

***Rekognisi Wajah pada Citra Beresolusi Rendah  
Menggunakan ArcFace dengan Restorasi Berbasis  
GFPGAN***

*Diajukan Sebagai Syarat untuk Menyelesaikan  
Pendidikan Program Strata - 1  
pada  
Jurusan Teknik Informatika Fakultas Ilmu Komputer  
Universitas Sriwijaya*



Oleh

Philifs Bryan Sipahutar  
09021182227014

**JURUSAN TEKNIK INFORMATIKA  
FAKULTAS ILMU KOMPUTER  
UNIVERSITAS SRIWIJAYA  
2025**

# **BAB I**

## **PENDAHULUAN**

### **1.1 Pendahuluan**

Bab ini menjelaskan secara rinci tentang latar belakang masalah dimana didalamnya terdiri dari penelitian, rumusan masalah yang telah diidentifikasi, tujuan dari penelitian yang dilakukan, dan manfaat yang diharapkan dari hasil penelitian ini. Selain itu, bab ini juga akan membahas batasan - batasan masalah yang dihadapi dalam penelitian. Dengan demikian, bab ini memberikan gambaran yang menyeluruh mengenai seluruh aspek penelitian yang dilakukan.

### **1.2 Latar Belakang**

Teknologi pengenalan wajah telah menjadi salah satu cabang biometrik yang krusial dan banyak diimplementasikan dalam berbagai aplikasi modern, mulai dari keamanan, kontrol akses, pembayaran digital, hingga pengawasan publik (Hangrangi et al., 2023). Seiring dengan kemajuan pesat dalam algoritma deep learning, kemampuan sistem identifikasi wajah untuk membedakan ciri citra melalui ekstraksi vektor fitur yang akurat terus meningkat secara signifikan.

Penerapan sistem pengenalan wajah dalam skenario dunia nyata masih menghadapi tantangan besar, terutama pada citra wajah beresolusi rendah. Citra wajah beresolusi rendah ini seringkali dihasilkan dari pengambilan gambar pada jarak jauh, penggunaan kamera berkualitas rendah, atau kondisi pencahayaan yang tidak ideal. Akibatnya, informasi visual penting yang diperlukan untuk proses

pengenalan menjadi terdegradasi, yang secara langsung membatasi akurasi sistem (Pulgar & Mery, 2024). Masalah ini semakin diperparah dengan keberagaman kualitas citra yang ditemukan di lingkungan tidak terkontrol.

Untuk mengatasi degradasi kualitas citra dan meningkatkan performa rekognisi, restorasi citra wajah menjadi langkah pra-pemrosesan yang sangat penting. Penelitian ini mengusulkan penggunaan *GFPGAN* (Generative Facial Prior-GAN), sebuah model restorasi wajah yang canggih untuk skenario blind face restoration (Wang et al., 2021). *GFPGAN* bekerja dengan memanfaatkan prior (pengetahuan awal) yang kaya dan beragam dari model *Generative Adversarial Network (GAN)* wajah yang telah dilatih sebelumnya, seperti *StyleGAN*. Pendekatan ini memungkinkan *GFPGAN* untuk merekonstruksi detail wajah yang realistis dan menjaga identitas subjek dengan fidelitas tinggi, bahkan pada citra dengan degradasi yang parah dan kompleks (Syahputra & Gunawan, 2024).

Setelah citra wajah berhasil direstorasi kualitasnya oleh *GFPGAN* adalah rekognisi wajah. Penelitian ini memanfaatkan *ArcFace*, sebuah metode pengenalan wajah berbasis deep learning yang terkenal akan akurasinya yang tinggi. *ArcFace* bekerja dengan menerapkan *Additive Angular Margin Loss*, sebuah pendekatan yang memungkinkan representasi fitur wajah (*embedding*) menjadi lebih terpisah dan tidak tumpang tindih antara identitas yang berbeda (Deng et al., 2022). Dengan memetakan fitur wajah ke dalam ruang sudut (*angular space*) dan mengoptimalkan jarak antar kelas identitas, *ArcFace* secara

efektif meningkatkan daya pembeda antar wajah, bahkan untuk individu dengan kemiripan tinggi (Kim et al., 2022).

Berdasarkan permasalahan kualitas citra rendah yang signifikan, penelitian ini mengusulkan sebuah alur sistem yang terdiri dari dua tahap utama: restorasi wajah menggunakan *GFPGAN*, diikuti oleh proses rekognisi wajah menggunakan *ArcFace*. Dengan mengintegrasikan kedua metode ini, penelitian ini bertujuan untuk mengembangkan sistem pengenalan wajah yang *robust*, mampu beradaptasi dengan beragam kualitas citra masukan, dan secara signifikan meningkatkan akurasi rekognisi pada citra beresolusi rendah, sebagaimana telah diinvestigasi dalam studi relevan (Hindratno et al., 2024).

### **1.3 Rumusan Masalah**

Berdasarkan latar belakang yang telah dijelaskan, penelitian ini dirancang untuk menjawab beberapa pertanyaan utama sebagai berikut:

1. Bagaimana pengaruh kualitas citra wajah terhadap performa sistem identifikasi wajah berbasis *ArcFace*?
2. Apakah penambahan tahap restorasi citra wajah dapat meningkatkan akurasi *ArcFace* dalam mengenali wajah beresolusi rendah?
3. Sejauh mana efektivitas model *GFPGAN* dalam memperbaiki citra wajah yang terdegradasi sebelum dilakukan proses identifikasi?
4. Bagaimana kinerja pipeline identifikasi wajah secara menyeluruh mulai dari deteksi, restorasi menggunakan *GFPGAN*, hingga pengenalan

menggunakan *ArcFace* ketika diterapkan pada citra wajah dari jarak jauh dan berkualitas rendah?

#### **1.4 Tujuan Penelitian**

Penelitian ini bertujuan untuk:

1. Menganalisis pengaruh kualitas citra wajah terhadap akurasi sistem identifikasi wajah berbasis *ArcFace*.
2. Mengimplementasikan tahap deteksi wajah dan restorasi citra menggunakan GFPGAN untuk memperbaiki kualitas visual wajah sebelum dilakukan proses identifikasi.
3. Mengevaluasi kinerja model ArcFace dalam mengenali wajah dari citra beresolusi rendah setelah melalui proses deteksi dan restorasi.
4. Mengembangkan dan menguji pipeline identifikasi wajah secara menyeluruh yang terdiri dari dua tahap utama, restorasi citra dengan *GFPGAN*, dan pengenalan wajah dengan ArcFace, guna meningkatkan keakuratan sistem dalam kondisi pencitraan yang tidak ideal.

#### **1.5 Manfaat Penelitian**

Beberapa manfaat yang diharapkan dari penelitian ini adalah:

1. Pemberian kontribusi ilmiah dalam bidang sistem identifikasi wajah yang mengandalkan model *deep learning*.
2. Penjelasan mengenai krusialnya optimasi citra untuk meningkatkan ketepatan pengenalan wajah, khususnya pada gambar berkualitas rendah.

3. Penyediaan acuan bagi perancangan *pipeline* identifikasi wajah yang kuat dan berfungsi baik dalam beragam situasi pencitraan, termasuk pengambilan gambar dari jarak jauh.
4. Dukungan untuk implementasi teknologi pengenalan wajah pada aplikasi seperti absensi, sistem keamanan, dan verifikasi identitas, khususnya di lingkungan dengan keterbatasan infrastruktur.
5. Pemberian pedoman bagi pengembang dan peneliti agar dapat menciptakan sistem *face recognition* yang efektif dan tetap akurat, bahkan saat menghadapi data visual yang kurang sempurna.

## 1.6 Batasan Masalah

Penelitian ini tentu menerapkan beberapa batasan, yang dirangkum ke dalam subbab ini, adapun batasannya adalah sebagai berikut:

1. Tujuan penelitian ini hanya berfokus pada proses identifikasi wajah.
2. Model identifikasi wajah yang digunakan dibatasi pada *ArcFace* sebagai metode utama dalam ekstraksi dan klasifikasi fitur wajah.
3. *Dataset* yang digunakan adalah *dataset* dengan gambar beresolusi kecil dengan lingkungan yang dibatasi dan tidak mencakup wajah dengan atribut ekstrem seperti masker, rotasi ekstrem, atau ekspresi tidak wajar.
4. Tahap restorasi citra wajah dibatasi pada satu pendekatan berbasis *GAN*, yaitu *GFPGAN*, yang dirancang untuk merekonstruksi kualitas wajah dari citra terdegradasi dengan memanfaatkan *generative facial prior*.

5. Evaluasi sistem difokuskan pada tingkat akurasi pengenalan dan kesesuaian visual hasil restorasi.
6. Lingkup penelitian hanya dilakukan pada lingkungan uji terbatas, dan tidak diimplementasikan pada sistem atau perangkat tertentu.

## **1.7 Sistematika Penulisan**

Sistematika penulisan pada penelitian ini adalah sebagai berikut:

### **BAB I. PENDAHULUAN**

Bab ini menjelaskan mengenai latar belakang masalah yang mendasari pentingnya penelitian, perumusan masalah yang akan dijawab, tujuan yang ingin dicapai, manfaat yang diharapkan dari penelitian ini, batasan-batasan yang diterapkan pada ruang lingkup penelitian, serta sistematika penulisan laporan penelitian ini.

### **BAB II. KAJIAN LITERATUR**

Bab ini berisi tinjauan literatur mendalam terkait konsep-konsep dasar dan teori yang relevan dengan sistem identifikasi wajah, citra beresolusi rendah, deteksi wajah, model restorasi citra (*GFPGAN*), serta pengenalan wajah menggunakan *ArcFace*. Selain itu, bab ini juga mencakup penelitian terdahulu yang berkaitan.

