

**PENGELOMPOKAN FOTO OTOMATIS BERDASARKAN
IDENTIFIKASI WAJAH MENGGUNAKAN ALGORITMA
DENSITY BASED SPATIAL CLUSTERING
*OF APPLICATION WITH NOISE***



Diajukan untuk Memenuhi Salah Satu Syarat Memperoleh Gelar
Sarjana Komputer Jurusan Teknik Informatika
pada Fakultas Sains dan Teknologi
UIN Alauddin Makassar

Oleh:

RAFIUL MUIZ.K
NIM 60200121003

UNIVERSITAS ISLAM NEGERI
ALAUDDIN
MAKASSAR

**FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI ALAUDDIN MAKASSAR
2025**

PERNYATAAN KEASLIAN SKRIPSI

Mahasiswa yang bertanda tangan di bawah ini:

Nama : Rafiul Muiz.K
NIM : 60200121003
Tempat/Tanggal Lahir : Mattoanging, 21 Februari 2003
Jurusan : Teknik Informatika
Fakultas : Sains dan Teknologi
Alamat : Mattoanging
Judul : Pengelompokan Foto Otomatis Berdsarkan Identifikasi Wajah Menggunakan Algoritma Density Based Spatial Clustering of Application with Noise

menyatakan dengan sesungguhnya dan penuh kesadaran bahwa skripsi ini benar adalah hasil karya sendiri. Jika di kemudian hari terbukti bahwa ia merupakan duplikat, tiruan, plagiat, atau dibuat oleh orang lain, sebagian atau seluruhnya, maka skripsi dan gelar yang diperoleh karenanya batal demi hukum.

Gowa, 25 Agustus 2025

Penyusun,

Rafiul Muiz.K
NIM. 60200121003

PERSETUJUAN UJIAN SKRIPSI (MUNAQASAH)



PENGESAHAN SKRIPSI

Skripsi berjudul "**Pengelompokan Foto Otomatis Berdasarkan Identifikasi Wajah Menggunakan Algoritma Density Based Spatial Clustering of Application with Noise**", yang disusun oleh Saudara Rafiul Muiz.K, NIM: 60200121003, mahasiswa Program Studi Teknik Informatika Fakultas Sains dan Teknologi UIN Alauddin Makassar, telah diuji dan dipertahankan dalam sidang Ujian Skripsi (Munaqasyah) yang diselenggarakan pada hari Rabu, 20 Agustus 2025 M. bertepatan dengan tanggal 25 Safar 1447 H, dinyatakan telah dapat diterima sebagai salah satu syarat untuk memperoleh gelar Sarjana Komputer pada Jurusan Teknik Informatika dengan beberapa perbaikan.

Gowa, 25 Agustus 2025

DEWAN PENGUJI:
Nomor SK 3353 Tahun 2025

Ketua Sidang : Dr. Ermawati, S.Pd., M.Si. ()

Sekretaris Sidang : Hariani, S.Kom., M.Kom. ()

DEWAN PENGUJI :

1. Mustikasari, S.Kom., M.Kom. ()
2. Dr. Hj. Besse Ruhaya, S.Pd.I., M.Pd.I. ()
3. Dr. Ridwan A. Kambau, S.T., M.Kom. ()
4. M. Hasrul H, S.Kom., M.Kom. ()

Disahkan Oleh:
Dekan FST UIN Alauddin Makassar,

Ar. Fahmyddin A'raaf Tauhid, S.T., M. Arch., Ph.D., IAI.
NIP. 197606102006041004

KATA PENGANTAR

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

Segenap puji dan syukur ke hadirat Allah SWT, atas berkat rahmat dan hidayah-Nya, penulis berhasil merampungkan skripsi yang berjudul, "Pengelompokan Foto Otomatis Berdasarkan Identifikasi Wajah Menggunakan Algoritma Density Based Spatial Clustering of Application with Noise". Penyusunan skripsi ini merupakan salah satu prasyarat untuk memperoleh gelar Sarjana (S1) pada program studi Teknik Informatika, Fakultas Sains dan Teknologi, Universitas Islam Negeri Alauddin Makassar. Shalawat serta salam senantiasa tercurah kepada junjungan kita, Nabi Muhammad saw. yang menjadi teladan utama dalam setiap aspek kehidupan.

Dengan segala kerendahan hati, penulis mengakui bahwa skripsi ini masih jauh dari sempurna akibat keterbatasan wawasan dan pengalaman. Untuk itu, segala bentuk masukan, kritik, dan saran yang membangun sangat dinantikan demi penyempurnaan karya ini.

Keberhasilan penulis dalam menyelesaikan skripsi ini tentu berkat kontribusi banyak pihak. Ucapan terima kasih tertinggi penulis sampaikan kepada kedua orang tua, ayahanda **Kasmadi** dan ibunda **Harming, S.Pd.I.** Terima kasih banyak atas segala dukungan moril, perhatian, doa yang tak pernah berhenti, serta motivasi yang menjadi sumber kekuatan penulis. Semua pengorbanan, kasih sayang, dan perjuangan yang telah diberikan adalah fondasi tak ternilai bagi kehidupan penulis.

Penulis juga tak lupa mengucapkan terima kasih yang sebesar-besarnya kepada:

1. **Bapak Prof. Dr. Hamdan Juhannis, M.A., Ph.D** Rektor UIN Alauddin Makassar beserta wakil rektor UIN Alauddin Makassar.

2. **Bapak Ar. Fahmyddin A'raf Tauhid, S.T., M.Arch., Ph.D., IAI.** Dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Alauddin Makassar, beserta Wakil Dekan I, Wakil Dekan II, Wakil Dekan III, dan seluruh staf administrasi yang telah memberikan berbagai fasilitas kepada kami selama masa pendidikan.
3. **Ibu Mustikasari, S.Kom., M.Kom.** Selaku Ketua Jurusan Teknik Informatika dan selaku Penguji I penulis, serta **Ibu Hariani, S.Kom., M.Kom,** selaku Sekretaris Jurusan Teknik Informatika yang selalu memberikan saran dan motivasi kepada mahasiswanya
4. **Bapak Dr. Ridwan A Kambau, S.T., M.Kom.** sebagai Pembimbing I yang selalu meluangkan waktu, memberikan motivasi dan membimbing penulis mulai dari pembuatan proposal sampai pada penyusunan skripsi ini selesai.
5. **Bapak M. Hasrul H, S.Kom., M.Kom.** sebagai pembimbing II yang selalu meluangkan waktu, memberikan motivasi dan membimbing penulis mulai dari pembuatan proposal sampai pada penyusunan skripsi ini selesai.
6. **Ibu Dr. Besse Ruhaya, S.Pdi., M.Pdi.** Selaku penguji II yang telah memberikan gambaran serta masukan yang sangat membantu penulis dalam menyelesaikan skripsi ini.
7. **Seluruh dosen, pejabat, dan staff** Fakultas Sains dan Teknologi dan dosen Jurusan Teknik Informatika yang telah mengajar penulis selama menempuh studi di Jurusan Teknik Informatika.
8. Teman-teman seperjuangan selama di Teknik Informatika angkatan 2021 “21MBRA” yang setia bersamai penulis dalam suka maupun duka dan memberikan semangat dalam menempuh pendidikan.

