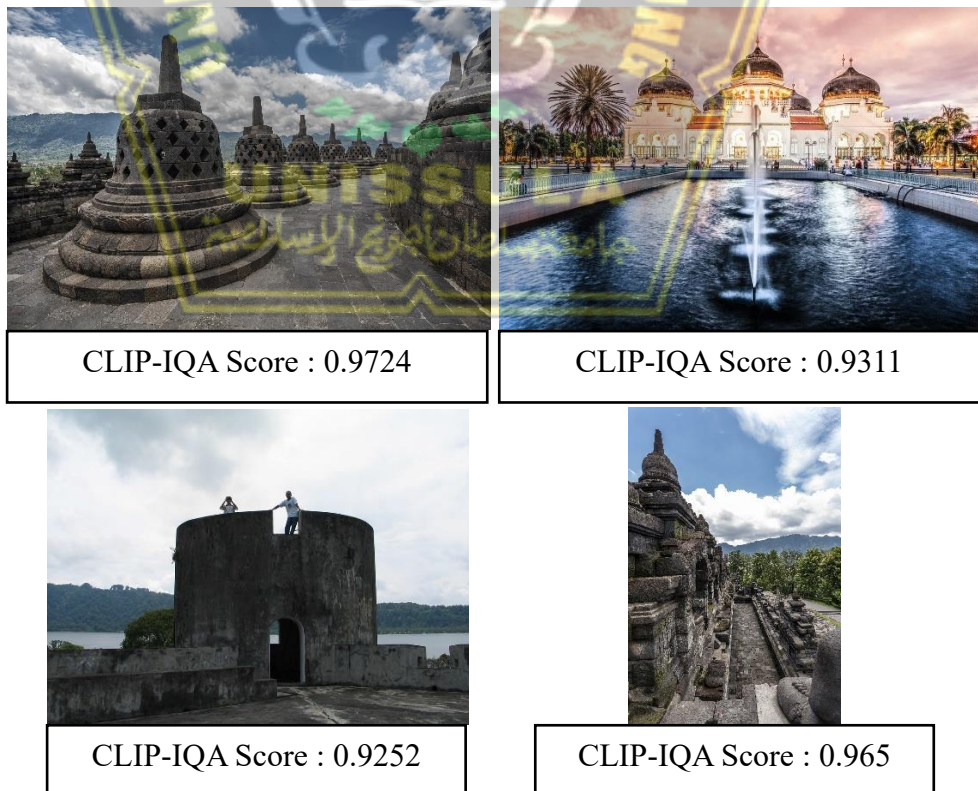
HASIL DAN ANALISIS PENELITIAN

4.1 Hasil

4.1.1 *Bussiness Understanding*

Business Understanding atau pemahaman bisnis yang didapatkan pada penelitian ini dilatar belakangi permasalahan citra bangunan bersejarah yang beresolusi rendah dengan degradasi yang kompleks dan tak terduga seperti kerusakan fisik dan penyimpanan kurang memadai. Untuk mengatasi permasalahan tersebut, resolusi citra ditingkatkan dengan metode Real-ESRGAN. Perencanaan yang baik diperlukan untuk mencapai tujuan dalam menerapkan Real-ESRGAN dan mengetahui kinerja model tersebut sesuai tujuan penelitian. Hasil dari pemahaman bisnis adalah penelitian ini dapat berjalan dengan baik. Setiap tahapan dapat dilakukan sesuai rencana dan menghasilkan hasil yang diharapkan.

4.1.2 *Data Understanding*



Gambar 4. 1 Contoh citra *training* dari Flickr2k

Pada gambar 4.1 merupakan contoh data *training* berasal dari dataset Flickr2k. Data tersebut adalah kumpulan foto beresolusi tinggi dari situs Flickr.com, dataset ini memiliki berbagai variasi seperti bangunan, hewan, tumbuhan dan manusia. Data yang dipakai berjumlah 2000 citra dengan resolusi yang bervariasi dan memiliki format *file Portable Network Graphics* (PNG). Berdasarkan hasil pengukuran menggunakan CLIP-IQA, citra dari flickr.com memperoleh skor dengan rata-rata diatas 0,9. Nilai ini menunjukkan bahwa kualitas citra mendekati nilai maksimal, yaitu 1 (satu).



Gambar 4. 2 Citra Candi Borobudur (Leiden University Libraries abad 19)

Gambar 4.2 merupakan citra bangunan bersejarah yang diperoleh dari koleksi Leiden University Libraries. Citra ini akan digunakan sebagai salah satu data uji (*testing*) dalam penelitian, dimana nanti citra tersebut akan ditambahkan berbagai degradasi seperti *blur*, *noise*, *resize* dan kompresi. Format file citra tersebut adalah Portable Network Graphics (PNG) dengan model warna (*Red, Green, Blue*) RGB dan saat ini masih dalam kondisi asli dari sumbernya tanpa penambahan degradasi. Namun citra ini telah

mengalami degradasi visual akibat proses digitalisasi arsip dan usia dokumen meskipun jenis degradasi tersebut tidak dapat diidentifikasi secara pasti. Berdasarkan hasil pengukuran menggunakan CLIP-IQA, citra Candi Borobudur memperoleh skor sebesar 0,8264. Nilai ini menunjukkan bahwa kualitas citra masih bisa ditingkatkan lebih lanjut agar mendekati nilai maksimal, yaitu 1 (satu).



"Stupa Borobudur, Jogjakarta. Stupa of Borobudur, Jogjakarta, Central Java"

Shelfmark

KITLV 1407025

Published

[Between the 19th and 20th century]

Gambar 4. 3 Keterangan citra Candi Borobudur (Leiden University Libraries abad 19)

Gambar 4.3 merupakan keterangan dari citra Candi Borobudur yang diambil dari koleksi digital milik Leiden University Libraries. Citra tersebut menampilkan stupa Candi Borobudur yang berlokasi di Yogyakarta, Jawa Tengah dan diberi judul "*Stupa Borobudur, Jogjakarta. Stupa of Borobudur, Jogjakarta, Central Java*". Berdasarkan informasi dari pustaka, citra ini dari lembaga akademik belanda yang fokus studi pada Asia Tenggara dengan nama *Koninklijk Instituut voor Taal-, Land- en Volkenkunde* (KITLV) dan memiliki kode *shelfmark* atau penanda unik dalam pengarsipan citra yaitu KITLV 1407025 dan dipublikasikan antara abad ke-19 hingga abad ke-20.



Gambar 4. 4 Citra Menara Kudus (Leiden University Libraries abad 19)

Gambar 4.4 merupakan citra bangunan bersejarah yang diperoleh dari koleksi Leiden University Libraries. Citra ini akan digunakan sebagai salah satu data uji (*testing*) dalam penelitian, dimana nanti citra tersebut akan ditambahkan berbagai degradasi seperti *blur*, *noise*, *resize* dan kompresi. Format file citra tersebut adalah Portable Network Graphics (PNG) dengan model warna hitam putih dan saat ini masih dalam kondisi asli dari sumbernya tanpa penambahan degradasi. Namun citra ini telah mengalami degradasi visual akibat proses digitalisasi arsip dan usia dokumen meskipun jenis degradasi tersebut tidak dapat diidentifikasi secara pasti. Berdasarkan hasil pengukuran menggunakan CLIP-IQA, citra menara kudus memperoleh skor sebesar 0,4007. Nilai ini menunjukkan bahwa kualitas citra masih bisa ditingkatkan lebih lanjut agar mendekati nilai maksimal yaitu 1 (satu).



Islamitische begraafplaats Soenan Koedoes te Koedoes

Shelfmark

KITLV 166424

Published

[Between 1941 and 1953]

Gambar 4. 5 Keterangan citra Menara Kudus (Leiden University Libraries abad 19)

Gambar 4.5 merupakan keterangan dari citra Menara Kudus yang diambil dari koleksi digital milik Leiden University Libraries. Citra tersebut menampilkan Menara Kudus yang berlokasi di Kudus, Jawa Tengah dan diberi judul "*Islamitische begraafplaats Soenan Koedoes te Koedoes*". Berdasarkan informasi dari pustaka, citra ini dari lembaga akademik belanda yang fokus studi pada Asia Tenggara dengan nama *Koninklijk Instituut voor Taal-, Land- en Volkenkunde* (KITLV) dan memiliki kode *shelfmark* atau penanda unik dalam pengarsipan citra yaitu KITLV 166424 dan dipublikasikan antara abad ke-1941 hingga abad ke-1953.

4.1.3 Data Preparation

Citra *training* yang didapatkan pada tahap *data understanding* selanjutnya dilakukan *preprocessing* pada tahap *data preparation*. Tahap ini dilakukan untuk mensimulasikan data *trainng* seperti citra yang mengalami degradasi yang kompleks dan tak terduga seperti kerusakan fisik sehingga membuat data citra optimal ketika diterapkan pada tahap *modeling* atau pemodelan. Tahap ini menggunakan *high-order degradation model* dengan mensimulasikan proses degradasi citra resolusi rendah secara bertahap dan kompleks agar menyerupai kondisi nyata. Tahap ini dilakukan melalui beberapa proses berurutan, yaitu *blur*, *resize (downsampling)*, *noise*, dan *JPEG compression*.

Adapun hasil dari proses *high-order degradation model* menunjukkan bahwa citra yang telah mengalami simulasi degradasi kompleks dengan berbagai jenis kerusakan nyata, seperti *blur*, *noise*, dan kompresi, sehingga memberikan kerusakan yang lebih realistis dalam proses pelatihan model.



Gambar Asli
(180 kb) 1024 x 768 px

Gambar *Final Degraded*
(5.65 kb) 256 x 192 px

Gambar 4. 6 Perbedaan antara gambar asli dengan setelah proses *high-order degradation model*

Sebagai ilustrasi hasil proses tersebut, Gambar 4.6 sebelah kiri merupakan citra asli bangunan bersejarah yang digunakan dalam penelitian, sedangkan Gambar 4.6 sebelah kanan menunjukkan hasil akhir citra setelah mengalami proses degradasi sesuai skema.

4.1.3.1 Blur

Proses *blur* dilakukan menggunakan *Generalized Gaussian filter* dan *2D sinc filter* untuk mensimulasikan degradasi citra resolusi rendah secara lebih realistis. *Gaussian filter* digunakan dalam bentuk *isotropic* (merata ke segala arah) dan *anisotropic* (menyebarkan dominan ke satu arah) untuk meniru efek *blur* dari lensa maupun gerakan kamera, sedangkan *2D sinc filter* berfungsi sebagai *low-pass filter* untuk mengurangi frekuensi tinggi dan memperhalus citra. Gambar 4.7 menunjukkan perbandingan citra sebelum dan sesudah diterapkan proses *blur* dengan pendekatan ini.



Gambar Asli
1024 x 768 px

Gausan Filter : Isotropic
512 x 384 px



Gausan Filter : Anisotropic

512 x 384 px



2D Sinc Filter

512 x 384 px

Gambar 4. 7 Citra dengan degradasi *blur*

4.1.3.2 Resize (downsampling)

Proses *resize (downsampling)* dengan sebuah contoh citra dilakukan untuk menghasilkan citra resolusi rendah dari citra resolusi tinggi berukuran 1024×768 piksel menjadi 307×230 piksel. Tiga metode interpolasi digunakan, yaitu *bicubic*, *bilinear*, dan *area*. *Bicubic* menghasilkan citra lebih halus dan detail, *bilinear* menawarkan perhitungan ringan dengan hasil cukup halus, sedangkan metode *area* menjaga informasi keseluruhan citra saat memperkecil ukuran. Gambar 4.8 menunjukkan perbandingan citra sebelum dan sesudah diterapkan proses *resize (downsampling)* dengan pendekatan ini.