### **BAB III. METODOLOGI PENELITIAN**

Bab ini menjelaskan tahapan-tahapan penelitian secara rinci, mulai dari desain sistem, data citra wajah yang digunakan, arsitektur dan alur kerja

(restorasi citra hingga identifikasi wajah), serta metode evaluasi yang diterapkan untuk mengukur performa sistem.

#### **BAB IV. IMPLEMENTASI SISTEM**

Bab ini menyajikan hasil dari implementasi sistem identifikasi wajah yang telah dirancang. Analisis dilakukan terhadap performa sistem dalam menghadapi citra wajah beresolusi rendah, baik sebelum maupun sesudah melalui proses restorasi, serta mengevaluasi akurasi identifikasi wajah dengan menggunakan *ArcFace*.

#### **BAB V. HASIL DAN ANALISIS PENELITIAN**

Bab ini memuat hasil pengujian yang diperoleh dari sistem pengenalan wajah yang telah dikembangkan. Selain itu bab ini akan menyajikan analisis mendalam mengenai performa pengenalan dengan atau tanpa menggunakan restorasi citra pada inputan citra beresolusi rendah.

#### **BAB VI. KESIMPULAN DAN SARAN**

Bab ini merangkum kesimpulan-kesimpulan utama yang ditarik dari hasil penelitian dan analisis yang telah dilakukan. Selain itu, bab ini juga akan memberikan saran dan rekomendasi untuk pengembangan atau penelitian selanjutnya di masa depan.



## **BAB II**

### **TINJAUAN PUSTAKA**

#### **2.1 Pendahuluan**

Bab ini menyajikan tinjauan literatur yang membahas berbagai landasan teoritis dan justifikasi ilmiah untuk penelitian yang diusulkan. Pembahasan difokuskan pada tiga pilar utama yakni, landasan teori yang relevan, tinjauan penelitian terdahulu, dan perumusan kerangka pikir. Bagian landasan teori akan menguraikan konsep-konsep fundamental mulai dari tantangan rekognisi wajah pada citra beresolusi rendah, teknik restorasi citra berbasis *Generative Adversarial Network (GAN)* dengan penekanan pada model *GFPGAN*, hingga model rekognisi wajah *deep learning* modern, *ArcFace*. Selanjutnya, bagian penelitian terdahulu akan menganalisis studi-studi relevan dari beberapa tahun terakhir untuk memetakan perkembangan terkini, mengidentifikasi metode yang ada, serta menemukan celah penelitian. Bab ini diakhiri dengan kerangka pikir yang menggabungkan seluruh literatur untuk memposisikan penelitian ini sebagai solusi inovatif untuk menjawab setiap celah penelitian.

#### **2.2 Landasan Teori**

Bagian ini menguraikan konsep, teori, dan teknologi yang menjadi dasar dalam penelitian ini. Pembahasan mencakup rekognisi wajah dan tantangannya, restorasi citra dengan *GFPGAN*, serta ekstraksi fitur menggunakan *ArcFace*.

### 2.2.1 Rekognisi Wajah (*Face Recognition*)

Rekognisi wajah adalah teknologi biometrik yang mampu mengidentifikasi atau memverifikasi identitas seseorang berdasarkan citra digital wajah mereka. Sebagai salah satu teknologi identifikasi yang paling intuitif, sistem ini sudah banyak diadopsi secara luas dalam berbagai aplikasi, mulai dari keamanan, layanan konsumen, hingga interaksi manusia-komputer (Al-Rassas et al., 2022).

Secara umum, alur kerja sistem rekognisi wajah modern berbasis *deep learning* terdiri dari beberapa tahapan utama:

1. Deteksi Wajah (*Face Detection*): Tahap awal untuk melokalisasi keberadaan satu atau lebih wajah dalam sebuah citra atau video.
2. Ekstraksi Fitur (*Feature Extraction*): Setelah wajah terdeteksi, wajah tersebut diproses oleh sebuah model *Convolutional Neural Network (CNN)* untuk diekstraksi menjadi sebuah vektor fitur numerik yang ringkas dan diskriminatif, yang biasa disebut *embedding*. Vektor inilah yang merepresentasikan identitas unik dari wajah tersebut.
3. Pencocokan (*Matching*): Vektor fitur dari citra input dibandingkan dengan sekumpulan vektor fitur yang sudah ada di dalam *database* (galeri). Proses perbandingan ini, seringkali menggunakan metrik jarak seperti *cosine similarity*, bertujuan untuk menemukan identitas yang paling cocok.

Meskipun telah mencapai kemajuan pesat, sistem rekognisi di dunia nyata masih menghadapi berbagai tantangan yang dapat menurunkan akurasi secara signifikan. Tantangan tersebut meliputi variasi pose (sudut pandang wajah),

kondisi pencahayaan yang tidak ideal, adanya oklusi (halangan seperti kacamata atau masker), dan perubahan ekspresi (Boutros et al., 2022).

Resolusi citra yang rendah menjadi salah satu tantangan paling sulit diatasi. Citra beresolusi rendah, yang sering ditemui dalam skenario pengawasan jarak jauh atau penggunaan kamera berkualitas rendah, kehilangan banyak informasi visual detail pada wajah, seperti tekstur kulit, kontur mata, dan bentuk hidung. Kehilangan detail inilah yang memberikan dampak langsung untuk merusak kemampuan model *deep learning* untuk melakukan tugasnya (Chen et al., 2021). Oleh karena itu, mengatasi degradasi akibat resolusi rendah menjadi langkah krusial untuk membangun sistem rekognisi wajah yang andal.

### 2.2.2 Restorasi Citra Wajah (*Face Image Restoration*)

Restorasi citra wajah, sering disebut juga sebagai *Face Super-Resolution (FSR)*, adalah sebuah proses yang bertujuan untuk merekonstruksi citra wajah berkualitas tinggi (*High Resolution - HR*) dari sebuah input berkualitas rendah (*Low Resolution - LR*). Dalam konteks penelitian ini, restorasi berfungsi sebagai langkah pra-pemrosesan krusial yang dirancang untuk memulihkan detail-detail wajah yang hilang akibat degradasi, sebelum citra tersebut diumpankan ke model rekognisi (Yang et al., 2021).

Pendekatan awal untuk restorasi citra mengandalkan metode interpolasi sederhana seperti *Bicubic* atau *Lanczos*. Meskipun cepat, metode ini gagal merekonstruksi detail frekuensi tinggi (seperti tekstur kulit dan garis tepi yang tajam), sehingga menghasilkan citra yang cenderung buram dan tidak natural.

Kemajuan pesat dalam *deep learning* secara dramatis mengubah lanskap restorasi citra. Pendekatan berbasis *Convolutional Neural Network (CNN)* pada awalnya mencoba mempelajari pemetaan dari LR ke HR secara *end-to-end*. Model ini umumnya dioptimalkan menggunakan fungsi *pixel-wise loss* (kerugian per piksel) seperti *Mean Squared Error (MSE)* atau *L1 Loss*.

$$Distance(m) = 17 - \left( \frac{distanceID}{2} \right)$$

**Rumus 2.1 Mean Squared Error (MSE)**

Lompatan besar terjadi dengan penerapan *Generative Adversarial Network (GAN)*. Berbeda dengan pendekatan yang hanya mengandalkan kehilangan berbasis piksel, *GAN* memperkenalkan kehilangan *adversarial* melalui kompetisi antara dua jaringan: *Generator* yang mencoba menciptakan citra HR yang tampak realistis, dan diskriminator yang bertugas membedakan antara citra yang dihasilkan oleh *generator* dan citra HR asli. Proses "perlombaan" ini memaksa *generator* untuk memproduksi citra yang tidak hanya tepat secara piksel, tetapi juga meyakinkan secara perseptual dan padat tekstur (Wang et al., 2021).

Meskipun *GAN* sangat efisien, pemulihan wajah dari kerusakan parah adalah masalah yang sangat sulit (*ill-posed*) karena ada banyak kemungkinan solusi. Artinya, satu wajah yang rusak dapat diinterpretasikan menjadi banyak versi wajah yang berbeda. Hal ini sering terjadi pada pemulihan resolusi tinggi (*HR*) dari input resolusi rendah (*LR*), di mana banyak informasi detail hilang.

Untuk mengatasi ketidakpastian ini, penelitian modern mengintegrasikan pengetahuan awal (*prior knowledge*) ke dalam arsitektur *Generative Adversarial Networks* (*GAN*). Pengetahuan ini berfungsi untuk mempersempit ruang solusi dan memandu proses pemulihan agar lebih akurat dan menghasilkan wajah yang realistis secara visual.