9. Seluruh pihak yang tidak dapat penulis sebutkan satu per satu namun telah banyak terlibat dalam membantu penulis selama proses penyusunan skripsi ini.

Penulis berharap skripsi ini dapat berkontribusi dalam memajukan ilmu pengetahuan serta memberikan manfaat yang luas bagi pembaca. Semoga semua bantuan dan dukungan yang diterima menjadi amal jariyah yang pahalanya dilipatgandakan oleh Allah SWT.

Samata, 25 Agustus 2025

Penulis,

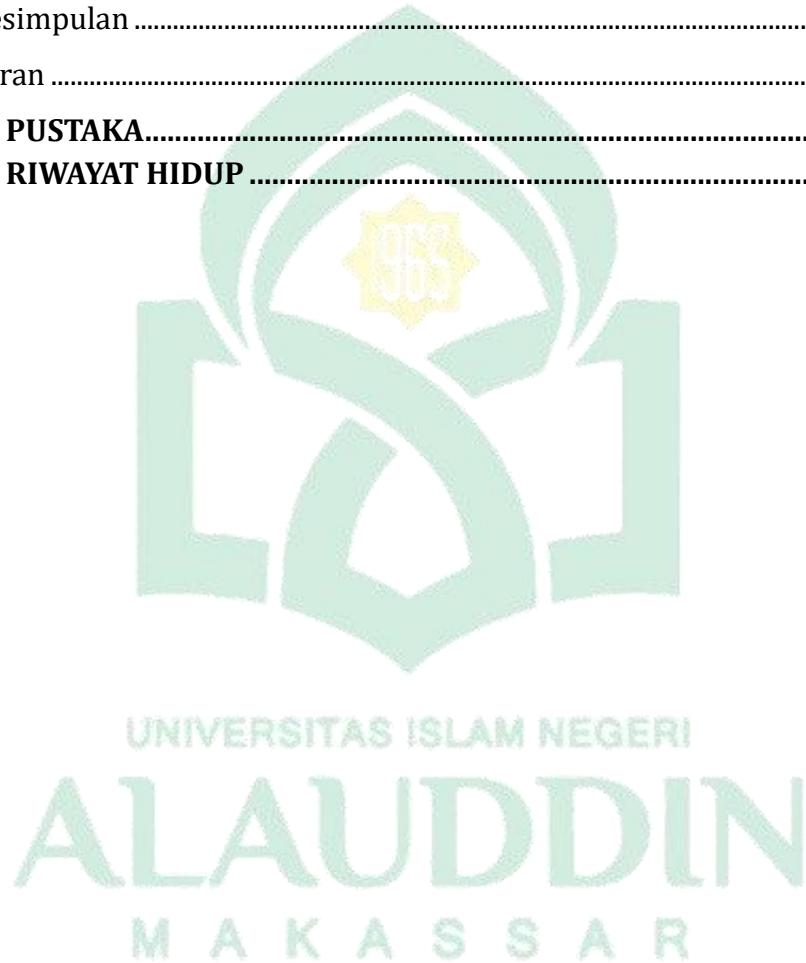
Rafiul Muiz.K
NIM: 60200121003



DAFTAR ISI

PERNYATAAN KEASLIAN SKRIPSI	ii
PERSETUJUAN UJIAN SKRIPSI (MUNAQASYAH)	iii
PENGESAHAN SKRIPSI	iv
KATA PENGANTAR	v
DAFTAR ISI	viii
DAFTAR GAMBAR.....	x
DAFTAR TABEL	xi
ABSTRAK.....	xii
BAB I PENDAHULUAN	1
A. Latar Belakang Masalah.....	1
B. Rumusan Masalah.....	8
C. Definisi Operasional Variabel dan Ruang Lingkup Penelitian	8
D. Kajian Pustaka/ Penelitian Terdahulu	10
E. Tujuan dan Kegunaan Penelitian.....	15
BAB II TINJAUAN TEORETIS	17
A. Kajian Integrasi Keilmuan	17
B. Pengelompokan Foto.....	19
C. Face Recognition	20
D. Feature Extraction	22
E. Density Based Spatial Clustering of Application With Noise	24
F. Facebook AI Similarity Search	27
G. InsightFace	29
H. Google Drive API.....	32
BAB III METODOLOGI PENELITIAN.....	35
A. Jenis Penelitian.....	35
B. Pendekatan Penelitian.....	35
C. Sumber Data.....	36
D. Metode Pengumpulan dan Pengolahan Data.....	36
E. Instrumen Penelitian	37
F. Teknik Pengolahan dan Analisis Data	38
G. Teknik Pengujian dan Evaluasi Model.....	40

H.	Perancangan Sistem.....	41
I.	Perancangan Antarmuka.....	45
BAB IV HASIL DAN PEMBAHASAN		51
A.	Hasil Penelitian	51
B.	Pembahasan	57
C.	Analisis Optimasi Parameter DBSCAN.....	64
BAB V PENUTUP		67
A.	Kesimpulan	67
B.	Saran	68
DAFTAR PUSTAKA.....		69
DAFTAR RIWAYAT HIDUP		72