Gambar Asli

1024 x 768 px



Resize Bicubic

307 x 230 px



Resize Biliner

307 x 230 px



Resize Area

307 x 230 px

Gambar 4. 8 Citra dengan degradasi *resize* (*downsampling*)

4.1.3.3 Noise

Proses penambahan *noise* dilakukan untuk mensimulasikan gangguan visual yang umum terjadi pada citra digital akibat keterbatasan sensor atau kondisi lingkungan. Beberapa jenis *noise* yang digunakan antara lain *Gaussian noise* (bintik halus), *Poisson noise* (bintik padah), *Color noise* (bintik halus berwarna), dan *Gray noise* (bintik halus abu-abu). Masing-masing jenis *noise* merepresentasikan gangguan spesifik, seperti gangguan elektronik, fluktuasi cahaya, serta gangguan warna dan intensitas keabuan sehingga menciptakan kondisi citra resolusi rendah yang lebih realistis. Gambar 4.9 menunjukkan perbandingan citra sebelum dan sesudah diterapkan proses *noise* dengan pendekatan ini.



Gambar Asli

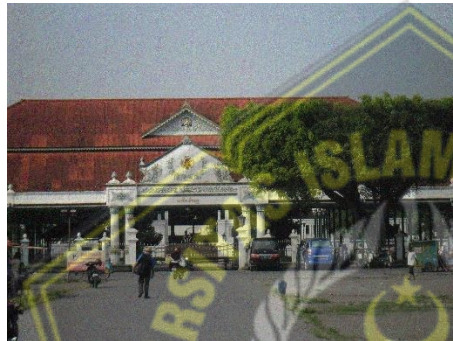
1024 x 768 px



Gaussian Noise
1024 x 768 px



Poisson Noise
1024 x 768 px



Color Noise
1024 x 768 px



Gray Noise
1024 x 768 px

Gambar 4. 9 Citra dengan degradasi *noise*

4.1.3.4 JPEG Compression

Penerapan kompresi JPEG dilakukan untuk mensimulasikan penurunan kualitas citra akibat pengurangan ukuran *file*. Proses ini dapat menimbulkan artefak seperti blok-blok kasar dan distorsi warna atau kecerahan, terutama pada tingkat kompresi tinggi, sehingga citra tampak kurang tajam dan kehilangan detail. Gambar 4.10 menunjukkan perbandingan citra sebelum dan sesudah diterapkan proses *JPEG Compression* dengan pendekatan ini.



Gambar Asli

(180 kb) 1024 x 768 px



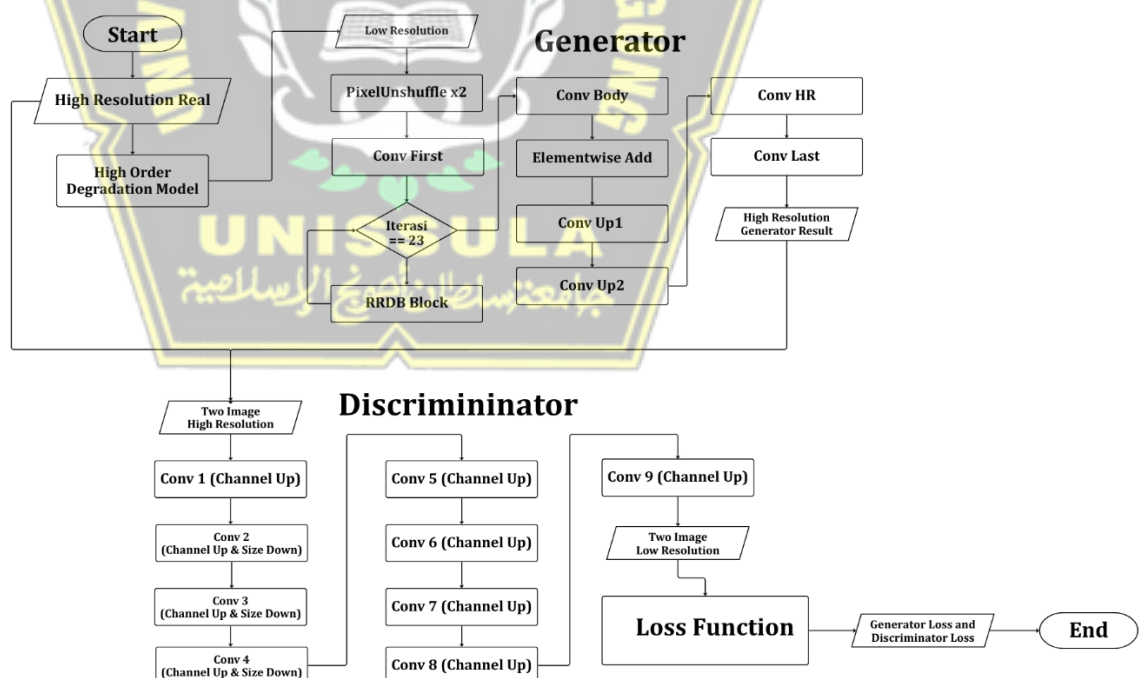
JPEG Compression

(27 kb) 1024 x 768 px

Gambar 4. 10 Degradasi dengan JPEG *compression*

4.1.4 Modeling

Modeling atau pemodelan merupakan tahap untuk membuat model yang akan menyelesaikan permasalahan penelitian. Pada tahap ini, dilakukan penerapan metode Real-ESRGAN sesuai arsitektur model yang dirancang pada bab sebelumnya. Adapun hasil pemodelan menggunakan metode Real-ESRGAN ditampilkan pada Gambar berikut:

Gambar 4. 11 Alur proses *modeling* pada metode Real-ESRGAN

Gambar di atas menunjukkan alur arsitektur dari model Real-ESRGAN yang terdiri dari 2 (dua) komponen utama, yaitu *generator* dan *discriminator*.

Proses diawali dengan input berupa citra resolusi tinggi yang kemudian diberikan ke *High Order Degradation Model* untuk mensimulasikan berbagai bentuk degradasi kompleks seperti *blur*, *noise*, *resize* dan kompresi. Hasil degradasi ini menjadi citra resolusi rendah yang kemudian masuk ke dalam *generator*. Di dalam *generator*, proses dimulai dari *PixelUnshuffle* dan *Conv First*, lalu dilanjutkan ke 23 blok RRDB (*Residual-in-Residual Dense Blocks*) yang bertugas mengekstraksi dan memperdalam citra. Setelahnya, citra yang telah diperoleh melalui jalur konvolusi diproses melalui lapisan upsampling (*Conv Up1*, *Conv Up2*) dan dilanjutkan hingga menghasilkan citra resolusi tinggi baru sebagai output dari *generator*.

Selanjutnya, citra hasil dari *generator* dan citra resolusi tinggi asli digunakan sebagai *input* bagi *discriminator*. *Discriminator* bertugas membedakan antara citra hasil rekonstruksi dengan citra asli melalui serangkaian lapisan konvolusi (*Conv 1–Conv 9*) yang bertahap menaikkan jumlah channel dan menurunkan ukuran spasial. Informasi dari kedua citra (real dan hasil) dikirimkan ke fungsi *loss* yang akan menghitung *generator loss* dan *discriminator loss*. Output dari fungsi *loss* ini menjadi dasar dalam proses pembelajaran model yang kemudian digunakan untuk memperbarui parameter jaringan. Dengan pendekatan ini, Real-ESRGAN mampu meningkatkan kualitas citra resolusi rendah menjadi lebih tajam, realistis dan kondisi degradasi nyata.

Penjelasan lebih lanjut mengenai komponen penyusun model Real-ESRGAN dapat dilihat pada uraian dibawah. Adapun berikut hasil pemodelan menggunakan metode Real-ESRGAN yang terdiri dari *generator* dan *discriminator*.

1. Model Generator

Model *generator* merupakan model yang dilatih untuk menghasilkan citra super resolusi (citra hasil yang telah diperbesar). Hasil citra yang dibuat oleh *generator* dapat meningkatkan resolusi citra sebesar 4x (empat kali) lipat dari ukuran sebelumnya. Adapun rincian struktur arsitektur *generator* ini ditunjukkan pada tabel 4.1.

Tabel 4. 1 Rincian arsitektur *generator*

No	Type Layer	Input Shape	Output Shape
1	Input	(B, 3, 64, 64)	(B, 3, 64, 64)
2	PixelUnshuffle x2	(B, 3, 64, 64)	(B, 12, 32, 32)
3	Conv First	(B, 12, 32, 32)	(B, 64, 32, 32)
4	RRDB $\times 23$	(B, 64, 32, 32)	(B, 64, 32, 32)
5	Conv Body	(B, 64, 32, 32)	(B, 64, 32, 32)
6	Elementwise Add	(B, 64, 32, 32)	(B, 64, 32, 32)
7	Conv Up1	(B, 64, 32, 32)	(B, 64, 64, 64)
8	Conv Up2	(B, 64, 64, 64)	(B, 64, 128, 128)
10	Conv HR	(B, 64, 128, 128)	(B, 64, 256, 256)
11	Conv Last	(B, 64, 256, 256)	(B, 3, 256, 256)
12	Output	(B, 3, 256, 256)	(B, 3, 256, 256)

2. Model Discriminator

Discriminator pada Real-ESRGAN merupakan model yang membandingkan apakah citra hasil dari *generator* cukup realistis dan mirip dengan citra asli. Model ini berfungsi sebagai pembanding dan lawan dari model *generator* agar dapat menghasilkan citra yang diharapkan. Adapun rincian struktur arsitektur *discriminator* ini ditunjukkan pada tabel 4.2

Tabel 4. 2 Rincian arsitektur *discriminator*

No	Type Layer	Input Shape	Output Shape
1	Input	(B, 3, 256, 256)	(B, 3, 256, 256)
2	Conv0	(B, 3, 256, 256)	(B, 64, 256, 256)
3	Conv1	(B, 64, 256, 256)	(B, 128, 128, 128)
4	Conv2	(B, 128, 128, 128)	(B, 256, 64, 64)
5	Conv3	(B, 256, 64, 64)	(B, 512, 32, 32)
6	Conv4	(B, 512, 32, 32)	(B, 256, 32, 32)
7	Conv5	(B, 256, 32, 32)	(B, 128, 32, 32)
8	Conv6	(B, 128, 32, 32)	(B, 64, 32, 32)

Tabel 4. 3 Rincian arsitektur *discriminator* (lanjutan)

9	Conv7	(B, 64, 32, 32)	(B, 64, 32, 32)
10	Conv8	(B, 64, 32, 32)	(B, 64, 32, 32)
12	Conv9	(B, 64, 32, 32)	(B, 1, 32, 32)
13	Output	(B, 1, 32, 32)	(B, 1, 32, 32)

4.1.5 Evaluation

Tahap evaluasi dilakukan untuk menguji kinerja model Real-ESRGAN apakah model yang dihasilkan dapat mengatasi masalah pada pemahaman bisnis sebelumnya. Evaluasi juga dilakukan untuk menemukan kinerja model yang optimal untuk dapat diterapkan pada tahap *deployment*. Dengan begitu, aplikasi dapat mengatasi masalah resolusi rendah pada citra dengan baik.