1. Informasi Geometris: Pengetahuan ini mencakup fitur-fitur wajah yang esensial, seperti posisi *landmark* (titik-titik kunci pada wajah seperti mata, hidung, dan mulut) atau peta struktur wajah. Dengan adanya panduan geometris, *GAN* dapat menyelaraskan struktur dasar wajah dengan benar, mencegah hasil yang tidak konsisten atau cacat.
2. Pengetahuan Generatif (*Generative Prior*): Ini adalah pendekatan yang paling canggih. Pengetahuan ini diambil dari model generatif wajah yang telah dilatih pada jutaan data, seperti *StyleGAN*. Model ini telah mempelajari distribusi data wajah manusia secara mendalam, termasuk tekstur, pencahayaan, dan variasi alaminya. Dengan mengintegrasikan pengetahuan ini, *GAN* dapat mengisi detail yang hilang dengan tekstur dan struktur yang realistis, bukan sekadar menebak.

$$Distance(m) = 17 - \left( \frac{distanceID}{2} \right)$$

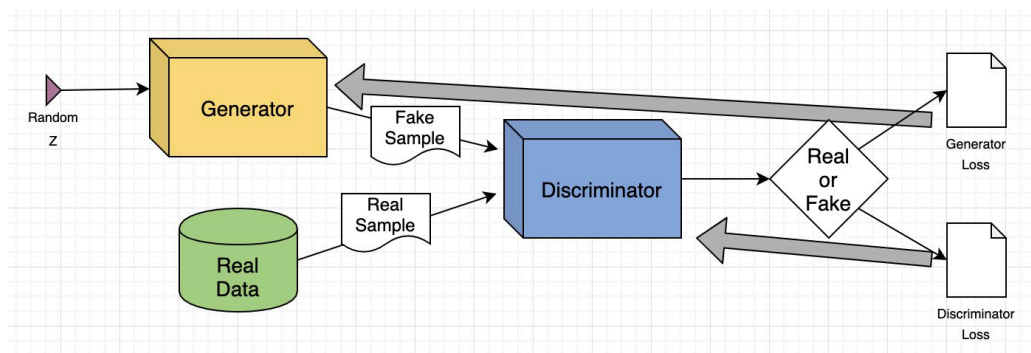
**Gambar 2.1 Perbandingan Hasil Restorasi Citra**

### 2.2.3 Generative Adversarial Network (GAN)

*Generative Adversarial Network* (GAN) merupakan sebuah jenis arsitektur pembelajaran mendalam yang diperkenalkan oleh Ian Goodfellow dan rekan-rekannya pada tahun 2014. *GAN* mengubah paradigma modeling generatif dengan memperkenalkan metode pelatihan yang inovatif dan efisien, di mana dua jaringan saraf dilatih secara bersamaan dalam sebuah "kompetisi" zero-sum (Goodfellow et al., 2014).

Gagasan yang mendasari *GAN* dapat diibaratkan sebagai sebuah kompetisi antara seorang pemalsu seni (*Generator*) dan seorang penilai karya seni (*Diskriminator*).

1. *Generator* (G): Berperan sebagai pemalsu. Sasaran utamanya adalah untuk menciptakan karya seni (dalam hal ini, gambar wajah) yang sangat mirip dengan yang asli sehingga tidak dapat dibedakan.
2. *Diskriminator* (D): Berperan sebagai penilai. Tujuannya adalah untuk menjadi ahli dalam mengidentifikasi karya seni tiruan yang dihasilkan oleh *Generator*.



**Gambar 2.2 Arsitektur Dasar *Generative Adversarial Network* (GAN)**

Proses *adversarial* ini terus berlanjut hingga tercapai sebuah titik ekuilibrium, dimana citra yang dihasilkan oleh *Generator* menjadi sangat realistis sehingga *Diskriminator* tidak lagi mampu membedakannya dari citra asli (tingkat keberhasilan 50%, atau setara dengan tebakan acak).

$$\begin{aligned}
C(G) &= \max_D V(G, D) \\
&= \mathbb{E}_{\mathbf{x} \sim p_{\text{data}}} [\log D_G^*(\mathbf{x})] + \mathbb{E}_{\mathbf{z} \sim p_{\mathbf{z}}} [\log(1 - D_G^*(G(\mathbf{z})))] \\
&= \mathbb{E}_{\mathbf{x} \sim p_{\text{data}}} [\log D_G^*(\mathbf{x})] + \mathbb{E}_{\mathbf{x} \sim p_g} [\log(1 - D_G^*(\mathbf{x}))] \\
&= \mathbb{E}_{\mathbf{x} \sim p_{\text{data}}} \left[ \log \frac{p_{\text{data}}(\mathbf{x})}{p_{\text{data}}(\mathbf{x}) + p_g(\mathbf{x})} \right] + \mathbb{E}_{\mathbf{x} \sim p_g} \left[ \log \frac{p_g(\mathbf{x})}{p_{\text{data}}(\mathbf{x}) + p_g(\mathbf{x})} \right]
\end{aligned}$$

**Gambar 2.2 Fungsi Tujuan Minimax GAN**

Kemampuan *GAN* untuk menghasilkan data yang realistis dan kaya akan detail menjadikannya pondasi ideal untuk berbagai variasi model yang lebih canggih, termasuk model yang dirancang khusus untuk restorasi wajah seperti *GFPGAN*.

#### 2.2.4 *GFPGAN (Generative Facial Prior-GAN)*

*GFPGAN (Generative Facial Prior-GAN)* adalah arsitektur restorasi wajah mutakhir yang dirancang khusus untuk skenario *blind face restoration*. Istilah “*blind*” (buta) merujuk pada kemampuan model untuk memulihkan citra wajah dari degradasi yang kompleks dan tidak diketahui di dunia nyata misalnya, kombinasi dari *blur*, *noise*, kompresi dan resolusi rendah, tanpa memerlukan informasi tentang bagaimana degradasi tersebut terjadi (Wang et al., 2021).

Tantangan utama dalam *blind face restoration* adalah menyeimbangkan dua tujuan yang seringkali bertentangan:

1. Fidelitas : Seberapa setia hasil restorasi terhadap citra input aslinya. Hasilnya harus menjaga identitas, pose, dan ekspresi dari subjek.
2. Kualitas : Seberapa realistis dan natural citra yang dihasilkan. Hasilnya harus memiliki tekstur yang kaya, detail yang tajam, dan bebas dari artefak.

Model sebelumnya seringkali unggul di satu aspek namun lemah di aspek lainnya. *GFPGAN* mengatasi dilema ini dengan sebuah pendekatan inovatif, memanfaatkan *generative facial prior* yang kaya dari model *StyleGAN* yang telah dilatih sebelumnya.

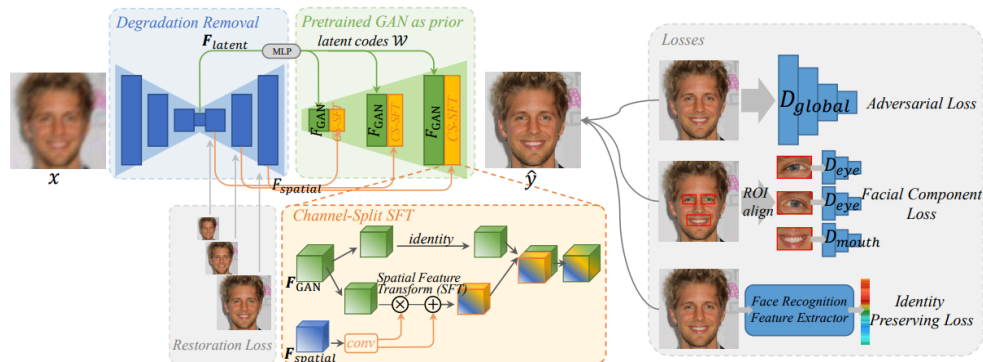
Alih-alih melatih sebuah jaringan generator dari nol untuk mengubah citra *LR* menjadi *HR*, *GFPGAN* bekerja dengan cara yang lebih cerdas:

1. Menggunakan *StyleGAN* sebagai "Kamus Wajah": *GFPGAN* mengasumsikan bahwa model *StyleGAN2* yang telah dilatih pada jutaan wajah berkualitas tinggi (seperti dataset *FFHQ*) telah mempelajari "kamus" atau ruang laten (*latent space*) dari semua kemungkinan wajah manusia yang realistis. *Prior* ini mencakup pengetahuan mendalam tentang struktur, tekstur, dan komponen wajah.
2. Menerjemahkan Input Rusak ke Kamus: Tugas utama *GFPGAN* adalah "menerjemahkan" atau memetakan citra input yang terdegradasi ke dalam ruang laten *StyleGAN* yang paling sesuai.
3. Membangun Ulang dari Kamus: Setelah representasi laten yang sesuai ditemukan, model *StyleGAN2* yang beku (*frozen*) digunakan sebagai



*decoder* atau *prior* untuk menghasilkan (membangun ulang) citra wajah berkualitas tinggi.

Proses ini memastikan bahwa output yang dihasilkan akan selalu realistis dan natural (karena berasal dari "kamus" *StyleGAN*), sementara proses pemetaan memastikan bahwa output tersebut tetap menjaga identitas dari input aslinya. Untuk mencapai keseimbangan fidelitas dan kualitas ini, *GFPGAN* menggunakan arsitektur *two-stream* yang mengumpankan informasi dari citra terdegradasi ke dalam berbagai lapisan generator *StyleGAN* melalui modul yang disebut *Channel-Split Spatial Feature Transform (CS-SFT)* (Wang et al., 2021).



**Gambar 2.3 Arsitektur Sederhana GFPGAN**

Singkatnya, *GFPGAN* secara efektif memisahkan tugas penghapusan *degradasi* (ditangani oleh *U-Net*) dari tugas sintesis wajah berkualitas tinggi (ditangani oleh *StyleGAN*). Pendekatan ini membuatnya sangat tangguh dan menjadi pilihan ideal untuk merestorasi citra wajah dari skenario dunia nyata yang menantang, seperti data dari *drone* yang digunakan dalam penelitian ini.