DAFTAR GAMBAR

Gambar 2.1 Alur Kerja <i>Clustering</i> DBSCAN dengan FAISS	26
Gambar 2.2 Alur Kerja InsightFace	31
Gambar 2.3 Google Drive API.....	33
Gambar 3.1 Kerangka Kerja CRISP-DM	37
Gambar 3.2 <i>Use Case</i> Diagram Sistem yang Diusulkan.....	42
Gambar 3.3 <i>Activity</i> Diagram Sistem yang Diusulkan	44
Gambar 3.4 <i>Wireframe</i> Halaman Utama.....	45
Gambar 3.5 <i>Wireframe</i> Informasi Link Google Drive	46
Gambar 3.6 <i>Wireframe</i> Informasi Lanjutan Link Google Drive	46
Gambar 3.7 <i>Wireframe</i> Input Tautan Google Drive	47
Gambar 3.8 <i>Wireframe</i> Progress Pengelompokan Foto	48
Gambar 3.9 <i>Wireframe</i> Hasil Pengelompokan Foto	48
Gambar 3.10 <i>Wireframe Preview Cluster</i>	49
Gambar 3.11 <i>Wireframe</i> Notifikasi Adaptif Mode ZIP	50
Gambar 4.1 Tampilan Halaman Utama Sistem	51
Gambar 4.2 Tampilan Halaman Informasi <i>Link</i> Google Drive	52
Gambar 4.3 Tampilan Halaman Informasi Lanjutan <i>Link</i> Google Drive	52
Gambar 4.4 Tampilan Input <i>Link</i> Google Drive	53
Gambar 4.5 Tampilan Status Proses dan Waktu Berjalan	53
Gambar 4.6 Tampilan Hasil Pengelompokan Foto.....	54
Gambar 4.7 Tampilan Pratinjau Gambar dalam Satu Cluster	55
Gambar 4.8 Notifikasi Jaringan Lambat dan Opsi Beralih ke Mode ZIP	55
Gambar 4.9 Diagram Waktu Pemrosesan	59
Gambar 4.10 Diagram Metrik Kerja DBSCAN.....	60
Gambar 4.11 Diagram Kualitas Data	61
Gambar 4.12 Diagram Perbandingan Metrik Kerja DBSCAN.....	61
Gambar 4.13 Diagram Perbandingan Waktu Proses	62
Gambar 4.14 <i>Unclustered Faces</i>	63

DAFTAR TABEL

Tabel 1.1 Komparasi Kajian Penelitian Sebelumnya dan Penelitian yang Diusulkan	12
Tabel 4.1 Hasil Pengujian Berdasarkan Kuantitas Data	56
Tabel 4.2 Hasil Pengujian Berdasarkan Kualitas Data	56
Tabel 4.3 Hasil Perbandingan Metode Input	57
Tabel 4.4 Hasil Pengujian Awal	65
Tabel 4.5 Hasil Pengujian Akhir / Terbaru	65



ABSTRAK

Nama : Rafiul Muiz.K

NIM : 60200121003

Judul : Pengelompokan Foto Otomatis Berdasarkan Identifikasi Wajah Menggunakan Algoritma *Density Based Spatial Clustering of Applications with Noise*

Tantangan dalam mengelola koleksi foto digital yang masif, terutama dalam menemukan foto personal dari dokumentasi kegiatan, menjadi masalah ineffisiensi di era digital. Penelitian ini bertujuan untuk merancang dan mengimplementasikan sistem pengelompokan foto otomatis untuk mempermudah pengguna dalam mengakses foto personal mereka berdasarkan identifikasi wajah. Sistem ini menerapkan algoritma *Density-Based Spatial Clustering of Applications with Noise* (DBSCAN) untuk proses pengelompokan. Data foto yang bersumber dari Google Drive diproses melalui beberapa tahapan yaitu deteksi wajah, ekstraksi fitur menggunakan *InsightFace* untuk menghasilkan *embedding* vektor, dan pengelompokan vektor fitur tersebut menggunakan DBSCAN.

Hasil penelitian menunjukkan bahwa sistem yang dikembangkan berhasil mengotomatisasi pengelompokan foto secara efisien dan akurat. Dalam pengujian, sistem mampu memproses 1.067 foto dalam waktu sekitar 20 menit dan mencapai *Silhouette Score* setinggi 0,817, yang menandakan bahwa kelompok-kelompok foto yang dihasilkan sangat padat dan terpisah dengan baik. Hal ini membuktikan bahwa penerapan algoritma DBSCAN efektif dalam mengidentifikasi dan mengelompokkan individu yang berbeda secara akurat.

Pemanfaatan sistem pengelompokan foto otomatis ini memberikan solusi praktis untuk mengatasi kesulitan pencarian foto personal. Implementasi ini tidak hanya meningkatkan efisiensi pengelolaan dokumentasi bagi organisasi, tetapi juga dapat diadaptasi untuk berbagai kebutuhan lain, seperti manajemen arsip foto keluarga atau perusahaan, sehingga berkontribusi pada ekosistem digital yang lebih terorganisir dan berpusat pada pengguna.

Kata Kunci: Pengelompokan Foto, Identifikasi Wajah, DBSCAN, *InsightFace*

BAB I

PENDAHULUAN

A. Latar Belakang Masalah

Era digital telah membawa perubahan signifikan dalam cara manusia mendokumentasikan momen-momen penting (Razec, 2022). Transformasi ini tidak hanya terjadi pada tingkat teknologi, tetapi juga telah mengubah perilaku sosial dan budaya masyarakat secara fundamental. Sebelum era digital, dokumentasi visual dibatasi oleh keterbatasan media fisik seperti *film* dan kertas foto, yang menyebabkan selektivitas tinggi dalam pengambilan gambar. Namun, dengan kemunculan kamera digital dan *smartphone* yang dilengkapi dengan kamera beresolusi tinggi, hambatan ini praktis telah hilang, memungkinkan pengambilan gambar tanpa batas dengan biaya marginal yang mendekati nol.

Menurut data Statista, lebih dari 1,8 miliar foto diunggah ke berbagai platform digital setiap harinya (Statista, 2022). Pertumbuhan visualisasi data ini juga didorong oleh kemajuan teknologi penyimpanan *cloud* seperti Google Drive yang semakin terjangkau, sehingga memungkinkan penyimpanan jangka panjang untuk volume data yang sebelumnya sulit dibayangkan. Namun, meskipun kemudahan pengambilan dan penyimpanan foto meningkat, pengelolaan dan pengaksesan foto-foto tersebut masih menjadi tantangan besar bagi pengguna.

Tantangan pengelolaan data visual ini semakin terasa dalam konteks aktivitas organisasi kemahasiswaan. Organisasi kemahasiswaan, sebagai entitas yang dinamis dan aktif, secara rutin menyelenggarakan berbagai kegiatan yang melibatkan puluhan hingga ratusan partisipan. Setiap kegiatan ini biasanya didokumentasikan secara ekstensif oleh tim dokumentasi yang

ditugaskan, menghasilkan ratusan foto yang merekam berbagai momen penting. Setelah mengikuti suatu kegiatan organisasi, dokumentasi kegiatan yang berjumlah sangat banyak dikirimkan dan disimpan dalam platform penyimpanan Google Drive. Proses penyimpanan yang umum dilakukan adalah dengan mengunggah seluruh foto ke dalam satu folder atau beberapa subfolder yang diorganisir berdasarkan tanggal atau nama kegiatan. Meskipun sistem pengorganisasian ini tampak logis dari perspektif administratif, pendekatan tersebut tidak memperhitungkan kebutuhan akses individual dari peserta kegiatan.