Pengujian model dilakukan menggunakan citra testing dari Leiden University Libraries. Akan tetapi, citra diberikan berbagai degradasi seperti *blur*, *noise*, *resize*, kompresi dan *high order degradation model* agar mensimulasikan degradasi yang nyata. Penelitian ini menggunakan metode pengujian *Contrastive Language-Image Pretraining for Image Quality Assessment* (CLIP-IQA) untuk mengevaluasi kinerja model dalam menghasilkan citra resolusi tinggi. CLIP-IQA merupakan metode *no-reference* yang tidak memerlukan citra asli. Evaluasi dilakukan dengan membandingkan representasi fitur dari citra *input* dan citra *output*. Model yang memiliki kinerja baik akan menghasilkan citra resolusi tinggi yang memperoleh skor mendekati nilai 1 (satu) sehingga menandakan kualitas visual dan estetika yang baik secara persepsi manusia.

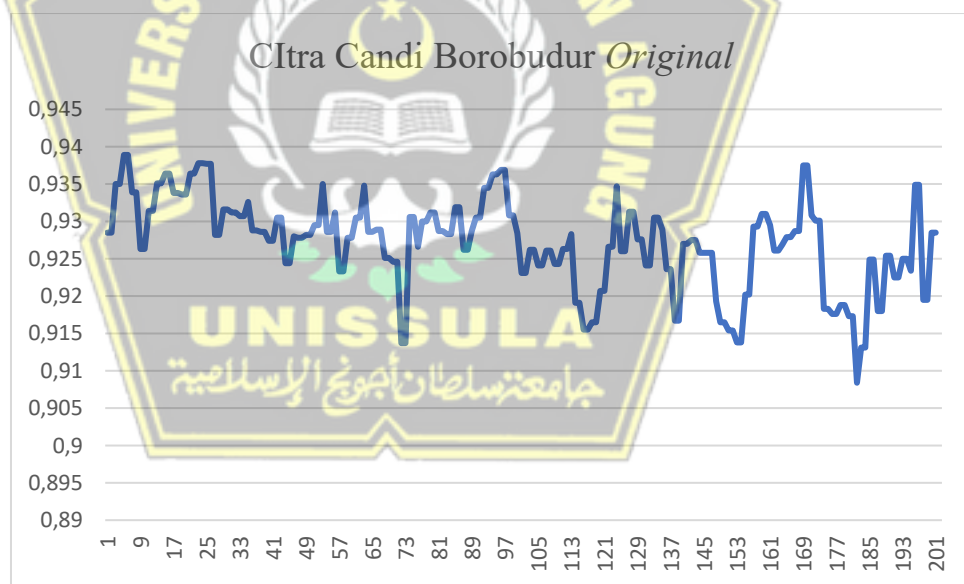
Pengujian *epoch* merupakan pengujian untuk mendapatkan model yang memiliki kinerja baik berdasarkan jumlah *epoch* yang dipakai. Adapun skenario pengujian dilakukan pada jumlah *epoch* yang berbeda. Berikut hasil skenario pengujian berdasarkan *epoch* yang berbeda dengan citra *testing* :

1. Citra Candi Borobudur Original

Tabel 4. 4 Hasil evaluasi citra Candi Borobudur *original*

Epoch	CLIP-IQA (Input)	CLIP-IQA (Output)
4	0,8264	0,9389
50	0,8264	0,9295
100	0,8264	0,9231
150	0,8264	0,9154
200	0,8264	0,9285

Tabel 4.4 menunjukkan hasil pengujian kualitas citra menggunakan metode CLIP-IQA pada citra asli Candi Borobudur *original* dengan berbagai model hasil pelatihan pada epoch berbeda. Skor tertinggi diperoleh pada epoch ke-4 sebesar 0,9389 sehingga menunjukkan memberikan peningkatan kualitas citra. Hasil tersebut dapat dilihat dari grafik pengujian berikut:



Gambar 4. 12 Grafik hasil evaluasi citra Candi Borobudur original

Gambar 4.12 menunjukkan grafik skor evaluasi CLIP-IQA *output* terhadap citra asli Candi Borobudur yang telah ditingkatkan resolusinya menggunakan model Real-ESRGAN pada berbagai epoch pelatihan. Terlihat bahwa nilai skor CLIP-IQA berfluktuasi seiring bertambahnya epoch. Secara umum, skor CLIP-IQA cenderung stabil menunjukkan bahwa model dapat mempertahankan kualitas perseptual citra hasil selama

pelatihan. Namun, perlu pemilihan epoch terbaik supaya mempertimbangkan kestabilan dan puncak skor CLIP-IQA bukan hanya berdasarkan banyaknya iterasi pelatihan. Hasil dari *input* dan *output* dapat dilihat pada gambar berikut :



CLIP-IQA Score : **0,8264**

Gambar 4. 13 Citra Candi Borobudur *original (input)*



CLIP-IQA Score : **0,9389**

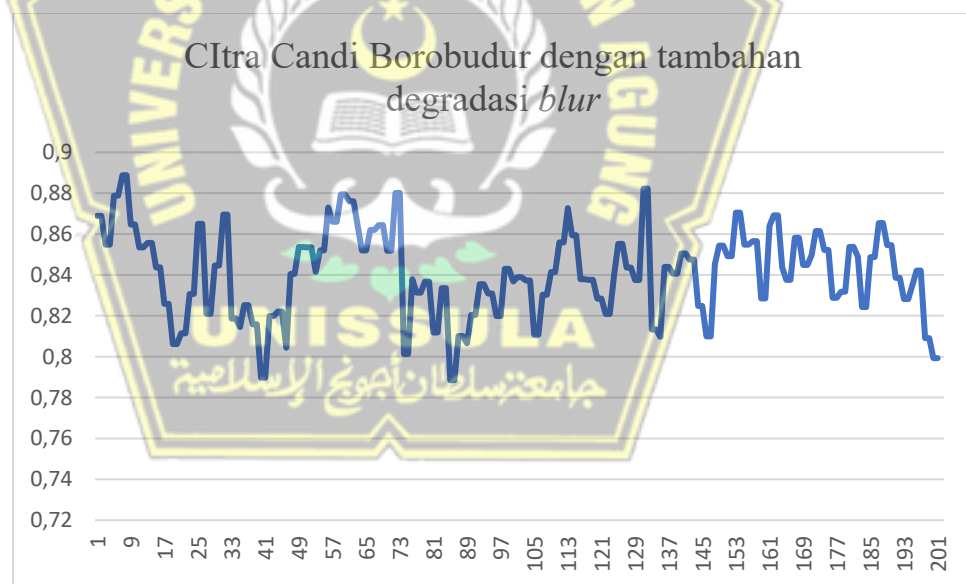
Gambar 4. 14 Citra Candi Borobudur *original (output)*

2. Citra Candi Borobudur dengan tambahan degradasi *blur*

Tabel 4. 5 Hasil evaluasi citra Candi Borobudur dengan tambahan degradasi *blur*

Epoch	CLIP-IQA (Input)	CLIP-IQA (Output)
6	0,7806	0,8888
50	0,7806	0,8535
100	0,7806	0,8389
150	0,7806	0,8492
200	0,7806	0,7993

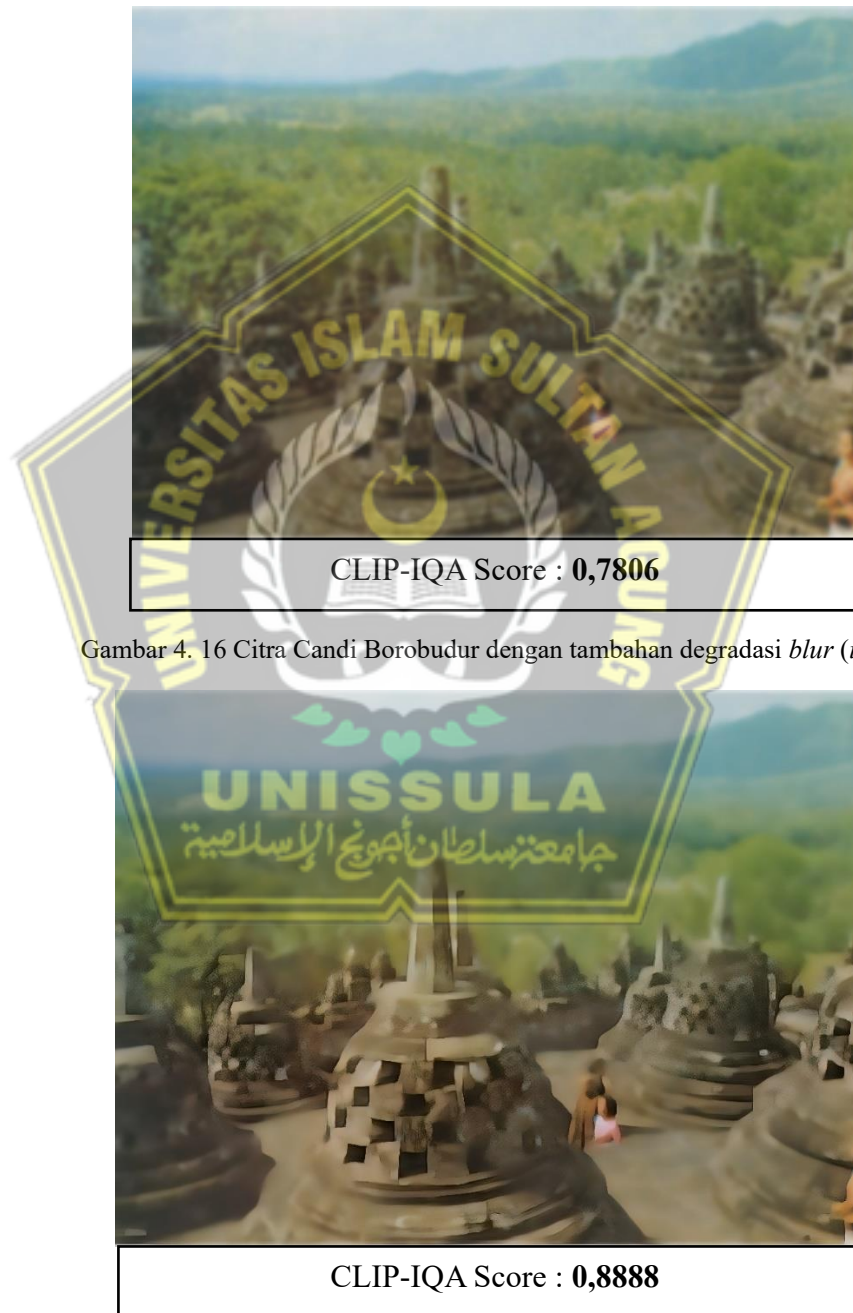
Tabel 4.5 menunjukkan hasil pengujian kualitas citra menggunakan metode CLIP-IQA pada citra asli Candi Borobudur dengan tambahan degradasi *blur* dengan berbagai model hasil pelatihan pada epoch berbeda. Skor tertinggi diperoleh pada epoch ke-6 sebesar 0,8888 sehingga menunjukkan memberikan peningkatan kualitas citra. Hasil tersebut dapat dilihat dari grafik pengujian berikut:



Gambar 4. 15 Grafik hasil evaluasi citra Candi Borobudur dengan tambahan degradasi *blur*