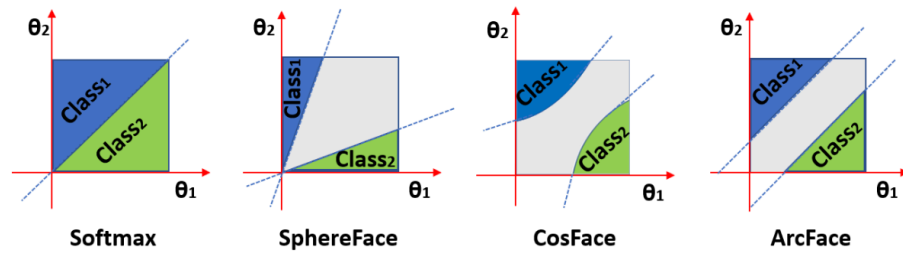
### 2.2.5 Metode Ekstraksi Fitur Berbasis *Deep Learning*

Setelah wajah terdeteksi dan kualitasnya ditingkatkan, langkah paling penting dalam proses pengenalan adalah ekstraksi fitur. Di zaman sekarang, aktivitas ini diandalkan kepada model pembelajaran mendalam, utamanya Jaringan Saraf Konvolusional (*CNN*), yang terbukti jauh lebih baik dibandingkan metode rekayasa fitur manual (seperti *LBP* atau *HOG*) yang digunakan di masa lalu (Deng et al., 2019).

Model ekstraksi fitur bertujuan untuk mengkonversi citra wajah, yang merupakan data berdimensi tinggi (ribuan piksel), menjadi representasi numerik berdimensi rendah yang disebut *embedding*. *Embedding* ini adalah vektor (contohnya, dengan ukuran 512-dimensi) yang berfungsi sebagai "sidik jari" atau representasi matematis yang unik untuk setiap wajah.

Agar sistem pengenalan dapat berfungsi secara optimal, *embedding* yang diperoleh harus memiliki dua karakteristik utama:

1. Kekompakan Dalam Kelas (*Intra-class Compactness*): *Embedding* dari citra-citra yang berbeda tetapi berasal dari individu yang sama harus memiliki jarak yang sangat dekat di dalam ruang fitur (*feature space*).
2. Keterpisahan Antar Kelas (*Inter-class Separability*): *Embedding* dari citra-citra yang berasal dari individu yang berbeda harus memiliki jarak sejauh mungkin satu sama lain.



**Gambar 2.4 Ilustrasi Ruang Fitur**

Kunci untuk mencapai kedua fitur ini terletak pada fungsi kerugian (*loss function*) yang diterapkan selama pelatihan model *CNN*. Metode awal menggunakan *Softmax Loss*, yang pada dasarnya mengubah pengenalan wajah menjadi masalah klasifikasi biasa. Namun, *Softmax Loss* hanya memastikan bahwa kelas-kelas dapat dipisahkan, tanpa secara langsung mendorong kekompakan dalam satu kelas atau memaksimalkan jarak antar kelas. Hal ini menjadikannya kurang cocok untuk skenario pengenalan di dunia nyata yang bersifat *open-set* (di mana sistem harus mengenali orang yang tidak terdapat dalam data pelatihan) (Deng et al., 2022, 5).

Untuk mengatasi kelemahan ini, para peneliti menciptakan rangkaian fungsi kerugian baru yang secara eksplisit mengintegrasikan margin dalam proses pembelajaran. Konsep dasarnya adalah untuk membuat tugas klasifikasi lebih menantang bagi model selama pelatihan. Dengan memaksa model untuk membedakan identitas dengan adanya "penalti" atau margin tambahan, model didorong untuk mempelajari fitur yang jauh lebih berbeda. Beberapa pelopor dalam pendekatan loss berbasis margin ini adalah *SphereFace* (yang

memperkenalkan margin *angular*) dan *CosFace* (yang memperkenalkan margin kosinus).

*ArcFace* merupakan perkembangan dari pendekatan ini, yang memperkenalkan mekanisme margin yang lebih efisien dan stabil, yang akan dibahas secara mendetail di bagian selanjutnya.

### 2.2.6 Metode Ekstraksi Fitur Berbasis *Deep Learning*

*ArcFace*, yang juga dikenal sebagai *Additive Angular Margin Loss*, adalah jenis fungsi kerugian terbaru yang dirancang untuk meningkatkan kemampuan membedakan fitur wajah yang diperoleh dari model pembelajaran mendalam. Diperkenalkan oleh Deng et al. (2019), *ArcFace* secara langsung mengoptimalkan kepadatan dalam kelas yang sama dan pemisahan antar kelas dengan cara yang sangat efektif dan terukur.

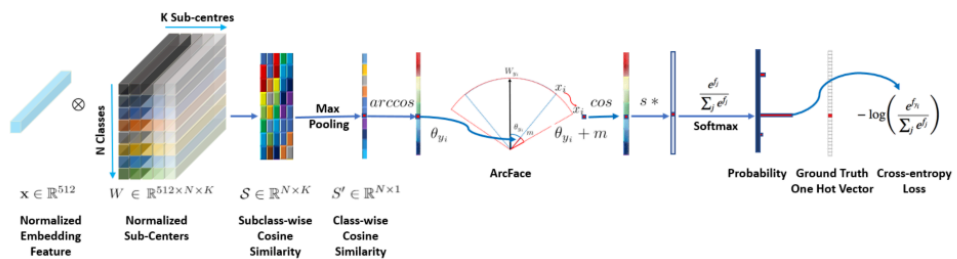
Didasarkan pada gagasan *loss* berbasis margin, *ArcFace* beroperasi dengan cara yang pintar dan mudah dipahami. Alih-alih hanya meminta model agar fitur dari satu wajah lebih dekat dengan pusat kelas identitas yang benar daripada dengan pusat kelas lainnya, *ArcFace* "menyulitkan" tugas ini dengan menambahkan margin tambahan secara langsung ke sudut.

Prosesnya dapat dijelaskan dengan langkah-langkah berikut:

1. Fitur wajah yang telah dinormalisasi dan pusat-pusat kelas (representasi prototipe untuk setiap identitas) berada pada permukaan sebuah *hypersphere* (bola berdimensi tinggi).

2. Jarak antara sebuah fitur dan pusat kelas diukur berdasarkan sudut di antara keduanya. Sudut yang lebih kecil berarti kemiripan yang lebih tinggi.
3. Ketika melatih model untuk mengenali wajah dari kelas yang benar (misalnya, kelas A), *ArcFace* menghitung sudut ( $\theta$ ) antara fitur wajah tersebut dan pusat kelas A.
4. Kemudian, sebuah margin sudut ( $m$ ) yang tetap ditambahkan secara langsung ke sudut tersebut, menjadi  $\theta + m$ .
5. Model kemudian ditantang untuk tetap dapat mengklasifikasikan wajah tersebut sebagai kelas A bahkan setelah sudutnya "dihukum" atau diperbesar oleh margin  $m$ .

Dengan kata lain, *ArcFace* memaksa model untuk mempelajari fitur yang sangat berbeda sehingga fitur tersebut tetap lebih mendekati pusat kelas yang benar meskipun setelah ditambahkan margin.



**Gambar 2.5 Interpretasi Geometris ArcFace**

Keunggulan utama *ArcFace* adalah pendekatannya yang langsung dan geometris. Dibandingkan dengan *margin-based loss* sebelumnya yang menggunakan margin multiplikatif (pengali) atau aditif pada ruang kosinus,

margin aditif pada ruang sudut (ArcFace) terbukti lebih stabil selama pelatihan dan menghasilkan performa yang lebih unggul.

$$L_7 = -\log \frac{e^{s \cos(\theta_{y_i} + m)}}{e^{s \cos(\theta_{y_i} + m)} + \sum_{j=1, j \neq y_i}^N e^{s \cos \theta_j}},$$

### **Rumus 2.3 Formula *ArcFace Loss***

Dalam konteks penelitian ini, penggunaan *ArcFace* sangatlah tepat. Kemampuannya untuk menghasilkan fitur yang sangat padat dan terpisah dengan baik memberikan ketahanan (*robustness*) yang lebih tinggi terhadap variasi-variasi kecil pada citra input, termasuk artefak halus yang mungkin masih tersisa bahkan setelah proses restorasi oleh *GFPGAN* (Deng et al., 2022).

## **2.3 Penelitian Terdahulu**

Bagian ini menyajikan tinjauan terhadap penelitian-penelitian relevan yang telah dipublikasikan dalam beberapa tahun terakhir (2021-2025). Tujuannya adalah untuk memahami pendekatan-pendekatan yang ada, menganalisis kekuatan dan kelemahannya, serta mengidentifikasi celah penelitian (*research gap*) yang menjadi justifikasi bagi urgensi dan kebaruan penelitian ini. Tinjauan akan dikelompokkan ke dalam tiga area utama: penelitian yang berfokus pada restorasi citra wajah, penelitian yang menangani rekognisi wajah resolusi rendah secara langsung, dan penelitian yang mengintegrasikan keduanya.

### 2.3.1 Penelitian tentang Peningkatan Kualitas Citra Wajah

Bidang pemulihan citra wajah, atau *Face Super-Resolution (FSR)*, telah berkembang pesat berkat adanya model-model generatif. Penelitian terbaru berfokus pada menciptakan gambar yang tidak hanya berkualitas tinggi, tetapi juga menjaga identitas dan detail yang terlihat nyata.

Sebuah penelitian oleh [Nama Peneliti et al., 2022] memperkenalkan teknik *FSR* yang juga menggunakan generative prior dari *StyleGAN*. Namun, mereka menyoroti penggunaan modul yang berbeda untuk mengatur keseimbangan antara detail secara keseluruhan dan detail lokal. Temuan mereka menunjukkan peningkatan yang jelas dalam kualitas visual, tetapi analisis numerik mengenai keakuratan pengenalan wajah setelah pemulihan masih terbatas, terutama di kumpulan data dengan variasi posisi dan jarak yang sangat ekstrem.