Permasalahan muncul ketika seseorang hanya ingin mengakses foto yang memuat wajahnya saja, suatu kebutuhan yang sangat wajar mengingat relevansi personal dari foto-foto tersebut. Dalam konteks organisasi kemahasiswaan, ilustrasi konkret dari permasalahan ini dapat dilihat pada kegiatan organisasi yang menghasilkan banyaknya dokumentasi kegiatan khususnya berupa foto. Perlu dipertimbangkan bahwa kegiatan organisasi mahasiswa yang berlangsung selama beberapa hari, seperti rapat kerja, musyawarah besar, atau pelaksanaan program kerja, dengan mudah menghasilkan volume foto dalam jumlah yang banyak, ratusan hingga ribuan foto. Proses pencarian manual pada kumpulan foto yang berjumlah ratusan hingga ribuan tersebut membutuhkan waktu yang sangat lama dan tidak efisien. Menurut penelitian yang dilakukan oleh International Data Corporation (IDC), sekitar 30% waktu kerja terbuang hanya untuk mencari dan mengorganisir data, termasuk file foto (IDC, 2022). Dalam konteks pencarian wajah pada kumpulan foto, ratusan hingga ribuan, seseorang mungkin mulai melewatkhan foto yang sebenarnya memuat wajahnya, atau sebaliknya, keliru mengidentifikasi foto yang tidak relevan sebagai foto yang

memuat wajahnya.

Waktu yang terbuang untuk mencari dan mengorganisir foto termasuk bentuk inefisiensi dan pemborosan waktu. Dalam perspektif Islam, pemborosan atau *tabdzir* adalah perilaku yang tidak tepat sasaran dan tidak membawa manfaat, sebagaimana terkandung dalam firman-Nya pada Surah Al-Isra' /17:26 sebagai berikut:

وَاتِّهَا الْقُرْبَىٰ حَقَّهُ وَالْمِسْكِينَ وَابْنَ السَّبِيلِ وَلَا تُبَذِّرْ تَبْذِيرًا (٢٦)

Terjemahnya:

"Berikanlah kepada kerabat dekat haknya, (juga kepada) orang miskin, dan orang yang dalam perjalanan. Janganlah kamu menghambur hamburkan (hartamu) secara boros" (Kementerian Agama, 2022a)

Berdasarkan Tafsir Al-Misbah, ayat di atas mengajarkan untuk menghindari pemborosan harta atau sumber daya, termasuk waktu (Shihab, 2021). Kaitan penafsiran ayat dengan penelitian ini bahwa meskipun di dalam ayat disebutkan secara spesifik larangan perilaku *tabdzir* atau pemborosan harta, namun harta yang dimaksud adalah segala sumber daya yang dimanfaatkan untuk kepentingan pribadi maupun kepentingan secara luas. Oleh karena itu, penekanan untuk menghindari pemborosan sumber daya khususnya sumber daya yang sering dilalaikan yaitu waktu. Mengapa demikian hal tersebut harus dihindari, karena islam mengajarkan pentingnya menghargai waktu dan menggunakan dengan bijak agar tidak ada lagi waktu yang terbuang sia-sia, sebagaimana terkandung dalam firman-Nya pada Surah Al-Ashr /103:1-3 sebagai berikut:

وَالْعَصْرِ (١) إِنَّ الْإِنْسَانَ لَفِي خُسْرٍ (٢) إِلَّا الَّذِينَ آمَنُوا وَعَمِلُوا الصِّلَاخَتِ

وَتَوَاصَوْا بِالْحَقِّ وَتَوَاصَوْا بِالصَّبْرِ (٣)

Terjemahnya:

"Demi masa, sesungguhnya manusia benar-benar berada dalam kerugian, kecuali orang-orang yang beriman dan beramal saleh serta

saling menasihati untuk kebenaran dan kesabaran" (Kementerian Agama, 2022b)

Berdasarkan Tafsir Quran Kemenag dengan metode penafsiran Tahlili (Kementerian Agama, 2022b), dalam Surah Al-Ashr ayat 1-3, Allah Swt. bersumpah "Demi masa" untuk menekankan pentingnya waktu sebagai anugerah berharga yang harus diperhatikan dan dimanfaatkan dengan optimal oleh manusia. Allah kemudian menegaskan bahwa manusia secara keseluruhan berada dalam kerugian bila tidak menggunakan waktunya dengan baik atau malah digunakan untuk perbuatan buruk, yang menjadi sumber kecelakaan dan kebinasaan. Pelanggaran dan dosa terhadap Allah yang telah memberi nikmat tak terhingga adalah kerugian yang tak terukur. Sebagai jalan keluar, Allah memberikan solusi agar manusia terhindar dari kerugian yaitu dengan beriman kepada-Nya, melaksanakan ibadah sesuai perintah-Nya, berbuat baik untuk diri sendiri, serta senantiasa berusaha memberikan manfaat kepada sesama, sehingga waktu yang dimiliki menjadi bernilai dan penuh berkah.

Kaitan dengan penelitian yang dilakukan bahwa dengan kesadaran akan pentingnya waktu seperti pada sumpah Allah Swt. Terhadap waktu yang menjadi landasan penlitian ini dilakukan dikarenakan tanpa melakukan pengelompokan foto otomatis, banyak waktu yang terbuang dengan sia-sia hanya untuk mencari ratusan bahkan ribuan foto dalam kumpulan foto dokumentasi kegiatan yang ada. Dalam kaitannya pula pada konteks modern, pemanfaatan waktu ini juga tercermin dari upaya manusia mengembangkan berbagai teknologi untuk mempermudah dan mengoptimalkan kehidupan, termasuk dalam memanfaatkan kemajuan teknologi informasi dan kecerdasan buatan yang berkembang pesat saat ini.

Dalam beberapa tahun terakhir, teknologi pengenalan wajah (*facial*

recognition) telah berkembang pesat, didorong oleh kemajuan dalam bidang kecerdasan buatan, khususnya *deep learning*. Kemajuan ini didorong oleh pengembangan arsitektur jaringan saraf yang khusus dirancang untuk tugas pengenalan wajah, seperti InsightFace yang diperkenalkan oleh DeepGlint pada tahun 2019, dan berbagai varian metode berbasis *Convolutional Neural Networks* (CNN) seperti ArcFace. Algoritma-algoritma ini mampu mengekstrak fitur wajah yang *robust* dan *invariant* terhadap berbagai transformasi, memungkinkan pengenalan wajah yang lebih akurat dan *reliable* dalam berbagai kondisi (Destiana dkk., 2024).