Gambar 4.15 menunjukkan grafik skor evaluasi CLIP-IQA *output* terhadap citra asli Candi Borobudur yang telah ditingkatkan resolusinya menggunakan model Real-ESRGAN pada berbagai epoch pelatihan. Terlihat bahwa nilai skor CLIP-IQA berfluktuasi seiring bertambahnya epoch. Secara umum, skor CLIP-IQA cenderung stabil menunjukkan

bahwa model dapat mempertahankan kualitas perseptual citra hasil selama pelatihan. Namun, perlu pemilihan epoch terbaik supaya mempertimbangkan kestabilan dan puncak skor CLIP-IQA bukan hanya berdasarkan banyaknya iterasi pelatihan. Hasil dari *input* dan *output* dapat dilihat pada gambar berikut :



Gambar 4. 16 Citra Candi Borobudur dengan tambahan degradasi *blur* (*input*)

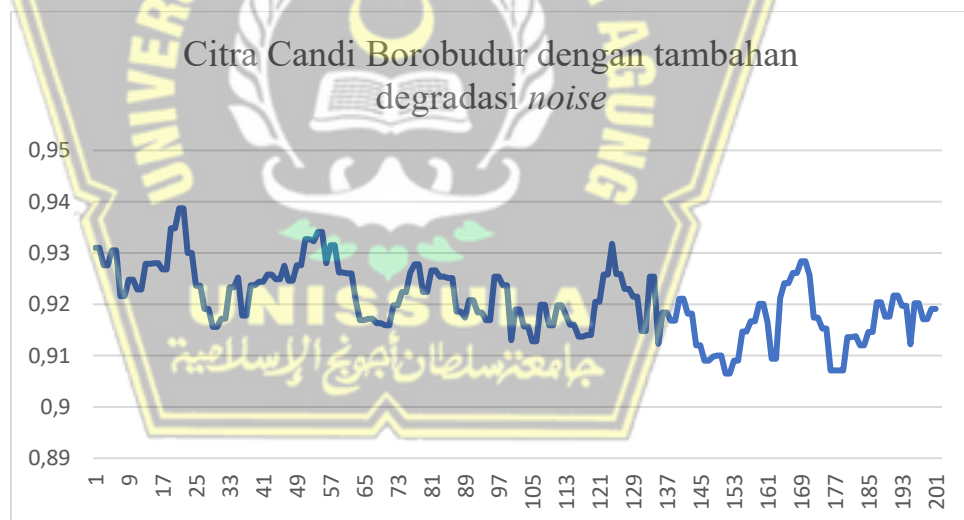
Gambar 4. 17 Citra Candi Borobudur dengan tambahan degradasi *blur* (*output*)

3. Citra Candi Borobudur dengan tambahan degradasi *noise*

Tabel 4. 6 Hasil evaluasi Citra Candi Borobudur dengan tambahan degradasi *noise*

Epoch	CLIP-IQA (Input)	CLIP-IQA (Output)
20	0,8154	0,9387
50	0,8154	0,9327
100	0,8154	0,919
150	0,8154	0,9065
200	0,8154	0,9191

Tabel 4.6 menunjukkan hasil pengujian kualitas citra menggunakan metode CLIP-IQA pada citra asli Candi Borobudur dengan tambahan degradasi *noise* dengan berbagai model hasil pelatihan pada epoch berbeda. Skor tertinggi diperoleh pada epoch ke-20 sebesar 0,9387 sehingga menunjukkan memberikan peningkatan kualitas citra. Hasil tersebut dapat dilihat dari grafik pengujian berikut:



Gambar 4. 18 Grafik hasil evaluasi Citra Candi Borobudur dengan tambahan degradasi *noise*

Gambar 4.18 menunjukkan grafik skor evaluasi CLIP-IQA *output* terhadap citra asli Candi Borobudur yang telah ditingkatkan resolusinya menggunakan model Real-ESRGAN pada berbagai epoch pelatihan. Terlihat bahwa nilai skor CLIP-IQA berfluktuasi seiring bertambahnya epoch. Secara umum, skor CLIP-IQA cenderung stabil menunjukkan bahwa model dapat mempertahankan kualitas perseptual citra hasil selama

pelatihan. Namun, perlu pemilihan epoch terbaik supaya mempertimbangkan kestabilan dan puncak skor CLIP-IQA bukan hanya berdasarkan banyaknya iterasi pelatihan. Hasil dari *input* dan *output* dapat dilihat pada gambar berikut :



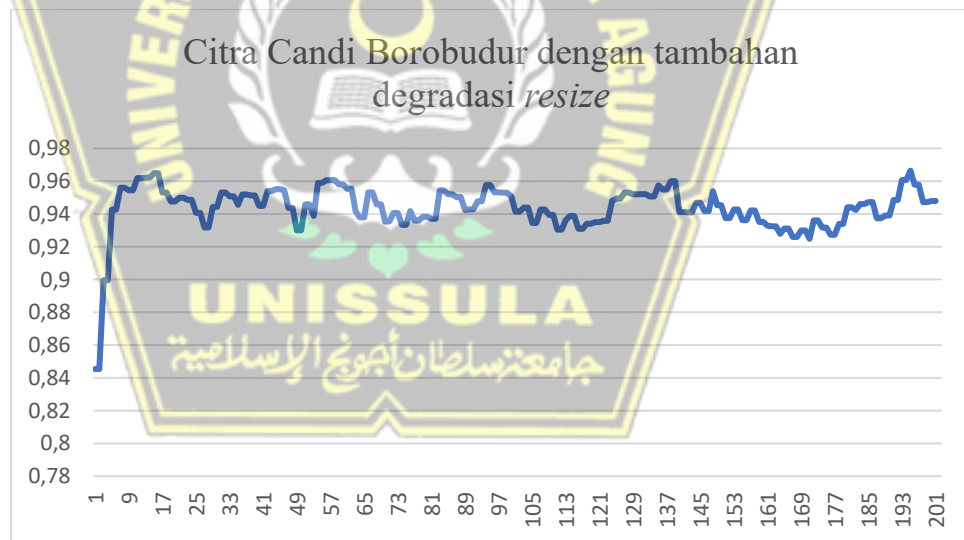
Gambar 4. 20 Citra Candi Borobudur dengan tambahan degradasi *noise* (*output*)

4. Citra Candi Borobudur dengan tambahan degradasi *resize*

Tabel 4. 7 Hasil evaluasi Citra Candi Borobudur dengan tambahan degradasi *resize*

Epoch	CLIP-IQA (Input)	CLIP-IQA (Output)
50	0,7979	0,9458
75	0,7979	0,9417
100	0,7979	0,9413
194	0,7979	0,9664
200	0,7979	0,9479

Tabel 4.7 menunjukkan hasil pengujian kualitas citra menggunakan metode CLIP-IQA pada citra asli Candi Borobudur dengan tambahan degradasi *resize* dengan berbagai model hasil pelatihan pada epoch berbeda. Skor tertinggi diperoleh pada epoch ke-194 sebesar 0,9664 sehingga menunjukkan memberikan peningkatan kualitas citra. Hasil tersebut dapat dilihat dari grafik pengujian berikut:



Gambar 4. 21 Grafik hasil evaluasi Citra Candi Borobudur dengan tambahan degradasi *resize*

Gambar 4.21 menunjukkan grafik skor evaluasi CLIP-IQA *output* terhadap citra asli Candi Borobudur yang telah ditingkatkan resolusinya menggunakan model Real-ESRGAN pada berbagai epoch pelatihan. Terlihat bahwa nilai skor CLIP-IQA berfluktuasi seiring bertambahnya epoch. Secara umum, skor CLIP-IQA cenderung stabil menunjukkan

bahwa model dapat mempertahankan kualitas perseptual citra hasil selama pelatihan. Namun, perlu pemilihan epoch terbaik supaya mempertimbangkan kestabilan dan puncak skor CLIP-IQA bukan hanya berdasarkan banyaknya iterasi pelatihan. Hasil dari *input* dan *output* dapat dilihat pada gambar berikut :



Gambar 4. 22 Citra Candi Borobudur dengan tambahan degradasi *resize (input)*



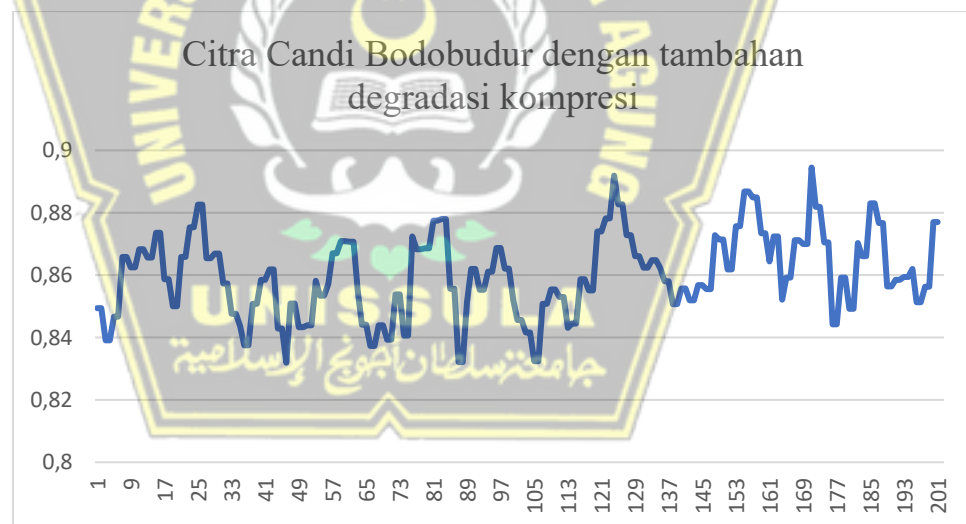
Gambar 4. 23 Citra Candi Borobudur dengan tambahan degradasi *resize (output)*

5. Citra Candi Borobudur dengan tambahan degradasi kompresi

Tabel 4. 8 Hasil evaluasi Citra Candi Borobudur dengan tambahan degradasi kompresi

Epoch	CLIP-IQA (Input)	CLIP-IQA (Output)
50	0,3918	0,8439
125	0,3918	0,8827
150	0,3918	0,8618
170	0,3918	0,8945
200	0,3918	0,877

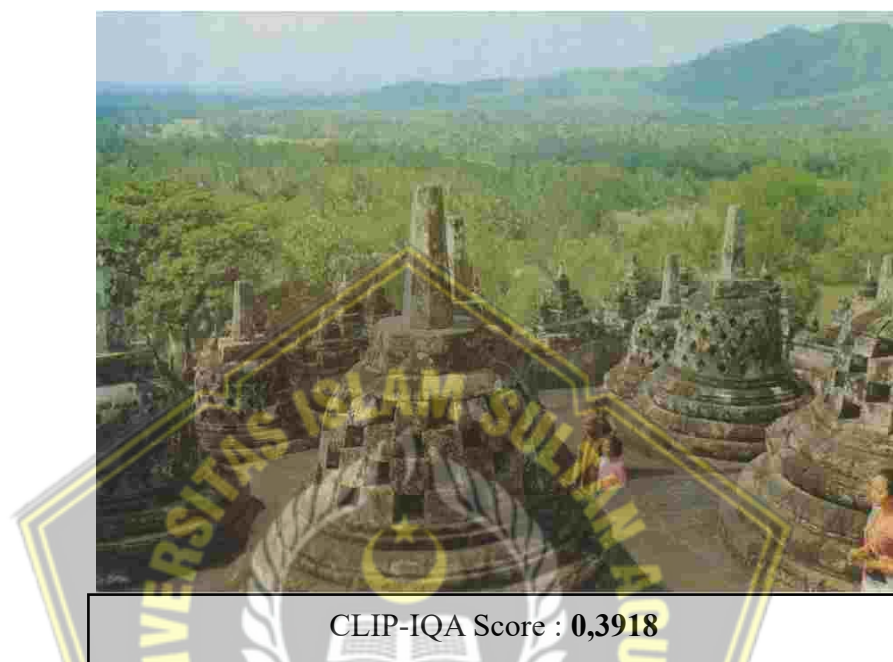
Tabel 4.8 menunjukkan hasil pengujian kualitas citra menggunakan metode CLIP-IQA pada citra asli Candi Borobudur dengan tambahan degradasi kompresi berbagai model hasil pelatihan pada epoch berbeda. Skor tertinggi diperoleh pada epoch ke-170 sebesar 0,8945 sehingga menunjukkan memberikan peningkatan kualitas citra. Hasil tersebut dapat dilihat dari grafik pengujian berikut:



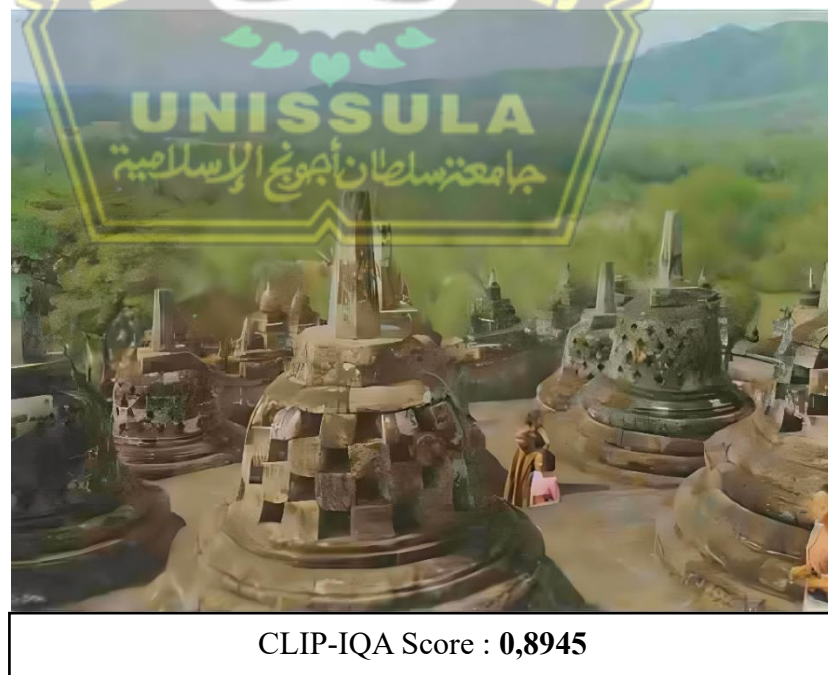
Gambar 4. 24 Grafik hasil evaluasi citra Candi Borobudur dengan tambahan degradasi kompresi

Gambar 4.24 menunjukkan grafik skor evaluasi CLIP-IQA *output* terhadap citra asli Candi Borobudur yang telah ditingkatkan resolusinya menggunakan model Real-ESRGAN pada berbagai epoch pelatihan. Terlihat bahwa nilai skor CLIP-IQA berfluktuasi seiring bertambahnya epoch. Secara umum, skor CLIP-IQA cenderung stabil menunjukkan bahwa model dapat mempertahankan kualitas perseptual citra hasil selama

pelatihan. Namun, perlu pemilihan epoch terbaik supaya mempertimbangkan kestabilan dan puncak skor CLIP-IQA bukan hanya berdasarkan banyaknya iterasi pelatihan. Hasil dari *input* dan *output* dapat dilihat pada gambar berikut :



Gambar 4. 25 Citra Candi Borobudur dengan tambahan degradasi kompresi (*input*)



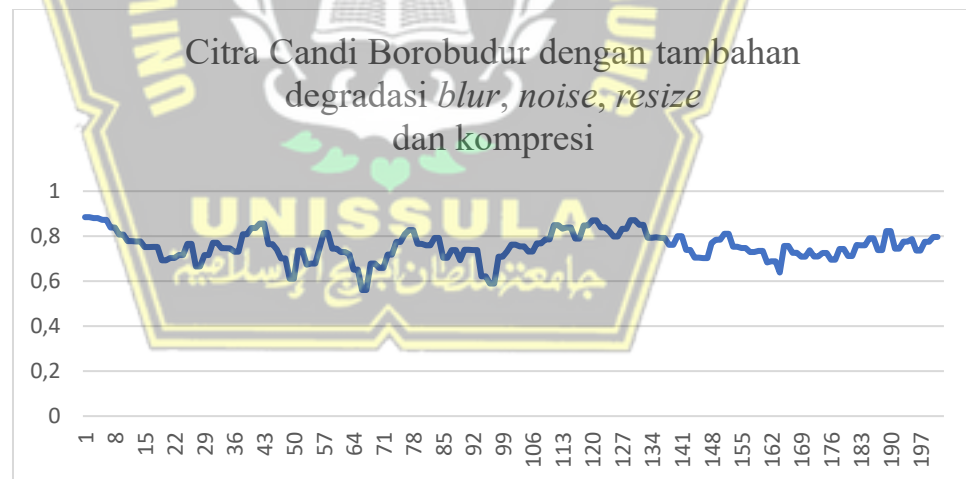
Gambar 4. 26 Citra Candi Borobudur dengan tambahan degradasi kompresi (*output*)

6. Citra Candi Borobudur dengan tambahan degradasi *blur*, *noise*, *resize* dan kompresi

Tabel 4. 9 Hasil evaluasi Citra Candi Borobudur dengan tambahan degradasi *blur*, *noise*, *resize* dan kompresi

Epoch	CLIP-IQA (Input)	CLIP-IQA (Output)
0	0,4026	0,8846
50	0,4026	0,736
100	0,4026	0,7626
150	0,4026	0,811
200	0,4026	0,7965

Tabel 4.9 menunjukkan hasil pengujian kualitas citra menggunakan metode CLIP-IQA pada citra asli Candi Borobudur dengan tambahan 4 (empat) degradasi pada berbagai model hasil pelatihan pada epoch berbeda. Skor tertinggi diperoleh pada epoch ke-0 sebesar 0,8846 sehingga menunjukkan memberikan peningkatan kualitas citra. Hasil tersebut dapat dilihat dari grafik pengujian berikut:



Gambar 4. 27 Grafik hasil evaluasi Citra Candi Borobudur dengan tambahan degradasi *blur*, *noise*, *resize* dan kompresi

Gambar 4.27 menunjukkan grafik skor evaluasi CLIP-IQA *output* terhadap citra asli Candi Borobudur yang telah ditingkatkan resolusinya menggunakan model Real-ESRGAN pada berbagai epoch pelatihan. Terlihat bahwa nilai skor CLIP-IQA berfluktuasi seiring bertambahnya epoch. Secara umum, skor CLIP-IQA cenderung stabil menunjukkan

bahwa model dapat mempertahankan kualitas perseptual citra hasil selama pelatihan. Namun, perlu pemilihan epoch terbaik supaya mempertimbangkan kestabilan dan puncak skor CLIP-IQA bukan hanya berdasarkan banyaknya iterasi pelatihan. Hasil dari *input* dan *output* dapat dilihat pada gambar berikut :



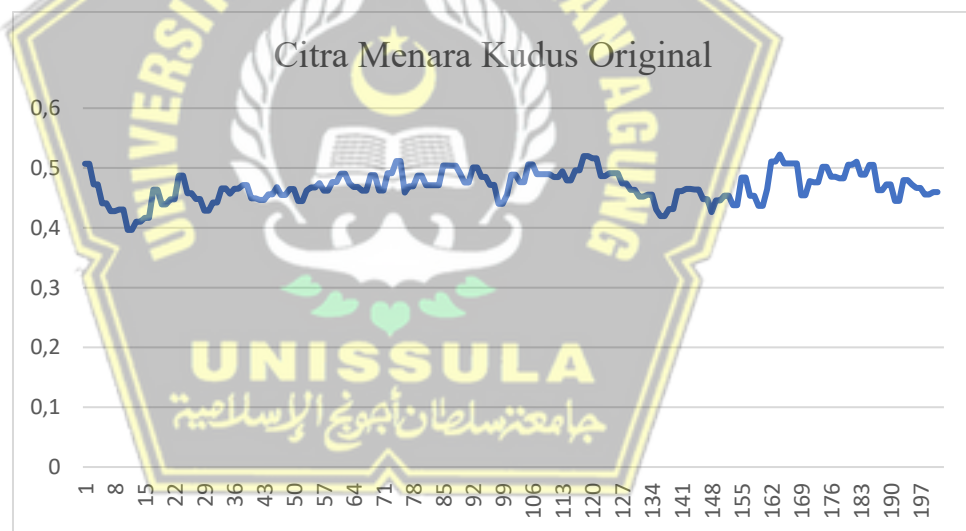
Gambar 4. 29 Citra Candi Borobudur dengan tambahan degradasi *blur, noise, resize* dan kompresi (*output*)

7. Citra Menara Kudus Original

Tabel 4. 10 Hasil evaluasi citra Menara Kudus Original

Epoch	CLIP-IQA (Input)	CLIP-IQA (Output)
50	0,4007	0,4447
100	0,4007	0,4889
125	0,4007	0,4912
163	0,4007	0,5225
200	0,4007	0,4598

Tabel 4.10 menunjukkan hasil pengujian kualitas citra menggunakan metode CLIP-IQA pada citra asli Menara Kudus pada berbagai model hasil pelatihan pada epoch berbeda. Skor tertinggi diperoleh pada epoch ke-163 sebesar 0,5225 sehingga menunjukkan memberikan peningkatan kualitas citra. Hasil tersebut dapat dilihat dari grafik pengujian berikut:



Gambar 4. 30 Grafik hasil evaluasi citra Menara Kudus Original

Gambar 4.30 menunjukkan grafik skor evaluasi CLIP-IQA *output* terhadap citra asli Menara Kudus yang telah ditingkatkan resolusinya menggunakan model Real-ESRGAN pada berbagai epoch pelatihan. Terlihat bahwa nilai skor CLIP-IQA berfluktuasi seiring bertambahnya epoch. Secara umum, skor CLIP-IQA cenderung stabil menunjukkan bahwa model dapat mempertahankan kualitas perseptual citra hasil selama pelatihan. Namun, perlu pemilihan epoch terbaik supaya mempertimbangkan kestabilan dan

puncak skor CLIP-IQA bukan hanya berdasarkan banyaknya iterasi pelatihan. Hasil dari *input* dan *output* dapat dilihat pada gambar berikut :