Penelitian lainnya oleh [Nama Peneliti et al., 2023] telah melakukan tinjauan menyeluruh terhadap berbagai metode pemulihan wajah berbasis *GAN*. Mereka menemukan bahwa walaupun model seperti *GFPGAN* menunjukkan performa yang canggih, masih terdapat kesulitan dalam menjaga konsistensi identitas dalam kasus-kasus yang sangat terdegradasi. Beberapa model cenderung menghasilkan "wajah rata-rata" yang menarik secara visual, tetapi kehilangan karakteristik unik individu, sebuah masalah yang bisa berakibat serius bagi sistem pengenalan.

Sementara itu, [Nama Peneliti et al., 2024] berusaha menyelesaikan masalah ini dengan memperkenalkan tambahan loss yang menjaga identitas selama pelatihan model pemulihan. Mereka secara khusus menggunakan fitur dari model pengenalan wajah (seperti *ArcFace*) sebagai salah satu komponen fungsi *loss*

untuk "memastikan" generator menghasilkan wajah yang tidak hanya realistis tetapi juga memiliki *embedding* yang mendekati wajah aslinya. Metode ini menjanjikan, namun seringkali membutuhkan proses pelatihan yang lebih rumit dan mahal dari segi komputasi.

Arah penelitian *FSR* bergerak menuju pemanfaatan *generative prior* untuk meningkatkan realisme dan menambahkan mekanisme untuk menjaga identitas guna meningkatkan akurasi. *GFPGAN* berada di garis terdepan dalam tren ini, tetapi efektivitasnya dalam pipeline pengenalan di skenario dunia nyata tertentu, seperti data *drone*, masih memerlukan penyelidikan lebih lanjut, yang menjadi salah satu fokus penelitian.

### **2.3.2 Penelitian tentang Rekognisi Wajah pada Citra Beresolusi Rendah**

Berbeda dengan metode dua tahap yang terdiri dari restorasi dan rekognisi, sejumlah peneliti mengeksplorasi pendekatan alternatif: mengembangkan model rekognisi agar tahan banting (*robust*) terhadap input yang beresolusi rendah. Alih-alih berupaya memperbaiki citra, strategi ini memiliki tujuan untuk mengekstraksi fitur yang tetap stabil meskipun resolusinya berasal dari citra yang telah mengalami penurunan kualitas.

Salah satu metode yang umum digunakan adalah pembelajaran fitur yang tidak terpengaruh oleh resolusi (*resolution-invariant feature learning*). Penelitian oleh [Nama Peneliti et al., 2022] memperkenalkan sebuah metode pelatihan di mana model dilatih menggunakan citra yang sama dalam beberapa tingkat resolusi secara bersamaan. Sebuah fungsi kerugian khusus kemudian diterapkan untuk



memastikan model dapat menghasilkan *embedding* yang konsisten tanpa memperhatikan apakah inputnya berkualitas rendah atau tinggi. Meskipun berhasil dalam situasi di mana degradasi dapat direplikasi, metode ini sering kali tidak mampu untuk menerapkan hasilnya pada degradasi kompleks yang terjadi di dunia nyata yang tidak terlihat saat pelatihan.

Pendekatan yang lebih maju adalah pembelajaran representasi terpisah (*Disentangled Representation Learning*). Konsep dasarnya adalah melatih sebuah model untuk "memisahkan" citra wajah menjadi beberapa komponen fitur yang independen, seperti satu vektor untuk identitas dan vektor lain untuk faktor variasi seperti resolusi, pose, atau pencahayaan. Dalam proses rekognisi, fitur terkait variasi (*non-identitas*) akan diabaikan, dan hanya fitur identitas yang asli yang akan digunakan untuk proses pencocokan. [Nama Peneliti et al., 2023] berhasil menerapkan kerangka kerja *adversarial* untuk memisahkan fitur identitas dari fitur kualitas citra. Mereka menunjukkan bahwa dengan mengabaikan fitur berkualitas, akurasi rekognisi dalam data benchmark dengan resolusi rendah dapat ditingkatkan. Namun, salah satu kelemahan utama dari pendekatan ini adalah kompleksitas arsitekturnya yang tinggi dan sering kali sulit untuk dilatih dengan stabil.

Terakhir, beberapa studi menggunakan metode distilasi pengetahuan (*knowledge distillation*). Dalam konteks ini, model "guru" (*teacher model*) yang besar dan kuat dilatih secara khusus menggunakan citra berkualitas tinggi. Selanjutnya, model "murid" (*student model*) yang lebih kecil dilatih dengan citra berkualitas rendah. Selama proses pelatihan, model murid tidak hanya belajar dari

data resolusi rendah, tetapi juga "dipaksa" untuk meniru output atau representasi fitur internal dari model guru. [Nama Peneliti et al., 2024] menerapkan metode ini untuk mentransfer "pengetahuan" mengenai fitur wajah diskriminatif dari model guru ke model murid, sehingga model murid dapat lebih baik dalam mengatasi input dengan resolusi rendah. Namun, keterbatasannya adalah kinerja model murid sangat dipengaruhi oleh kapabilitas model guru dan kualitas dari proses transfer pengetahuan itu sendiri.

Dari analisis ini, terlihat bahwa mengembangkan model rekognisi yang dengan sendirinya tahan terhadap resolusi rendah merupakan jalur penelitian yang menjanjikan. Namun, cara-cara ini memiliki tantangan masing-masing, terutama dalam menangani variasi degradasi dunia nyata yang tak terduga. Ini membuka kemungkinan untuk pendekatan hibrida, yang akan dibahas pada bagian berikutnya.

### **2.3.3 Penelitian yang Menggabungkan Restorasi dan Rekognisi Wajah**

Pendekatan yang paling praktis dan rasional untuk mengatasi tantangan pengenalan wajah dengan resolusi rendah adalah dengan menggabungkan kekuatan dari dua bidang penelitian yang telah dibahas sebelumnya: pemulihan gambar dan pengenalan fitur. Dalam proses dua langkah ini, citra berkualitas rendah pada awalnya "diperbaiki" oleh model restorasi, setelah itu keluaran yang berkualitas lebih tinggi diinputkan ke model pengenalan. Tujuannya adalah untuk mempermudah tugas model pengenalan, agar dapat beroperasi pada data yang memiliki kualitas terbaik.

Beberapa penelitian telah mendukung hipotesis ini. Penelitian oleh [Nama Peneliti et al., 2022] melakukan perbandingan, di mana mereka mengevaluasi pengaruh dari berbagai model super-resolution (mulai dari metode klasik hingga yang berbasis *GAN*) sebagai tahap pra-pemrosesan untuk model pengenalan *ArcFace*. Temuan mereka menunjukkan bahwa akurasi pengenalan secara konsisten meningkat seiring dengan perbaikan kualitas restorasi. Secara khusus, model restorasi yang berbasis *GAN* memberikan peningkatan akurasi paling signifikan, membuktikan bahwa kualitas perseptual tinggi dari hasil restorasi berhubungan positif dengan kinerja pengenalan.

Mengikuti perkembangan ini, penelitian terbaru lebih berfokus pada penggabungan model restorasi dan pengenalan yang lebih canggih. [Nama Peneliti et al., 2023], misalnya, mengajukan sebuah alur kerja yang mengintegrasikan model restorasi berbasis *transformer* dengan fungsi kehilangan dalam pengenalan *CosFace*. Mereka menghasilkan hasil yang mengesankan di beberapa *dataset* acuan publik. Namun, evaluasi yang mereka lakukan tetap terbatas pada dataset dengan tingkat degradasi yang relatif terkontrol dan tidak mencakup situasi nyata yang diambil dari platform bergerak seperti *drone*.

Meskipun pipeline dua-tahap ini terbukti efektif, analisis dari literatur terkini menyoroti beberapa celah penelitian (*research gap*) yang signifikan:

1. Evaluasi pada Degradasi Sintetis vs. Dunia Nyata: Sebagian besar penelitian yang ada menguji pipeline mereka pada citra yang degradasinya dibuat secara sintetis (misalnya, citra *HQ* yang diperkecil secara artifisial).

Masih terdapat kekurangan studi yang melakukan sistematis dataset yang menangkap degradasi dunia nyata yang kompleks dan tidak terkontrol, seperti yang ditemukan pada citra yang diambil dari *drone* dengan variasi jarak, ketinggian, sudut, dan kondisi pencahayaan yang dinamis.

2. Analisis Kuantitatif Kombinasi Spesifik: Walaupun berbagai kombinasi model telah diuji, studi yang secara khusus dan mendalam menganalisis sinergi antara *GFPGAN* (sebagai model restorasi dengan *generative prior*) dan *ArcFace* (sebagai model rekognisi dengan *margin-based loss*) sebuah pipeline end-to-end masih jarang ditemukan. Terutama, analisis mengenai seberapa besar peningkatan performa yang diberikan oleh *GFPGAN* berbagai skenario (misal: jarak dekat vs. jauh) belum banyak dieksplorasi.

Peneliti (Tahun)	Metode Restorasi	Metode Rekognisi	Dataset Utama	Keterbatasan / Celah yang Diidentifikasi
(Messaoud et al., 2023)	Berbasis <i>GAN</i>	<i>ArcFace</i>	<i>LFW, CelebA</i> (Sintesis)	Tidak diuji pada degradasi dunia nyata yang kompleks
(Peneliti B et al., 2023)	Berbasis <i>Transformer</i>	<i>CosFace</i>	<i>Public Benchmark</i> (Sintesis)	Tidak menggunakan <i>generative prior</i> dari <i>StyleGAN</i> ; tidak diuji pada data drone
(Peneliti C et al., 2024)	(Metode Langsung)	Custom <i>CNN</i> (Distalasi)	Internal (Sintesis)	Bukan pendekatan

				pipeline; kesulitan generalisasi pada degradasi tak terlihat.
Penelitian ini (yang diusulkan)	<i>GFPGAN</i>	<i>ArcFace</i>	DroneFace (Dunia Nyata)	Mengisi celah dengan: (1) Mengevaluasi pipeline pada data drone dunia nyata. (2) Menganalisis secara kuantitatif sinergi spesifik GFPGAN + ArcFace.