Di sisi lain, algoritma *clustering* khususnya *Density Based Spatial Clustering of Application With Noise* (DBSCAN) telah terbukti efektif dalam mengelompokkan data dengan pola spasial yang kompleks. DBSCAN, yang diperkenalkan pertama kali oleh Ester, Kriegel, Sander, dan Xu pada tahun 1996, berfungsi dengan mencari titik-titik data yang memiliki kepadatan tinggi (*density*) dan mengelompokkannya, sambil mengidentifikasi titik-titik dengan kepadatan rendah sebagai *noise* atau *outlier* (Sujatmiko, 2023). Pendekatan berbasis kepadatan ini memberikan DBSCAN keunggulan dibandingkan algoritma *clustering* tradisional seperti K-means, terutama dalam hal kemampuannya untuk menangani *cluster* dengan bentuk tidak beraturan dan mendeteksi *outlier* secara eksplisit. Berbagai studi komparatif telah ditunjukkan bahwa DBSCAN mampu mengidentifikasi *cluster* dengan bentuk tidak beraturan dan menangani *noise* dengan baik, karakteristik yang sangat relevan untuk data visual seperti foto yang memiliki variasi pose, pencahayaan, dan ekspresi wajah. Dalam konteks pengenalan wajah, fitur wajah yang diekstrak oleh model *deep learning* membentuk ruang fitur multidimensional yang kompleks, di mana wajah yang sama akan membentuk *cluster* dalam

ruang fitur tersebut. DBSCAN, dengan kemampuannya menangani data dalam ruang dimensi tinggi dan mengidentifikasi *cluster* dengan kepadatan berbeda, menjadi pilihan yang ideal untuk mengelompokkan representasi fitur wajah. Integrasi antara teknologi pengenalan wajah dan algoritma DBSCAN menawarkan potensi solusi untuk masalah pengelompokan foto otomatis. Dengan mengidentifikasi wajah pada setiap foto dan mengelompokkannya berdasarkan kesamaan fitur, sistem dapat secara otomatis mengategorikan foto berdasarkan individu yang terdapat di dalamnya. Pendekatan ini mampu mengatasi kelemahan sistem pengelompokan konvensional yang sering kali gagal dalam menangani variasi pose dan pencahayaan pada foto wajah (Mutiah dkk., 2024).

Sistem yang diusulkan akan mengkombinasikan detektor wajah modern untuk mengidentifikasi dan mengekstrak wajah dari foto, dengan model ekstraksi fitur wajah berbasis *deep learning* untuk mengubah citra wajah menjadi *embedding* vektor dalam ruang *Euclidean*. *Embedding* ini kemudian akan dikelompokkan menggunakan algoritma DBSCAN untuk menemukan *cluster* natural yang merepresentasikan individu yang sama. Pendekatan ini memiliki beberapa keunggulan dibandingkan dengan metode konvensional. Pertama, tidak seperti algoritma *clustering* tradisional seperti K-means yang memerlukan spesifikasi jumlah *cluster* di awal, DBSCAN dapat menemukan jumlah *cluster* secara otomatis berdasarkan distribusi data. Hal ini sangat penting dalam konteks pengelompokan foto berdasarkan wajah, di mana jumlah individu yang muncul dalam koleksi foto tidak diketahui sebelumnya. Kedua, DBSCAN memiliki kemampuan untuk mengidentifikasi *outlier* atau *noise*, yang dalam konteks foto wajah dapat berupa deteksi wajah yang salah atau wajah dengan kualitas rendah yang tidak dapat

dikelompokkan dengan percaya diri. Ketiga, algoritma ini dapat menangani *cluster* dengan kepadatan bervariasi, memungkinkan pengenalan individu yang muncul dalam jumlah foto yang sangat berbeda (misalnya, beberapa individu mungkin muncul dalam puluhan foto, sementara yang lain hanya muncul dalam beberapa) (Mutiah dkk., 2024).

Implementasi teknologi ini dalam konteks organisasi kemahasiswaan akan memberikan manfaat signifikan. Dari perspektif pengguna individu, sistem ini akan memungkinkan pencarian foto yang lebih cepat dan akurat berdasarkan kehadiran wajah mereka. Proses yang sebelumnya memakan waktu belasan hingga puluhan menit dapat direduksi. Dari perspektif organisasi, sistem ini dapat meningkatkan efisiensi dalam pengelolaan dan distribusi dokumentasi kegiatan. Tantangan implementasi yang mungkin dihadapi dalam pengembangan sistem ini mencakup variasi dalam kualitas foto, kondisi pencahayaan, dan pose wajah dapat mempengaruhi akurasi deteksi dan pengenalan wajah. Sistem perlu dirancang dengan *robustness* yang tinggi untuk menangani variasi ini. Skalabilitas juga menjadi perhatian penting, terutama ketika jumlah foto dan individu yang perlu diproses meningkat secara signifikan.

Berdasarkan uraian di atas, penelitian dengan judul "*Pengelompokan Foto Otomatis Berdasarkan Identifikasi Wajah Menggunakan Algoritma Density Based Spatial Clustering of Application With Noise*" diharapkan dapat memberikan solusi untuk permasalahan pengelolaan dan pengaksesan foto berdasarkan identitas wajah. Dengan memanfaatkan kemajuan teknologi pengenalan wajah dan algoritma *clustering*, penelitian ini bertujuan untuk mengembangkan sistem yang dapat secara otomatis mengklasifikasikan foto berdasarkan individu yang terdapat di dalamnya, sehingga meningkatkan

efisiensi dan akurasi dalam proses pencarian foto personal di tengah ledakan data visual era digital. Implementasi sistem ini diharapkan tidak hanya bermanfaat dalam konteks organisasi kemahasiswaan, tetapi juga dapat diadaptasi untuk berbagai konteks lain seperti manajemen foto keluarga, dokumentasi acara perusahaan, atau pengelolaan arsip visual institusi pendidikan. Pada akhirnya, penelitian ini berkontribusi pada pengembangan ekosistem digital yang lebih terorganisir, efisien, dan berpusat pada pengguna.

B. Rumusan Masalah

Berdasarkan uraian pada latar belakang di atas, dapat dirumuskan suatu permasalahan sebagai berikut:

1. Bagaimana sistem pengelompokan foto otomatis dapat membantu pengguna agar dapat dengan mudah mengakses foto personal mereka berdasarkan identitas wajah?
2. Bagaimana algoritma DBSCAN dapat diterapkan untuk otomatisasi proses pengelompokan foto berdasarkan identitas wajah?