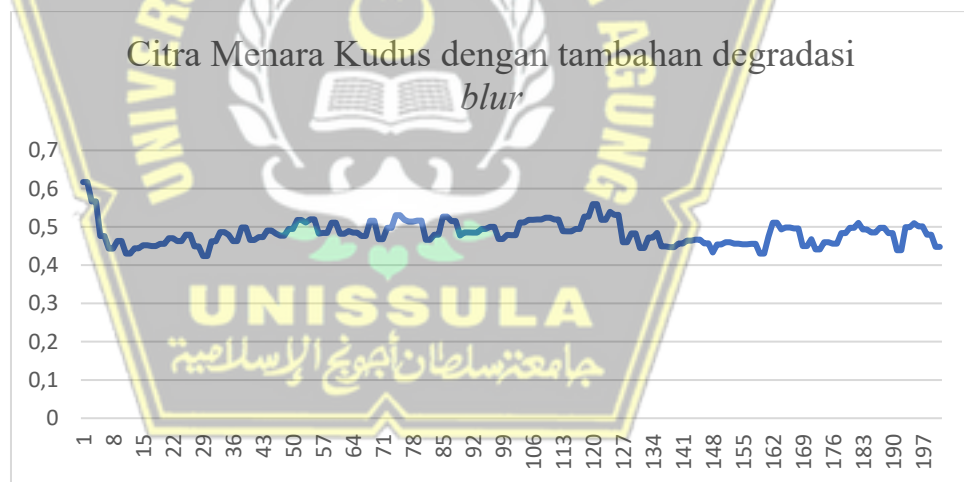
Gambar 4. 32 Citra Menara Kudus Original (*output*)

8. Citra Menara Kudus dengan tambahan degradasi *blur*

Tabel 4. 11 Hasil evaluasi citra Menara Kudus dengan tambahan degradasi *blur*

Epoch	CLIP-IQA (Input)	CLIP-IQA (Output)
0	0,357	0,6171
50	0,357	0,5182
100	0,357	0,4782
150	0,357	0,46
200	0,357	0,448

Tabel 4.11 menunjukkan hasil pengujian kualitas citra menggunakan metode CLIP-IQA pada citra asli Menara Kudus dengan tambahan degradasi *blur* pada berbagai model hasil pelatihan pada epoch berbeda. Skor tertinggi diperoleh pada epoch ke-0 sebesar 0,6171 sehingga menunjukkan memberikan peningkatan kualitas citra. Hasil tersebut dapat dilihat dari grafik pengujian berikut:



Gambar 4. 33 Grafik hasil evaluasi citra Menara Kudus dengan tambahan degradasi *blur*

Gambar 4.33 menunjukkan grafik skor evaluasi CLIP-IQA *output* terhadap citra asli Menara Kudus yang telah ditingkatkan resolusinya menggunakan model Real-ESRGAN pada berbagai epoch pelatihan. Terlihat bahwa nilai skor CLIP-IQA berfluktuasi seiring bertambahnya epoch. Secara umum, skor CLIP-IQA cenderung stabil menunjukkan bahwa model dapat mempertahankan kualitas perseptual citra hasil selama pelatihan. Namun, perlu pemilihan epoch terbaik supaya mempertimbangkan kestabilan dan

puncak skor CLIP-IQA bukan hanya berdasarkan banyaknya iterasi pelatihan. Hasil dari *input* dan *output* dapat dilihat pada gambar berikut :



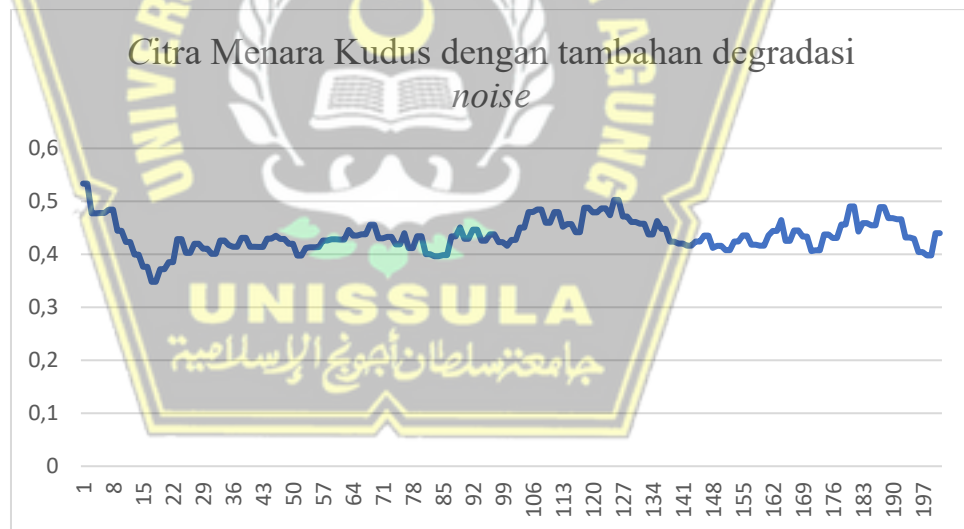
Gambar 4. 35 Citra Menara Kudus dengan tambahan degradasi blur (*output*)

9. Citra Menara Kudus dengan tambahan degradasi *noise*

Tabel 4. 12 Hasil evaluasi citra Menara Kudus dengan tambahan degradasi *noise*

Epoch	CLIP-IQA (Input)	CLIP-IQA (Output)
0	0,3808	0,5335
50	0,3808	0,3976
100	0,3808	0,4271
150	0,3808	0,4079
200	0,3808	0,4399

Tabel 4.12 menunjukkan hasil pengujian kualitas citra menggunakan metode CLIP-IQA pada citra asli Menara Kudus dengan tambahan degradasi *noise* pada berbagai model hasil pelatihan pada epoch berbeda. Skor tertinggi diperoleh pada epoch ke-0 sebesar 0,5335 sehingga menunjukkan memberikan peningkatan kualitas citra. Hasil tersebut dapat dilihat dari grafik pengujian berikut:



Gambar 4. 36 Grafik Hasil evaluasi citra Menara Kudus dengan tambahan degradasi *noise*

Gambar 4.36 menunjukkan grafik skor evaluasi CLIP-IQA *output* terhadap citra asli Menara Kudus yang telah ditingkatkan resolusinya menggunakan model Real-ESRGAN pada berbagai epoch pelatihan. Terlihat bahwa nilai skor CLIP-IQA berfluktuasi seiring bertambahnya epoch. Secara umum, skor CLIP-IQA cenderung stabil menunjukkan bahwa model dapat mempertahankan kualitas perseptual citra hasil selama pelatihan. Namun,

perlu pemilihan epoch terbaik supaya mempertimbangkan kestabilan dan puncak skor CLIP-IQA bukan hanya berdasarkan banyaknya iterasi pelatihan. Hasil dari *input* dan *output* dapat dilihat pada gambar berikut :



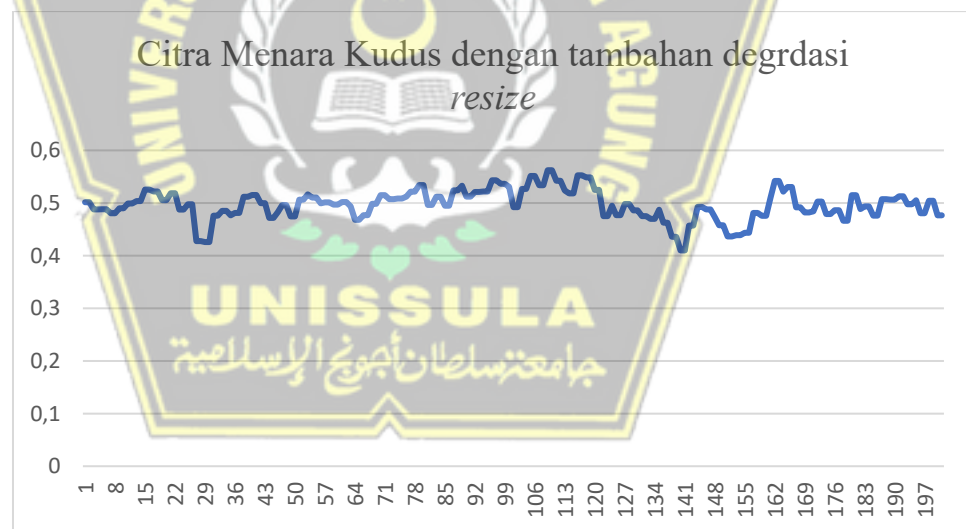
Gambar 4. 38 Citra Menara Kudus dengan tambahan degradasi *noise* (*output*)

10. Citra Menara Kudus dengan tambahan degradasi *resize*

Tabel 4. 13 Hasil evaluasi citra Menara Kudus dengan tambahan degradasi *resize*

Epoch	CLIP-IQA (Input)	CLIP-IQA (Output)
50	0,3896	0,5063
75	0,3896	0,5122
108	0,3896	0,5623
150	0,3896	0,4366
200	0,3896	0,4765

Tabel 4.13 menunjukkan hasil pengujian kualitas citra menggunakan metode CLIP-IQA pada citra asli Menara Kudus dengan tambahan degradasi *resize* pada berbagai model hasil pelatihan pada epoch berbeda. Skor tertinggi diperoleh pada epoch ke-108 sebesar 0,5623 sehingga menunjukkan memberikan peningkatan kualitas citra. Hasil tersebut dapat dilihat dari grafik pengujian berikut:



Gambar 4. 39 Grafik hasil evaluasi citra Menara Kudus dengan tambahan degradasi *resize*

Gambar 4.39 menunjukkan grafik skor evaluasi CLIP-IQA *output* terhadap citra asli Menara Kudus yang telah ditingkatkan resolusinya menggunakan model Real-ESRGAN pada berbagai epoch pelatihan. Terlihat bahwa nilai skor CLIP-IQA berfluktuasi seiring bertambahnya epoch. Secara umum, skor CLIP-IQA cenderung stabil menunjukkan bahwa model dapat mempertahankan kualitas perseptual citra hasil selama pelatihan. Namun,

perlu pemilihan epoch terbaik supaya mempertimbangkan kestabilan dan puncak skor CLIP-IQA bukan hanya berdasarkan banyaknya iterasi pelatihan. Hasil dari *input* dan *output* dapat dilihat pada gambar berikut :



Gambar 4. 40 Citra Menara Kudus dengan tambahan degradasi *resize* (*input*)

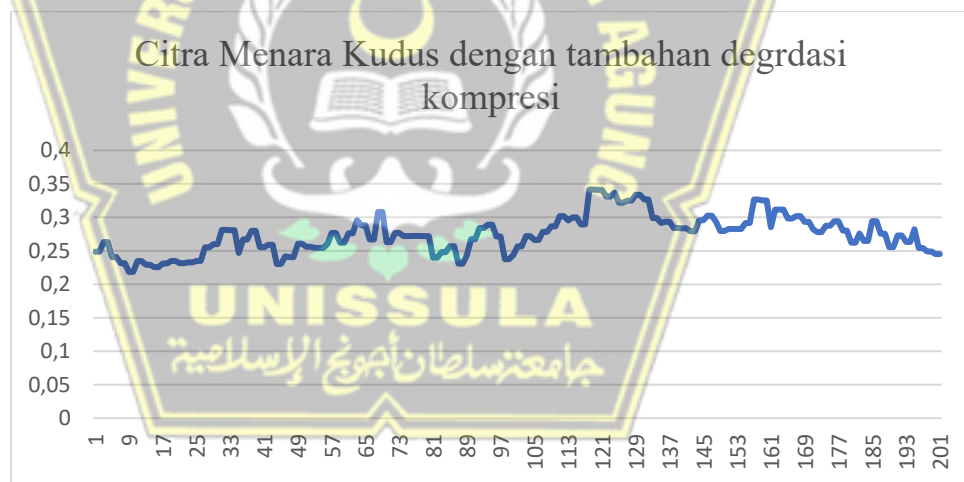
Gambar 4. 41 Citra Menara Kudus dengan tambahan degradasi *resize* (*output*)