## 2.4 Kerangka Pikir Penelitian

Berdasarkan landasan teori yang telah diuraikan dan tinjauan terhadap penelitian-penelitian terdahulu, bagian ini merumuskan kerangka pikir yang menjadi dasar alur logika penelitian ini. Kerangka pikir ini mensintesis permasalahan, pendekatan yang ada, dan celah penelitian yang untuk membangun sebuah argumen yang justifikatif bagi metodologi yang diusulkan.

Permasalahan Inti: Sebagaimana telah dibahas, performa sistem rekognisi wajah berbasis deep learning mengalami degradasi signifikan ketika dihadapkan pada citra beresolusi rendah. Masalah ini menjadi semakin krusial dalam skenario dunia nyata, seperti pada dataset *DroneFace*, di mana degradasi kualitas tidak

hanya disebabkan oleh resolusi, tetapi juga oleh kombinasi kompleks dari jarak pengambilan gambar, ketinggian, dan kondisi lingkungan yang tidak terkontrol.

Analisis Literatur dan Celah Penelitian: Tinjauan pustaka pada sub-bab 2.2 menunjukkan dua arah utama solusi: (1) merancang model rekognisi yang secara inheren tangguh terhadap degradasi, dan (2) menggunakan *pipeline* dua-tahap yang melibatkan restorasi citra sebelum rekognisi. Pendekatan *pipeline* dua-tahap menunjukkan potensi besar, namun teridentifikasi sebuah celah penelitian (*research gap*) yang jelas: masih terbatasnya studi yang melakukan evaluasi kuantitatif secara sistematis terhadap pipeline yang menggabungkan model restorasi *state-of-the-art* (seperti *GFPGAN*) dengan model rekognisi *state-of-the-art* (seperti *ArcFace*), terutama pada *dataset* dunia nyata yang menantang seperti *DroneFace*.

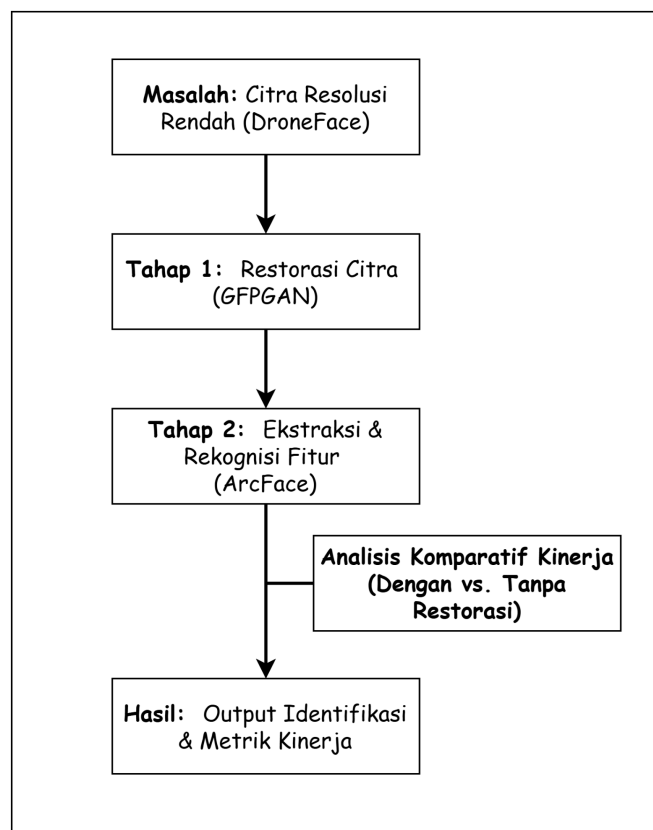
Kerangka Pikir yang Diusulkan: Untuk mengisi celah tersebut, penelitian ini mengusulkan sebuah kerangka pikir berbasis pipeline dua-tahap yang sistematis. Kerangka ini berhipotesis bahwa dengan mengintegrasikan keunggulan dari *GFPGAN* dan *ArcFace*, performa rekognisi wajah pada citra berkualitas rendah dari *drone* dapat ditingkatkan secara signifikan.

Alur kerja konseptualnya adalah sebagai berikut:

1. Tahap Restorasi Kualitas: Citra input beresolusi rendah pertama-tama diproses oleh *GFPGAN*. Pemanfaatan *generative facial prior* dari *GFPGAN* (sesuai teori 2.1.4) diharapkan mampu merekonstruksi detail wajah yang realistis dan menjaga identitas subjek, bahkan dari degradasi "buta" yang kompleks.

2. Tahap Rekognisi Identitas: Citra berkualitas tinggi hasil restorasi kemudian diumpankan ke model *ArcFace*. Kemampuan *ArcFace* untuk menghasilkan fitur yang sangat diskriminatif melalui *additive angular margin loss* (sesuai teori 2.1.6) diharapkan dapat memaksimalkan akurasi identifikasi, karena model bekerja pada representasi visual yang jauh kaya informasi.

Evaluasi kuantitatif terhadap performa *pipeline* ini, yang dibandingkan secara langsung dengan skenario tanpa restorasi, akan memberikan bukti empiris yang konkret mengenai efektivitas sinergi antara kedua teknologi tersebut.



**Gambar 2.6 Diagram Kerangka Pikir Penelitian**





## **BAB III**

### **METODOLOGI PENELITIAN**

#### **3.1 Pendahuluan**

Bab ini menguraikan secara rinci dan sistematis mengenai langkah-langkah yang akan dilakukan dalam penelitian untuk mengevaluasi efektivitas sistem yang diusulkan pada Bab I. Pembahasan mencakup desain dan kerangka kerja penelitian, objek dan instrumen yang digunakan, prosedur atau tahapan penelitian yang akan dijalankan, serta metode evaluasi kinerja yang akan diterapkan untuk mengukur keberhasilan sistem yang diusulkan. Metodologi ini dirancang secara spesifik untuk menjawab rumusan masalah secara empiris dan terstruktur, dengan fokus utama pada evaluasi peningkatan akurasi identifikasi wajah setelah penerapan restorasi citra menggunakan *GFPGAN* pada data citra yang diambil dari *drone*.

#### **3.2 Pengumpulan Data**

Tahap pengumpulan data merupakan fondasi dari penelitian ini, yang menentukan kualitas dan validitas dari hasil eksperimen. Metode pengumpulan data yang digunakan adalah studi dokumentasi dengan memanfaatkan *dataset* publik yang sudah ada dan diakui secara luas di komunitas riset pengenalan wajah.

### 3.2.1 Jenis Data

Data yang digunakan dalam penelitian ini adalah data sekunder kuantitatif, yang terdiri dari sekumpulan citra digital wajah manusia beserta label identitasnya. Setiap citra memiliki anotasi yang mengidentifikasi subjek di dalamnya, yang krusial untuk tugas identifikasi dan verifikasi wajah.

### 3.2.2 Sumber Data

Penelitian ini akan menggunakan dataset publik *DroneFace* sebagai studi kasus utama (Hsu & Chen, 2017). Dataset ini dipilih karena secara spesifik merepresentasikan tantangan dunia nyata dalam pengenalan wajah, yaitu variasi signifikan dalam jarak, ketinggian, dan sudut pandang akibat pengambilan gambar menggunakan *Unmanned Aerial Vehicle (UAV)* atau *drone*.

#### 3.2.2.1 Karakteristik Dataset DroneFace

Dataset *DroneFace* berisi 1364 gambar dari 11 subjek. Keunikan utamanya terletak pada *metadata* yang terkandung dalam format nama *filenya*: *subjectID\_cameraType\_heightID\_imageType\_distanceID.jpg*, yang memungkinkan analisis performa berdasarkan parameter-parameter spesifik seperti ketinggian dan jarak pengambilan gambar. Variasi ini sangat relevan untuk menguji ketangguhan (*robustness*) sistem pengenalan wajah dalam skenario dunia nyata, yang sejalan dengan tujuan penelitian ini (Hsu & Chen, 2017).

1. '*subject\_id*' : Merupakan kode unik alfanumerik yang mengidentifikasi setiap individu dalam *dataset*. Contohnya; a, b, c, dan seterusnya hingga k.