C. Definisi Operasional Variabel dan Ruang Lingkup Penelitian

Agar pelaksanaan tugas akhir ini lebih terarah, ruang lingkup penelitian ini akan difokuskan pada pembahasan sebagai berikut:

1. Tema penelitian yang digunakan adalah implementasi algoritma *Density-Based Spatial Clustering of Applications with Noise* (DBSCAN) dalam sistem pengelompokan foto otomatis berbasis identitas wajah untuk meningkatkan kemudahan pencarian foto personal.
2. Data yang digunakan mencakup kumpulan foto kegiatan pada 2 organisasi yaitu Inready Workgroup dan Generasi Baru Indonesia (GenBI) dalam format digital (JPEG, JPG, PNG, HEIC) yang tersimpan di

platform penyimpanan cloud (Google Drive), dengan fokus pada foto yang mengandung wajah manusia.

3. Teknologi yang digunakan adalah algoritma DBSCAN yang dikombinasikan dengan InsightFace untuk ekstraksi fitur wajah, bertujuan mengidentifikasi dan mengelompokkan foto berdasarkan kemiripan identitas wajah.
4. Sistem yang dibuat fokus untuk data bagian wajah pada proses identifikasi

Agar mempermudah pemahaman dan memberikan gambaran serta menyamakan persepsi antara penulis dan pembaca, penelitian ini mengemukakan penjelasan berdasarkan definisi operasional. Berikut adalah definisi operasional variabel yang digunakan dalam penelitian ini:

1. Variabel independen adalah parameter DBSCAN (eps dan minimum sampel) dan fitur wajah yang diekstraksi oleh model deep learning. Variabel dependen adalah hasil pengelompokan foto yang dikelompokkan berdasarkan kesamaan wajah dan seberapa tepat sistem mengelompokkan foto berdasarkan identitas wajah yang sama.
2. Data yang digunakan mencakup foto kegiatan Inready Workgroup dan GenBI yang diambil dari link Google Drive, dengan proses praprosesan meliputi deteksi wajah (*face detection*) dan ekstraksi fitur wajah menggunakan InsightFace.
3. Evaluasi sistem dilakukan untuk mengukur akurasi pengelompokan menggunakan metrik *silhouette score* dan *Davies Bouldin Index*, dan kemudahan pengguna dalam mengakses foto personal mereka.

D. Kajian Pustaka/ Penelitian Terdahulu

Berikut beberapa penelitian sebelumnya yang telah dilakukan dan dianggap memiliki perbedaan yang ditemukan:

Penelitian yang dilakukan oleh (Choo dkk., 2024) yang berjudul "*Automatic Photo Classification System Based on Face Feature Extraction and Clustering*" membahas tentang sistem otomatis untuk mengelompokkan foto berdasarkan fitur wajah dan teknik *clustering*. Penelitian ini bertujuan memecahkan masalah pengelolaan koleksi foto digital yang besar, terutama yang berisi banyak wajah manusia. Sistem yang dikembangkan terdiri dari tiga tahap utama yaitu: deteksi wajah menggunakan Multi Task Cascaded Convolutional Network atau MTCNN, ekstraksi fitur wajah menggunakan *FaceNet*, dan pengelompokan wajah menggunakan algoritma DBSCAN. Hasil pengelompokan kemudian digunakan untuk mengorganisir foto ke dalam folder-folder yang merepresentasikan individu yang terdeteksi. Studi ini menekankan keefektifan pendekatan kombinasi *deep learning* dan *clustering* dalam meningkatkan efisiensi klasifikasi foto secara otomatis, terutama dalam skenario personal seperti manajemen galeri pengguna. Sistem tersebut memiliki persamaan dan perbedaan dengan yang diusulkan oleh peneliti. Adapun persamaannya, yaitu kedua penelitian menggunakan algoritma DBSCAN untuk proses *clustering*. Persamaan berikutnya terkait dengan tujuan penelitian, yakni mempermudah pengelolaan foto berdasarkan identitas wajah. Namun, terdapat perbedaan dalam hal ekstraksi fitur wajah. Choo dkk. Menggunakan *FaceNet* untuk ekstraksi wajah, sedangkan pada penelitian ini menggunakan *InsightFace* untuk ekstraksi wajah.

Penelitian yang dilakukan oleh (Varun dkk., 2024) yang berjudul "*Face Clustering with Advanced Techniques*" membahas cara untuk mengelompokkan

foto wajah secara otomatis. Mereka menggunakan teknik *deep learning* (pembelajaran mendalam) dan algoritma canggih untuk meningkatkan akurasi pengenalan wajah, terutama pada dataset yang rumit. Fokusnya adalah mengatasi masalah seperti pose wajah yang berbeda, pencahayaan tidak konsisten, atau ekspresi wajah beragam. Caranya dengan menggabungkan metode seperti *graph-based clustering* (pengelompokan berbasis graf), *metric learning* (pengukuran jarak antar-data), dan *autoencoder* (jaringan saraf untuk ekstraksi fitur). Mereka juga mengevaluasi algoritma DBSCAN dan menemukan bahwa kombinasi ekstraksi fitur yang baik dengan algoritma berbasis kepadatan (seperti DBSCAN) bisa mengurangi kesalahan pengelompokan. Sistem tersebut memiliki persamaan dan perbedaan dengan yang diusulkan oleh peneliti. Adapun persamaannya, penelitian ini sama-sama menggunakan algoritma DBSCAN untuk mengatasi data yang tidak merata dan menghilangkan *noise*. Kedua penelitian juga menekankan pentingnya ekstraksi fitur wajah sebagai langkah awal. Namun, perbedaan utama terlihat pada cakupan dan tujuan penelitian. Varun dkk. mengintegrasikan berbagai metode kompleks seperti *graph-based clustering* dan *metric learning* untuk meningkatkan fleksibilitas sistem pada dataset besar, sementara penelitian ini berfokus pada algoritma DBSCAN untuk kondisi spesifik, agar sistem bisa berjalan efisien di perangkat komputasi sederhana. Selain itu, penelitian Varun dkk. bersifat komparatif dengan menguji berbagai algoritma, sedangkan penelitian ini bertujuan menciptakan sistem praktis yang mudah diimplementasikan dalam skenario nyata tanpa memerlukan sumber daya komputasi tinggi.