11. Citra Menara Kudus dengan tambahan degradasi kompresi

Tabel 4. 14 Hasil evaluasi citra Menara Kudus dengan tambahan degradasi kompresi

Epoch	CLIP-IQA (Input)	CLIP-IQA (Output)
50	0,2247	0,2558
117	0,2247	0,3415
150	0,2247	0,2827
175	0,2247	0,2942
200	0,2247	0,2451

Tabel 4.14 menunjukkan hasil pengujian kualitas citra menggunakan metode CLIP-IQA pada citra asli Menara Kudus dengan tambahan degradasi kompresi pada berbagai model hasil pelatihan pada epoch berbeda. Skor tertinggi diperoleh pada epoch ke-117 sebesar 0,3415 sehingga menunjukkan memberikan peningkatan kualitas citra. Hasil tersebut dapat dilihat dari grafik pengujian berikut:



Gambar 4. 42 Grafik hasil evaluasi citra Menara Kudus dengan tambahan degradasi kompresi

Gambar 4.42 menunjukkan grafik skor evaluasi CLIP-IQA *output* terhadap citra asli Menara Kudus yang telah ditingkatkan resolusinya menggunakan model Real-ESRGAN pada berbagai epoch pelatihan. Terlihat bahwa nilai skor CLIP-IQA berfluktuasi seiring bertambahnya epoch. Secara umum, skor CLIP-IQA cenderung stabil menunjukkan bahwa model dapat mempertahankan kualitas perseptual citra hasil selama pelatihan. Namun, perlu pemilihan epoch terbaik supaya mempertimbangkan kestabilan dan

puncak skor CLIP-IQA bukan hanya berdasarkan banyaknya iterasi pelatihan. Hasil dari *input* dan *output* dapat dilihat pada gambar berikut :



Gambar 4. 43 Citra Menara Kudus dengan tambahan degradasi kompresi (*input*)

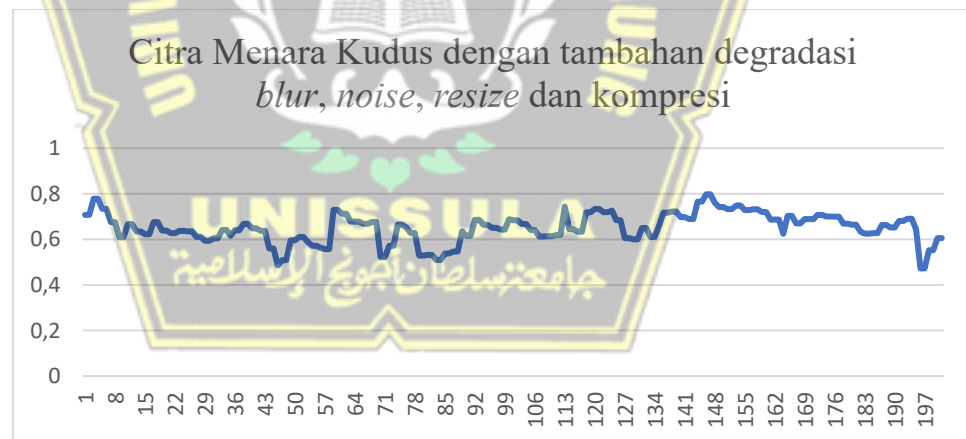
Gambar 4. 44 Citra Menara Kudus dengan tambahan degradasi kompresi (*output*)

12. Citra Menara Kudus dengan tambahan degradasi *blur*, *noise*, *resize* dan kompresi

Tabel 4. 15 Hasil evaluasi Citra Menara Kudus dengan tambahan degradasi *blur*, *noise*, *resize* dan kompresi

Epoch	CLIP-IQA (Input)	CLIP-IQA (Output)
50	0,3303	0,6109
100	0,3303	0,684
145	0,3303	0,7978
175	0,3303	0,7001
200	0,3303	0,6058

Tabel 4.15 menunjukkan hasil pengujian kualitas citra menggunakan metode CLIP-IQA pada citra asli Menara Kudus dengan tambahan 4 (empat) degradasi pada berbagai model hasil pelatihan pada epoch berbeda. Skor tertinggi diperoleh pada epoch ke-145 sebesar 0,7978 sehingga menunjukkan memberikan peningkatan kualitas citra. Hasil tersebut dapat dilihat dari grafik pengujian berikut:



Gambar 4. 45 Grafik hasil evaluasi Citra Menara Kudus dengan tambahan degradasi *blur*, *noise*, *resize* dan kompresi

Gambar 4.45 menunjukkan grafik skor evaluasi CLIP-IQA *output* terhadap citra asli Menara Kudus yang telah ditingkatkan resolusinya menggunakan model Real-ESRGAN pada berbagai epoch pelatihan. Terlihat bahwa nilai skor CLIP-IQA berfluktuasi seiring bertambahnya epoch. Secara umum, skor CLIP-IQA cenderung stabil menunjukkan bahwa model dapat mempertahankan kualitas perseptual citra hasil selama pelatihan. Namun,

perlu pemilihan epoch terbaik supaya mempertimbangkan kestabilan dan puncak skor CLIP-IQA bukan hanya berdasarkan banyaknya iterasi pelatihan. Hasil dari *input* dan *output* dapat dilihat pada gambar berikut :



Gambar 4. 47 Citra Menara Kudus dengan tambahan degradasi *blur*, *noise*, *resize* dan kompresi (*output*)

4.1.6 Deployment

Tahap *deployment* merupakan implementasi model pada media berbasis *website*. Pada tahap ini menampilkan hasil implementasi dari rancangan pada bab sebelumnya. Adapun implementasi dapat dilihat pada beberapa gambar 4.48 dan 4.49 berikut.

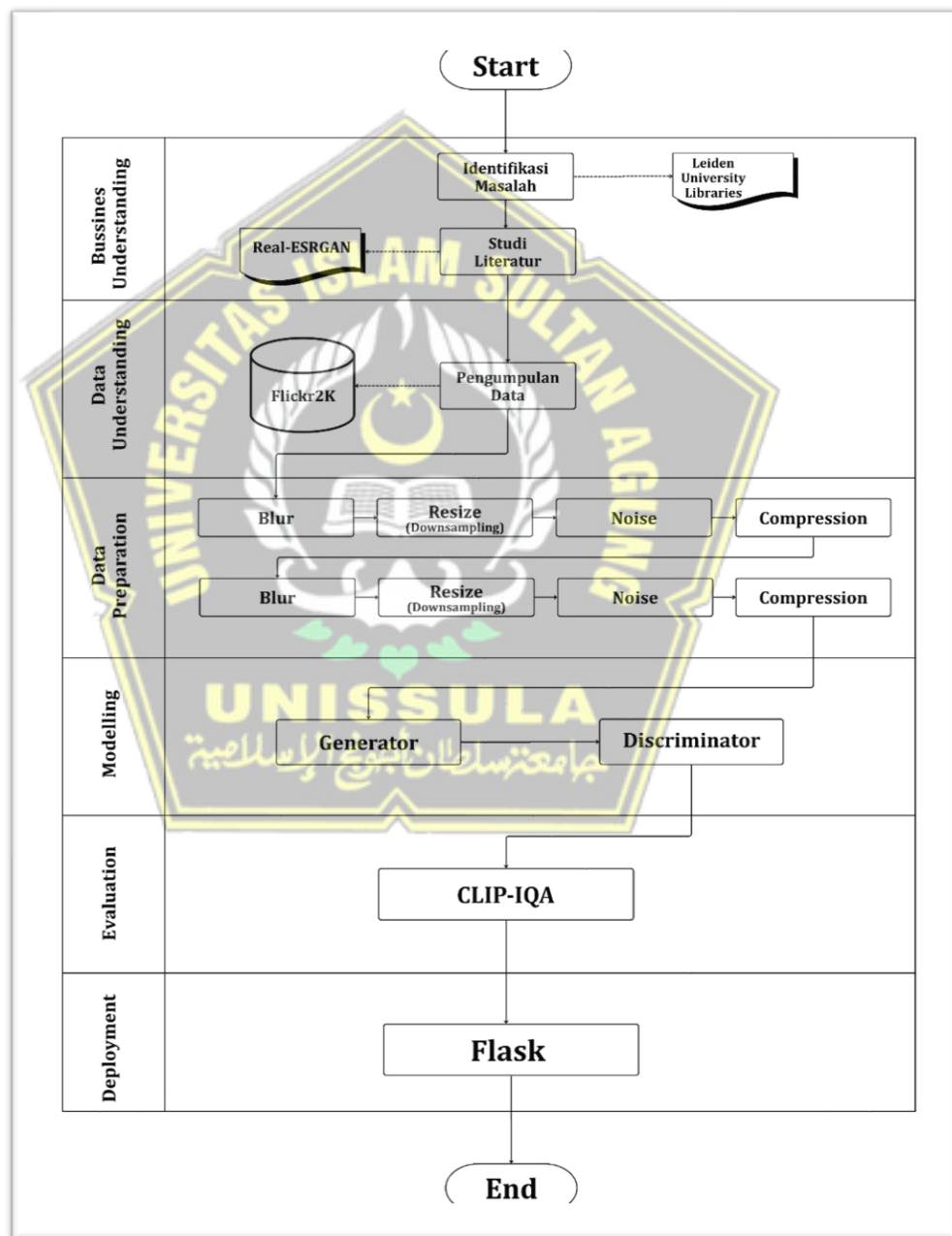


Gambar 4. 49 Halaman *output* citra resolusi tinggi

Halaman selanjutnya ialah output dari aplikasi setelah upload file. Halaman ini dapat dilihat pada gambar 4.49. Hasil pembesaran resolusi citra pada halaman ini. Selain itu, citra asli juga ditampilkan pada halaman ini. Citra dapat diunduh dengan menekan tombol download.

4.2 Analisis Penelitian

4.2.1 Implementasi Metode Real-ESRGAN



Gambar 4. 50 Flowchart Implementasi metode Real-ESRGAN

Penerapan metode Real-ESRGAN pada citra bangunan bersejarah di Indonesia dilakukan dengan menggunakan metode CRISP-DM. Tahapan-tahapan yang dilakukan pada penelitian ini dapat dilakukan sesuai rencana. Penerapan Real-ESRGAN dilakukan pada dataset citra bangunan bersejarah dengan jumlah data 2000 citra dari flickr2k untuk data *training* dan 2 citra dari Leiden University Libraries yang memiliki degradasi tetapi tidak diketahui jenis degradasinya untuk data *testing*. Pada tahap *preprocessing*, dilakukan *high-order degradation model* pada citra untuk menghasilkan citra resolusi rendah yang mensimulasikan degradasi kompleks.

Data yang telah dilakukan tahapan *preprocessing* kemudian dipakai untuk pemodelan. Pemodelan Real-ESRGAN dilakukan dengan menggunakan 2 jenis model. Model *generator* untuk menghasilkan citra resolusi tinggi dan model discriminator yang dipakai untuk melatih model secara *adversarial*. Pada penelitian ini, pelatihan model Real-ESRGAN dilakukan sebanyak 200 epoch. Hal ini tentunya cukup jauh dari pelatihan yang baik untuk model arsitektur GAN yang biasanya dapat mencapai ribuan epoch. Kekurangan-kekurangan ini disebabkan kurangnya sumber daya dan waktu yang memadai untuk melakukan pelatihan dengan lebih banyak epoch. Kekurangan tersebut dapat diatasi dengan menggunakan komputer dengan spesifikasi tinggi atau cloud processing seperti google colab atau kaggle notebook. *Graphic Processing Unit* (GPU) dengan kinerja tinggi juga diperlukan untuk pelatihan model Real-ESRGAN dengan lebih cepat.