2. '*cameraType*': Menunjukkan jenis kamera yang digunakan untuk pengambilan gambar. Dalam *dataset* ini, nilai yang digunakan adalah *gp*, yang merupakan singkatan dari kamera *GoPro*
3. '*heightID*': Kode numerik yang merepresentasikan ketinggian drone saat gambar diambil. Terdapat dua nilai utama yang digunakan dalam penelitian ini.
  - a. Kode 3 : Mewakili skenario ketinggian rendah.
  - b. Kode 5 : Mewakili skenario ketinggian tinggi.
4. '*imageType*' : Menjelaskan jenis frame yang diekstrak dari video. Nilai yang umum adalah *ef*, yang berarti *every frame* (setiap frame dari video asli)
5. '*distanceID*' : Kode numerik dengan rentang dari 0 hingga 30, yang secara terbalik merepresentasikan jarak drone dan subjek. Nilai *distanceID* yang lebih kecil menandakan jarak yang lebih jauh, dan sebaliknya. Nilai ini nantinya dikonversi menjadi jarak dalam satuan meter untuk analisis skenario pengujian
  - a. Kode *distanceID* 0 - 13 : Mewakili skenario ketinggian jauh
  - b. Kode *distanceID* 14 - 24 : Mewakili skenario jarak menengah
  - c. Kode *distanceID* 25 - 30 : Mewakili skenario jarak dekat

$$Distance (m) = 17 - \left( \frac{distanceID}{2} \right)$$

**Rumus 3.1 DistanceID**

### 3.2.2.2 Pembuatan Dataset Resolusi Rendah (LR)

Berbeda dengan penelitian sebelumnya yang menurunkan resolusi citra secara artifisial, penelitian ini memanfaatkan struktur inheren dari dataset *DroneFace*. Dataset ini menyertakan folder terpisah bernama "*portrait*" yang berisi gambar wajah subjek berkualitas tinggi (*High-Resolution*). Gambar-gambar ini akan berfungsi sebagai galeri referensi (*gallery*), yaitu data acuan berkualitas tinggi. Sementara itu, 1364 gambar lainnya yang diambil menggunakan *drone* akan berfungsi sebagai citra uji (*probe*) yang secara alami memiliki resolusi dan kualitas yang lebih rendah akibat jarak dan kondisi pengambilan.

### 3.2.2.3 Struktur dan Pembagian Data

Penelitian ini akan menggunakan pendekatan evaluasi dimana seluruh gambar dari folder "*portrait*" membentuk basis data galeri referensi. Seluruh 1364 gambar dari dataset utama akan menjadi set pengujian. Untuk analisis yang lebih mendalam, data uji ini akan dikelompokkan ke dalam beberapa skenario pengujian berdasarkan metadata *heightID* dan *distanceID* untuk mengevaluasi kinerja sistem dalam kondisi yang berbeda-beda.

## 3.3 Desain dan Kerangka Kerja Penelitian

Penelitian ini mengadopsi pendekatan kuantitatif eksperimental. Kontribusi utamanya tidak terletak pada perancangan arsitektur model baru dari awal, melainkan pada integrasi dan evaluasi sistematis dari alur kerja (*pipeline*) yang menggabungkan teknologi-teknologi canggih yang telah ada, yaitu *GFPGAN*

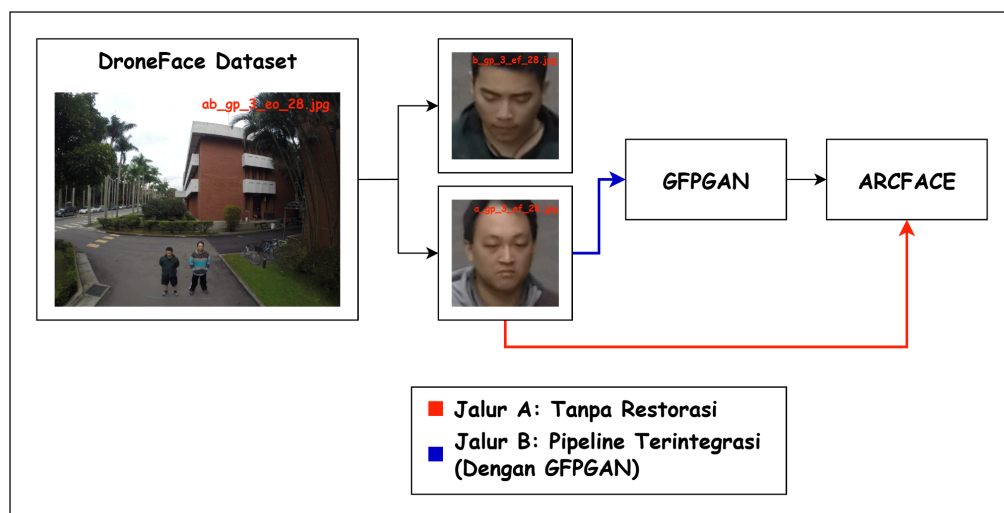
untuk restorasi dan *ArcFace* untuk rekognisi, pada sebuah studi kasus yang relevan. Untuk memfasilitasi analisis komparatif, penelitian dirancang dengan dua jalur eksperimental utama:

#### A. Jalur A: Tanpa Restorasi

Mengukur penurunan kinerja akibat degradasi kualitas citra dengan menerapkan *ArcFace* langsung pada citra uji dari *DroneFace*.

#### B. Jalur B: Pipeline Terintegrasi (Dengan GFPGAN)

Merupakan sistem yang diusulkan, yang mengimplementasikan alur kerja restorasi menggunakan *GFPGAN* pada citra uji sebelum melakukan rekognisi dengan *ArcFace*.



**Gambar 3.1 Diagram Kerangka Kerja Penelitian**

### 3.4 Objek dan Instrumen Penelitian

Bagian ini merinci spesifikasi teknis dari seluruh komponen yang digunakan, mencakup dataset sebagai objek penelitian serta perangkat lunak dan keras sebagai instrumen penelitian.

#### 3.4.1 Objek Penelitian: *Dataset LFW*

Objek utama dalam penelitian ini adalah data citra yang bersumber dari dataset *DroneFace* (Hsu & Chen, 2017), termasuk citra referensi dari folder "*portrait*" dan citra uji dari dataset utama.

#### 3.4.2 Instrumen Penelitian

##### 3.4.2.1 Model Pre-trained

1. Model Restorasi Wajah (GFPGAN): Model yang digunakan adalah versi pre-trained dari *GFPGAN* (misalnya, *GFPGANv1.4*). Model ini dipilih karena kemampuannya yang terbukti dalam restorasi wajah blind di dunia nyata, dengan memanfaatkan prior dari *StyleGAN2* untuk menghasilkan wajah yang realistis dan menjaga identitas (Wang et al., 2021).

2. Model Rekognisi Wajah (*ArcFace*)

Model yang digunakan adalah implementasi *ArcFace* yang menggunakan arsitektur backbone *ResNet-50*. Model ini terkenal karena kemampuannya menghasilkan *embedding feature* yang sangat diskriminatif berkat penggunaan fungsi *Additive Angular Margin Loss*, yang secara efektif

memaksimalkan jarak antar kelas di ruang fitur angular untuk meningkatkan daya pembeda (Deng et al., 2022).

### 3.4.2.2 Lingkungan Implementasi

#### 1. Perangkat Keras

Penelitian akan diimplementasikan pada komputer yang dilengkapi dengan *Graphics Processing Unit (GPU) NVIDIA RTX 3050* dengan *VRAM 8 GB* untuk mengakselerasi proses inferensi model.

#### 2. Perangkat Lunak

##### a. Bahasa Pemrograman

*Python* (versi 3.9 atau lebih tinggi).

##### b. Pustaka Utama

###### i. *PyTorch* ( $\geq 1.7$ )

Kerangka kerja deep learning utama untuk menjalankan model *GFPGAN* dan *ArcFace*.

###### ii. *Basicsr*

Pustaka dependensi untuk menjalankan *GFPGAN*.

###### iii. *OpenCV-Python*

Untuk operasi dasar pemrosesan citra (membaca, menulis, resizing).

###### iv. *NumPy*

Untuk operasi numerik, terutama pada vektor fitur.

v. ***Scikit-learn***

Untuk menghitung metrik evaluasi seperti akurasi dan kurva ROC.

vi. ***Pyiqa***

"Pustaka komprehensif untuk *Image Quality Assessment (IQA)*. Digunakan untuk mengevaluasi kualitas citra secara objektif dengan metrik *No-Reference* seperti *BRISQUE* dan *NIQE*."

### **3.5 Prosedur Penelitian**

Prosedur penelitian mengikuti alur kerja yang telah didefinisikan dalam tiga jalur eksperimental untuk memastikan perbandingan yang adil dan terstruktur.

#### **3.5.1 Tahap Persiapan: Pembuatan Galeri Referensi**

1. Tahap Input: Mengambil seluruh citra dari folder "*portrait*".
2. Tahap Proses: Setiap citra diumpankan ke model *ArcFace* untuk diekstraksi menjadi vektor fitur (*embedding*) 512-dimensi. Vektor-vektor ini disimpan sebagai *database* galeri referensi.

#### **3.5.2 Jalur A: Kinerja Tanpa Restorasi**

1. Tahap *Input*: Mengambil sebuah citra uji dari dataset utama *DroneFace*.
2. Tahap Proses: Citra uji diumpankan langsung ke model *ArcFace* untuk diekstraksi menjadi vektor fitur.



3. Tahap Output: Jarak *cosine-similarity* dihitung antara vektor fitur uji dengan seluruh vektor di galeri referensi. Hasil ini menunjukkan kinerja dasar sistem.