Penelitian yang dilakukan oleh (Cheung, 2023) yang berjudul "*Face Clustering for Connection Discovery from Event Images*" membahas pendekatan

dalam mengelompokkan wajah dari sekumpulan foto yang diambil dalam suatu acara atau kegiatan. Tujuan utama penelitian ini adalah untuk menemukan hubungan atau koneksi sosial antara individu berdasarkan frekuensi kemunculan wajah mereka dalam foto yang sama. Penelitian ini menggunakan pendekatan *face clustering* berbasis model *deep learning*, *FaceNet*, untuk mengekstraksi fitur wajah dan kemudian mengelompokkan wajah-wajah yang serupa menggunakan algoritma *clustering*. Sistem tersebut memiliki persamaan dan perbedaan dengan yang diusulkan oleh peneliti. Adapun persamaannya, penelitian ini memiliki kesamaan dalam hal tujuan dasar, yaitu melakukan pengelompokan wajah secara otomatis dari kumpulan foto. Keduanya sama-sama menggunakan pendekatan *face recognition* dan *clustering* untuk mengelola data visual, serta sama-sama bertujuan untuk meningkatkan efisiensi dalam proses dokumentasi foto. Namun, perbedaannya terletak pada tujuan akhir dan konteks penerapan. Penelitian tersebut lebih berfokus pada analisis hubungan sosial berdasarkan foto acara, sedangkan penelitian ini berfokus pada pengelompokan otomatis foto dokumentasi berdasarkan identitas wajah yang terdapat pada Google Drive, dengan penerapan langsung pada sistem dokumentasi kegiatan organisasi kemahasiswaan.

Untuk menyimpulkan persamaan dan perbedaan penelitian yang serupa, berikut tabel komparasi kajian penelitian sebelumnya dan yang diusulkan:

Tabel 1.1 Komparasi Kajian Penelitian Sebelumnya dan Penelitian yang Diusulkan

No	Judul Penelitian	Nama Peneliti	Perbedaan	Hasil Penelitian
1.	<i>Automatic Photo</i>	Choo, S, O Lee, S,	Penelitian ini berfokus pada	Sistem yang dikembangkan

	<i>Classification System Based on Face Feature Extraction and Clustering</i>	Y., Seok, H., Lee, M., Lee, S., Yoo, H	sistem yang dirancang untuk membantu staff <i>daycare</i> dalam mengelola foto kegiatan harian.	berhasil mendeteksi, menyimpan, dan mengelompokkan wajah secara otomatis menggunakan OpenCV dan DBSCAN, sehingga mengurangi beban kerja staff dalam memilah foto berdasarkan individu.
2.	<i>Face Clustering with Advanced Techniques</i>	Varun, Rohit, Maheshwari, Chaitanya, Prateek	Penelitian ini menggunakan pendekatan <i>link prediction</i> dan ekstraksi fitur wajah dengan model VGG16, serta membandingkan hasil clustering menggunakan K-means dan DBSCAN. Fokusnya adalah pada akurasi dan skalabilitas	DBSCAN memberikan skor <i>silhouette</i> yang lebih baik dibanding K-means, menunjukkan efektivitas pengelompokan wajah berdasarkan kemiripan fitur dalam ruang vektor.

			pengelompokan wajah.	
3.	<i>Face Clustering for Connection Discovery from Event Images</i>	Cheung, M	Fokus utama adalah menemukan koneksi sosial antar individu dari foto-foto acara melalui <i>co-occurrence</i> wajah dalam gambar, tanpa memerlukan identitas peserta. Penelitian ini lebih menekankan pada konstruksi <i>social graph</i> .	Sistem mampu mengelompokkan lebih dari 40.000 wajah dengan F1 <i>score</i> sebesar 80% dan menghasilkan <i>social graph</i> berdasarkan frekuensi kemunculan wajah secara bersamaan dalam gambar.
Penelitian yang Diusulkan				
Judul Penelitian	Nama Peneliti dan Tahun Penelitian	Pendekatan dan Metode Penelitian yang Diusulkan		
Pengelompokan Foto Otomatis Berdasarkan Identifikasi Wajah Menggunakan Algoritma <i>Density</i>	Rafiul Muiz.K, 2025	Pendekatan penelitian yang diusulkan adalah kuantitatif, berfokus pada algoritma Density Based Spatial Clustering of Applications with Noise (DBSCAN) untuk otomatisasi mengelompokkan foto		

<i>Based Spatial Clustering of Application with Noise</i>	 The logo of Universitas Islam Negeri Alauddin Makassar features a green circular emblem with a yellow star in the center. Below the emblem, the university's name is written in large, stylized green letters: "UNIVERSITAS ISLAM NEGERI ALAUDDIN MAKASSAR".	<p>berdasarkan identitas wajah. Metode penelitian melibatkan pengumpulan data dari foto dokumentasi kegiatan dua organisasi, yaitu Inready Workgroup dan Generasi Baru Indonesia (GenBI), yang disimpan di Google Drive. Proses ini mencakup identifikasi foto yang mengandung wajah, pra-pemrosesan seperti deteksi wajah dan ekstraksi fitur menggunakan InsightFace, serta penerapan DBSCAN untuk mengelompokkan vektor fitur wajah. Evaluasi dilakukan untuk mengukur akurasi pengelompokan menggunakan metrik Silhouette Score dan Davies Bouldin Index, dan kemudahan pengguna dalam mengakses foto personal mereka.</p>
---	--	--

E. Tujuan dan Kegunaan Penelitian

1. Tujuan Penelitian

Penelitian ini bertujuan untuk mengembangkan sistem pengelompokan foto otomatis berdasarkan identifikasi wajah dengan

menggunakan algoritma DBSCAN. Sistem ini dirancang untuk:

- a. Untuk mengetahui bagaimana sistem pengelompokan foto otomatis dalam membantu pengguna mengakses foto personal mereka dengan mudah berdasarkan identitas wajah
- b. Untuk mengetahui implementasi dan kinerja algoritma DBSCAN dalam proses otomatisasi pengelompokan foto berdasarkan identitas wajah

2. Kegunaan Penelitian

Adapun kegunaan dari penelitian ini baik secara teoretis maupun secara praktis yaitu :

a. Kegunaan Teoretis

Penelitian ini memberikan kontribusi dalam pengembangan ilmu pengetahuan, khususnya mengenai *clustering* berbasis kepadatan. Penelitian ini juga memperkaya literatur akademik terkait integrasi teknologi *deep learning* dan algoritma DBSCAN dalam pengelolaan data visual. Selain itu, penelitian ini dapat menjadi acuan konseptual bagi pengembangan sistem pencarian dan pengelompokan wajah dalam konteks pengelolaan arsip digital berbasis *cloud*.

b. Kegunaan Praktis

Penelitian ini memberikan manfaat praktis bagi organisasi kemahasiswaan, terutama dalam meningkatkan efisiensi dokumentasi dan pencarian foto personal dari kumpulan dokumentasi kegiatan. Sistem yang dikembangkan memungkinkan pengguna untuk menemukan foto yang memuat wajah mereka dengan cepat dan akurat, sehingga mengurangi waktu pencarian manual. Selain itu, sistem ini dapat diadaptasi oleh instansi pendidikan, komunitas, maupun perusahaan dalam pengelolaan arsip foto secara digital dan terorganisir.