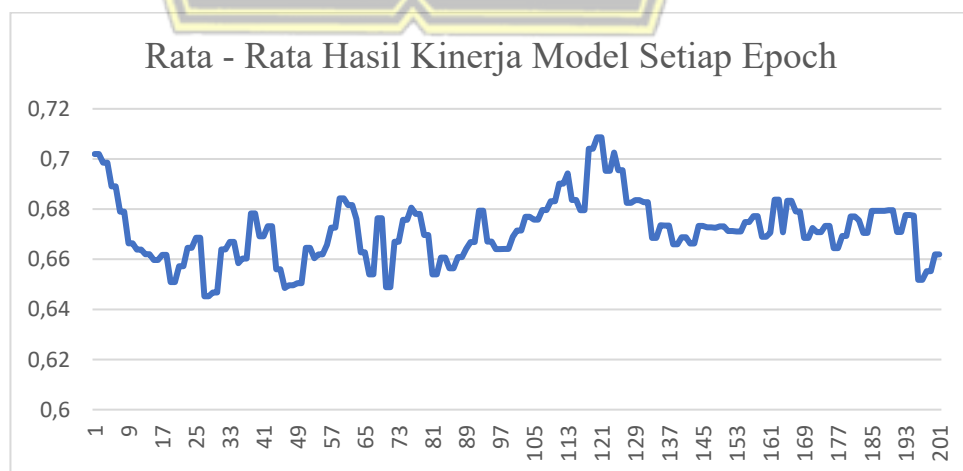
Setelah tahap pemodelan, dilakukan pengujian model. Model Real-ESRGAN yang terbaik kemudian diterapkan pada media berbasis *website* menggunakan *framework flask*. Dengan begitu, pengguna dapat menggunakan *website* untuk meningkatkan resolusi citra bangunan bersejarah. model Real-ESRGAN dapat diimplementasikan dengan baik pada media berbasis *website* dan dapat berjalan tanpa kendala. Aplikasi dapat menjalankan model untuk meningkatkan resolusi citra.

4.2.2 Kinerja Metode Real-ESRGAN

Tabel 4. 16 Hasil rata-rata evaluasi

Epoch	Rata-Rata CLIP-IQA Score
50	0,66457
119	0,70863
150	0,67124
175	0,66444
200	0,66195

Kinerja metode Real-ESRGAN dalam penelitian ini diuji menggunakan metode *Contrastive Language–Image Pretraining Image Quality Assessment* (CLIP-IQA) yang merupakan pendekatan *no-reference* dalam menilai kualitas persepsi visual citra. Pengujian dilakukan berdasarkan jumlah epoch pelatihan, dengan menilai skor CLIP-IQA pada citra hasil yang telah ditingkatkan resolusinya. Berdasarkan hasil pengujian mendapatkan hasil rata-rata dengan perhitungan *total* hasil dibagi *total* citra, model pada epoch ke-119 menunjukkan performa terbaik dengan rata-rata skor CLIP-IQA *output* sebesar 0,70863 sehingga menandakan bahwa citra hasil dari model tersebut memiliki tingkat kemiripan perseptual tertinggi dan membuktikan bahwa Real-ESRGAN memiliki kemampuan yang baik dalam meningkatkan kualitas persepsi visual citra bangunan bersejarah di Indonesia. Adapun grafik keseluruhan rata-rata dari kinerja model pada setiap *epoch* sebagai berikut :



Gambar 4. 51 Grafik rata – rata hasil kinerja model pada setiap *epoch*

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Penelitian ini bertujuan untuk merancang dan mengimplementasikan model Real-ESRGAN dalam upaya meningkatkan resolusi pada citra bangunan bersejarah di Indonesia. Seluruh tahapan yang dirancang dalam proses penelitian, mulai dari persiapan data, penerapan degradasi kompleks, hingga pelatihan model telah berhasil dijalankan sesuai rencana

Penelitian ini juga bertujuan untuk mengevaluasi kinerja Real-ESRGAN dalam meningkatkan resolusi citra bangunan bersejarah Indonesia dari koleksi Leiden University Libraries. Citra tersebut mengalami degradasi visual akibat usia, proses digitalisasi, dan keterbatasan teknologi dokumentasi. Namun, jenis degradasi yang terjadi tidak diketahui secara pasti sehingga pendekatan yang digunakan difokuskan pada peningkatan resolusi citra secara umum tanpa diketahui jenis degradasinya. Evaluasi kinerja model dilakukan menggunakan metrik *no-reference* CLIP-IQA, yang menilai kualitas citra berdasarkan kesesuaian perseptual terhadap representasi semantik. Hasil evaluasi menunjukkan peningkatan kualitas perseptual terbaik pada epoch ke-119 dengan skor CLIP-IQA sebesar 0,70863. Nilai tersebut mengindikasikan bahwa model Real-ESRGAN mampu dala meningkatkan resolusi citra bangunan bersejarah yang telah terdegradasi, meskipun tanpa informasi pasti mengenai jenis degradasi yang dialami.

5.2 Saran

Penelitian ini menghasilkan model Real-ESRGAN yang dapat meningkatkan resolusi citra bangunan bersejarah. Akan tetapi, masih terdapat kekurangan atau pun hal yang dapat diperbaiki pada penelitian selanjutnya. Adapun saran untuk penelitian selanjutnya ialah sebagai berikut:

1. Menggabungkan metode Real-ESRGAN dengan metode lain untuk meningkatkan kinerja model.

2. Menggunakan metode *preprocessing* yang berbeda untuk menghasilkan citra resolusi rendah yang beragam.
3. Melatih model dengan lebih banyak data dan jumlah epoch yang lebih banyak.
4. Mengidentifikasi jenis degradasi yang terjadi pada citra bangunan bersejarah agar lebih spesifik pada tahap *training*.



DAFTAR PUSTAKA

- Agustinova, D. E. (2022). *Strategy for Maintaining Cultural Heritage Objectives*. 18(2), 60–68.
- Darmarastri, H. A., Susanto, Sutirto, T. W., Dadtun, Y. S., & Kurniawati, A. (2023). Digitalisasi Katalog Foto Bangunan di Wilayah Mangkunegaran Koleksi Perpustakaan Reksa Pustaka Mangkunegaran Surakarta. *Memetika: Jurnal Kajian Budaya*, 5, 54–57. <https://jurnal.uns.ac.id/memetika/article/view/79632>
- Feng, R., Gu, J., Qiao, Y., & Dong, C. (2019). Suppressing model overfitting for image super-resolution networks. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, 2019-June*, 1964–1973. <https://doi.org/10.1109/CVPRW.2019.00248>
- Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). Generative adversarial networks. *Communications of the ACM*, 63(11), 139–144. <https://doi.org/10.1145/3422622>
- Hardiansyah, B., Armin, A. P., & Yunanda, A. B. (2019). Rekonstruksi Citra Pada Super Resolusi Menggunakan Interpolasi Bicubic. *INTEGER: Journal of Information Technology*, 4(2), 1–12. <https://doi.org/10.31284/j.integer.2019.v4i2.684>
- Kho, I. E., & Braganca, W. O. (2023). *Comparative Analysis of Python Microframeworks: Flask, Dash, and CherryPy*. July. <https://doi.org/10.13140/RG.2.2.17101.82402>
- Ledig, C., Theis, L., Husz, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A., Tejani, A., Totz, J., Wang, Z., & Shi, W. (2017). *Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network*. 4681–4690.
- Nandal, P., Pahal, S., Khanna, A., & Pinheiro, P. R. (2024). Super-resolution of medical images using real ESRGAN. *IEEE Access*, December, 176155–176170. <https://doi.org/10.1109/ACCESS.2024.3497002>

- Rohim, M. I. A., Nisa, A., Hindratno, M. N., Fajri, R., Wibowanto, G. S., Lestriandoko, N. H., & Normakristagaluh, P. (2024). Peningkatan Performa Pengenalan Wajah pada Gambar *Low-Resolution* Menggunakan Metode *Super-Resolution*. *Jurnal Teknologi Informasi Dan Ilmu Komputer*, 11(1), 199–208. <https://doi.org/10.25126/jtiik.20241117947>
- Rossum, G. van, & Boer, J. de. (1991). Interactively Testing Remote Servers Using the Python Programming Language. In *CWI Quarterly* (pp. 283–303).
- Safira, F., Salim, T. A., Rahmi, R., & Sani, M. K. J. A. (2020). Peran Arsip Dalam Pelestarian Cagar Budaya Di Indonesia: Sistematika Review. *Baca: Jurnal Dokumentasi Dan Informasi*, 41(2), 289. <https://doi.org/10.14203/j.baca.v41i2.593>
- Sarwito, J., Kesia, B., & Vania, C. (2024). *PENDEKATAN DIGITAL DALAM PENELITIAN SEJARAH ARSITEKTUR : TEKNOLOGI PEMODELAN DAN REKONSTRUKSI 3D PADA BANGUNAN KOLONIAL DI SEMARANG*. 4(1), 3–8. <https://doi.org/10.24167/joda.v4i1.12771>
- Shearer, C. (2000). The CRISP-DM model: the new blueprint for data mining. *Journal of Data Warehousing*, 5(4), 13–22.
- Sulaeman, A. (2022). *Implementasi algoritma srgan untuk meningkatkan resolusi pada citra wajah*. 2022.
- Wang, J., Chan, K. C. K., & Loy, C. C. (2022). Exploring CLIP for Assessing the Look and Feel of Images. *Proceedings of the 37th AAAI Conference on Artificial Intelligence, AAAI 2023*, 37, 2555–2563. <https://doi.org/10.1609/aaai.v37i2.25353>
- Wang, X., Xie, L., Dong, C., & Shan, Y. (2021). Real-ESRGAN: Training Real-World Blind Super-Resolution with Pure Synthetic Data. *Proceedings of the IEEE International Conference on Computer Vision, 2021-Octob*, 1905–1914. <https://doi.org/10.1109/ICCVW54120.2021.00217>
- Wang, X., Yu, K., Wu, S., Gu, J., Liu, Y., Dong, C., Qiao, Y., & Loy, C. C. (2018). ESRGAN: Enhanced super-resolution generative adversarial networks. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 11133 LNCS, 63–

79. https://doi.org/10.1007/978-3-030-11021-5_5

Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. (2004). Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4), 600–612. <https://doi.org/10.1109/TIP.2003.819861>

Ward, C. M., Harguess, J. D., Parameswaran, S., & Crabb, B. (2019). *Image quality assessment for determining efficacy and limitations of Super-Resolution Convolutional Neural Network (SRCNN)*. 5. <https://doi.org/10.1117/12.2275157>

Wolf, S., & Pinson, M. (2001). *Reference Algorithm for Computing Peak Signal to Noise Ratio (PSNR) of a Video Sequence with a Constant Delay*. 1–18.

Xu, T., Mi, P., Zheng, X., Li, L., Chao, F., Jiang, G., Zhang, W., Zhou, Y., & Ji, R. (2024). *Uncovering the Over-smoothing Challenge in Image Super-Resolution: Entropy-based Quantification and Contrastive Optimization*. XX(Xx), 1–21. <http://arxiv.org/abs/2201.01034>

Zhang, Y., & Yu, W. (2022). Comparison of DEM Super-Resolution Methods Based on Interpolation and Neural Networks. *Sensors*, 22(3). <https://doi.org/10.3390/s22030745>