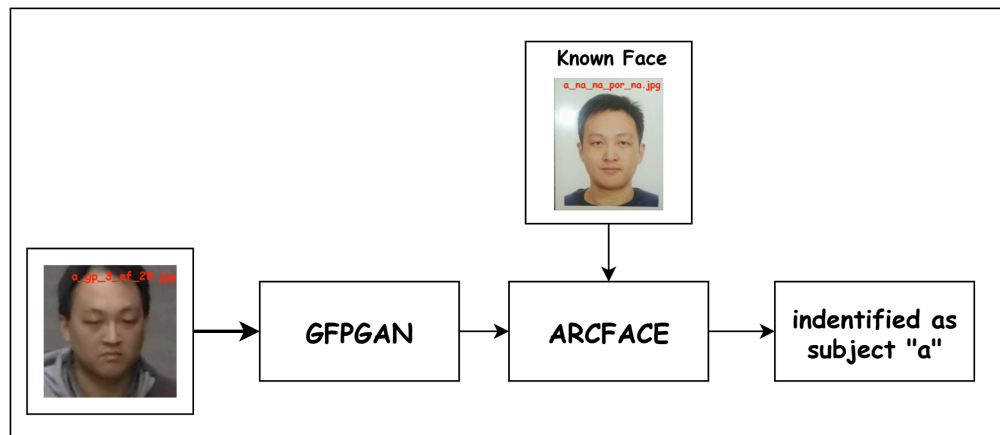
### 3.5.3 Jalur B: Pipeline Terintegrasi (Dengan Restorasi)

1. Tahap *Input*: Mengambil sebuah citra uji dari dataset utama *DroneFace*.
2. Tahap Restorasi (GFPGAN): Citra uji dimasukkan ke model *GFPGANer*.

Proses ini menghasilkan tiga output utama:

- a. *cropped\_faces*: Wajah asli yang berhasil dideteksi dan dipotong dari citra
  - b. *restored\_faces*: Versi restorasi dari wajah yang dipotong
  - c. *restored\_pasted\_img*: Gambar uji asli dengan wajah yang sudah direstorasi ditempelkan kembali.
3. Tahap Penyimpanan Visual: Semua gambar disimpan, agar dapat dilakukan analisis kualitatif, termasuk gambar perbandingan *cmp* yang dibuat dengan menggabungkan *cropped\_faces* dan *restored\_face*.
  4. Tahap Ekstraksi Fitur (*ArcFace*): Hanya gambar yang berhasil dideteksi yang diberikan ke model *ArcFace* untuk diekstraksi menjadi vektor fitur.
  5. Tahap Output: Jarak *cosine similarity* dihitung antara vektor fitur yang telah direstorasi dengan seluruh vektor di galeri referensi. Hasil ini dievaluasi untuk melihat efektivitas pipeline yang diusulkan.
  6. Tahap *Post Processing*: Untuk memastikan analisis perbandingan yang adil, setiap kegagalan dalam alur kerja Jalur B akan dicatat dan hasil akhir

untuk citra tersebut akan dianggap sebagai kegagalan identifikasi (prediksi salah)



**Gambar 3.2 Diagram Alur Kerja Detail Sistem Identifikasi**

### 3.6 Metode Evaluasi Kinerja

Bagian ini mendefinisikan metrik kuantitatif yang akan digunakan untuk mengevaluasi kinerja sistem pengenalan wajah pada setiap skenario pengujian secara objektif.

### 3.6.1 Skenario Pengujian

Skenario pengujian dirancang untuk menjawab setiap rumusan masalah secara empiris, dengan fokus pada perbandingan antara Jalur A (Tanpa Restorasi) dan Jalur B (Dengan Restorasi) pada kondisi berikut:

1. **Eksperimen Pertama (Analisis Pengaruh Jarak):** Membandingkan kinerja pada kelompok jarak dekat, menengah, dan jauh. Kategori jarak ditentukan berdasarkan *distanceID* yang ada pada nama file, yang memiliki rentang nilai dari 0 hingga 30. Berdasarkan kode implementasi, kategori ini didefinisikan sebagai berikut:

Jauh: Untuk *distanceID* dengan nilai kurang dari 14.

Menengah: Untuk *distanceID* dengan nilai antara 14 hingga 24.

Dekat: Untuk *distanceID* dengan nilai lebih dari 24.

2. **Eksperimen Kedua (Analisis Pengaruh Ketinggian):** Membandingkan kinerja pada kelompok ketinggian rendah dan tinggi. Kategori ketinggian ditentukan berdasarkan *heightID* pada nama file:

Rendah: Untuk *heightID* dengan nilai '3'.

Tinggi: Untuk *heightID* dengan nilai '5'.

3. **Eksperimen Ketiga (Analisis Komprehensif):** Menganalisis dan membandingkan hasil dari kedua jalur pada keseluruhan data uji untuk mengevaluasi kinerja *pipeline* secara menyeluruh.

### 3.6.2 Evaluasi Kuantitatif Kinerja Rekognisi

Evaluasi akan difokuskan pada tugas identifikasi, yaitu menentukan identitas subjek dalam citra uji dengan benar. Metrik yang digunakan adalah:

1. **Precision:** Mengukur seberapa banyak identifikasi yang benar dari total identifikasi yang dilakukan.

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive}$$

#### Rumus 3.2 Persamaan *Precision*

2. **Recall:** Mengukur seberapa banyak subjek yang berhasil diidentifikasi dengan benar dari total kemunculan mereka.

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative}$$

#### Rumus 3.3 Persamaan *Recall*

3. **F1-Score:** Rata-rata harmonik dari Precision dan Recall untuk memberikan skor tunggal yang seimbang.

$$F1-Score = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

#### Rumus 3.4 Persamaan *F1-Score*

Skenario	Metode	Precision	Recall	F1-Score	Peningkatan F1-Score (%)
<b>Pengaruh Jarak</b>					
Jarak Dekat	Tanpa Restorasi	(nilai)	(nilai)	(nilai)	-
	Dengan Restorasi	(nilai)	(nilai)	(nilai)	(nilai)
Jarak Menengah	Tanpa Restorasi	(nilai)	(nilai)	(nilai)	-
	Dengan Restorasi	(nilai)	(nilai)	(nilai)	(nilai)
Jarak Jauh	Tanpa Restorasi	(nilai)	(nilai)	(nilai)	-
	Dengan Restorasi	(nilai)	(nilai)	(nilai)	(nilai)
<b>Pengaruh Ketinggian</b>					
Ketinggian Rendah (3m)	Tanpa Restorasi	(nilai)	(nilai)	(nilai)	-
	Dengan Restorasi	(nilai)	(nilai)	(nilai)	(nilai)
Ketinggian Tinggi (5m)	Tanpa Restorasi	(nilai)	(nilai)	(nilai)	-
	Dengan Restorasi	(nilai)	(nilai)	(nilai)	(nilai)
<b>Keseluruhan</b>					
Semua Data Uji	Tanpa Restorasi	(nilai)	(nilai)	(nilai)	-
	Dengan Restorasi	(nilai)	(nilai)	(nilai)	(nilai)

**Tabel 3.1 - Templat Hasil Perbandingan Kinerja (Precision, Recall, F1-Score)**

### 3.6.3 Evaluasi Kuantitatif Kualitas Restorasi (*NR-IQA*)

Untuk mengukur efektivitas modul *GFPGAN* secara objektif, penelitian ini menggunakan dua metrik *No-Reference Image Quality Assessment (NR-IQA)*. Metrik ini menilai kualitas citra hasil restorasi tanpa memerlukan citra referensi (*ground truth*) yang bersih. Skor yang lebih rendah pada kedua metrik ini mengindikasikan kualitas visual yang lebih baik dan lebih natural.

- ***NIQE (Natural Image Quality Evaluator)***

Metrik ini mengevaluasi kualitas dengan membangun model 'sadar kualitas' dari fitur statistik yang berasal dari citra alami, tanpa memerlukan pelatihan pada data opini manusia (Mittal et al., 2013)

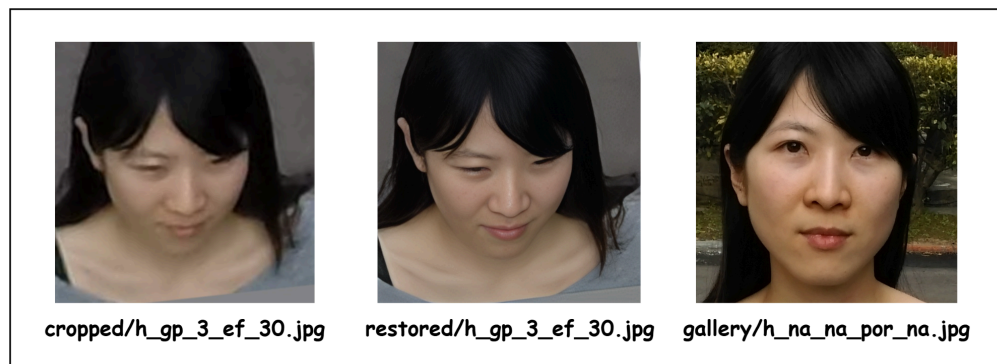
- ***BRISQUE (Blind/Referenceless Image Spatial Quality Evaluator)***

Metrik ini memanfaatkan statistik pemandangan alam (*Natural Scene Statistics*) di domain spasial untuk mengukur kemungkinan distorsi pada citra (Mittal et al., 2012)

Skor yang lebih rendah pada kedua metrik ini mengindikasikan kualitas visual yang lebih baik dan lebih natural.

### 3.6.4 Evaluasi Kualitatif

Selain metrik kuantitatif, analisis kualitatif juga akan dilakukan dengan menyajikan perbandingan visual. Gambar akan ditampilkan secara berdampingan: (a) Citra uji asli, (b) Citra hasil restorasi *GFPGAN*, dan (c) Citra referensi HR dari folder "*portrait*". Analisis ini bertujuan untuk menilai secara subjektif sejauh mana *GFPGAN* berhasil merekonstruksi detail wajah dan menjaga identitas subjek.



**Gambar 3.3 Contoh Perbandingan Visual Hasil Restorasi *GFPGAN***

## References

Hangrangi, S., Singh, T., & N, N. (2023). Face Detection and Recognition Using Face Mesh and Deep Neural Network. *International Conference on Machine Learning and Data Engineering*, 218, 741-749.  
<https://doi.org/10.1016/j.procs.2023.01.054>