BAB II

TINJAUAN TEORETIS

A. Kajian Integrasi Keilmuan

Penelitian ini mengintegrasikan kemajuan dalam bidang teknologi informasi dengan prinsip-prinsip fundamental dalam ajaran Islam. Al-Qur'an mendorong manusia untuk menggunakan ilmu pengetahuan dan teknologi yang Allah berikan untuk memudahkan dan meningkatkan kualitas hidup, sebagaimana tersirat dalam Surah Al-Jatsiyah/45:13 sebagai berikut:

وَسَخَّرَ لَكُمْ مَا فِي السَّمَاوَاتِ وَمَا فِي الْأَرْضِ جَمِيعًا مِنْهُ إِنَّ فِي ذَلِكَ لَذِي
الْقُوَّمِ يَتَفَكَّرُونَ

Terjemahnya:

"Dia telah menundukkan (pula) untukmu apa yang ada di langit dan apa yang ada di bumi semuanya (sebagai rahmat) dari-Nya. Sesungguhnya pada yang demikian itu benar-benar terdapat tanda-tanda (kebesaran Allah) bagi kaum yang berpikir" (Kementerian Agama, 2022c)

Menurut Tafsir Ibnu Katsir (Al-Sheikh, 2004), ayat di atas menekankan bahwa Allah telah menundukkan semua yang ada di langit dan bumi untuk kepentingan manusia. Hal ini dapat dimaknai sebagai dorongan untuk memanfaatkan teknologi dan ilmu pengetahuan yang telah Allah anugerahkan untuk memudahkan kehidupan manusia, termasuk dalam pengelolaan informasi visual.

Solusi teknologi yang diusulkan dalam penelitian ini tidak hanya menjawab tantangan praktis di era digital, tetapi juga sejalan dengan panduan etis yang ditekankan dalam sunnah Nabi Muhammad saw. Penggunaan ilmu pengetahuan dan teknologi yang Allah berikan untuk memudahkan dan meningkatkan kualitas hidup mendorong umat manusia untuk memanfaatkan sumber daya yang ada, khususnya yang berkaitan dengan sumber daya

berharga namun seringkali terabaikan, yaitu waktu. Salah satu tantangan utama yang diidentifikasi dalam penelitian ini adalah proses pencarian foto personal secara manual yang sangat tidak efisien dan memakan waktu. Aktivitas ini, meskipun terlihat sepele, merupakan cerminan dari bagaimana waktu luang dapat terbuang untuk kegiatan yang kurang produktif. Hal ini sangat relevan dengan hadis yang diriwayatkan oleh Ibnu Abbas ra. di mana Rasulullah saw. bersabda:

عَنْ أَبْنَى عَبَّاسٍ رَضِيَ اللَّهُ عَنْهُمَا قَالَ: قَالَ رَسُولُ اللَّهِ صَلَّى اللَّهُ عَلَيْهِ وَسَلَّمَ: نِعْمَتَانِ مَغْبُونُ فِيهِمَا كَثِيرٌ مِنَ النَّاسِ: الصَّحَّةُ وَالْفَرَاغُ

Artinya:

"Dari Ibnu 'Abbas radhiyallahu 'anhuma berkata: Rasulullah shallallahu 'alaihi wa sallam bersabda: 'Ada dua nikmat yang banyak dilalaikan manusia: kesehatan dan waktu luang'." (Al-Bukhari, 1997)

Hadir ini menegaskan bahwa waktu luang (*al-faragh*) adalah sebuah nikmat besar yang seringkali disia-siakan. Kata *maghbun* (tertipu atau merugi) menunjukkan kondisi di mana seseorang tidak menyadari nilai sebenarnya dari nikmat tersebut hingga ia kehilangannya (Fauzan & Zainarti, 2025). Dalam konteks penelitian ini, waktu yang dihabiskan berjam-jam untuk menyortir ratusan hingga ribuan foto adalah wujud nyata dari kerugian dalam memanfaatkan nikmat waktu luang.

Oleh karena itu, pengembangan sistem dalam penelitian ini dapat dipandang sebagai sebuah ikhtiar atau usaha intelektual untuk mengamalkan pesan hadis ini. Dengan mengotomatisasi proses identifikasi dan pengelompokan foto, sistem yang diusulkan bertujuan untuk membebaskan pengguna dari tugas yang repetitif dan tidak efisien, sehingga dapat mengurangi waktu pencarian secara signifikan.

Secara esensial, penelitian ini berupaya mengubah kondisi *maghbun* (merugi) menjadi kondisi yang lebih produktif dan bernilai. Waktu yang

berhasil dihemat dapat dialokasikan oleh pengguna untuk aktivitas lain yang lebih bermanfaat, baik untuk pengembangan diri, ibadah, maupun interaksi sosial. Dengan demikian, penelitian ini tidak hanya memberikan kontribusi teknologi dalam pengelolaan data visual, tetapi juga berfungsi sebagai alat bantu praktis bagi umat Muslim untuk lebih menghargai dan mengoptimalkan nikmat waktu luang yang telah dianugerahkan oleh Allah Swt.

B. Pengelompokan Foto

Foto adalah hasil dari proses fotografi, yaitu teknik merekam cahaya pada permukaan sensitif seperti film atau sensor digital untuk menghasilkan gambar visual dari objek nyata (Yunianto, 2021). Dalam konteks modern, foto tidak hanya berfungsi sebagai dokumentasi visual, tetapi juga sebagai media komunikasi, ekspresi artistik, dan menyampaikan informasi. Roland Barthes dalam *Camera Lucida* menegaskan bahwa foto memiliki kekuatan untuk merekam momen yang pernah ada, bukan sekadar menampilkan apa yang sudah tiada, melainkan menghadirkan kembali peristiwa yang telah berlangsung ke dalam bentuk visual yang dapat diinterpretasikan oleh siapa saja (Sholichin, 2022). Dari sisi teknis, perkembangan kamera digital dan smartphone telah mendemokratisasi proses pengambilan foto, sehingga siapa pun dapat menghasilkan dan menyebarkan foto dengan mudah. Hal ini menyebabkan ledakan data visual di era digital, di mana foto menjadi salah satu bentuk data yang paling banyak dihasilkan dan dibagikan setiap hari.

Pengelompokan foto (*image clustering*) adalah proses mengorganisasi kumpulan foto ke dalam kelompok-kelompok (*cluster*) berdasarkan kemiripan fitur visual, objek, atau identitas tertentu yang terdapat di dalam foto tersebut. Secara umum, pengelompokan foto merupakan bagian dari teknik *unsupervised learning* dalam ilmu komputer dan kecerdasan buatan,

yang bertujuan menemukan pola atau kategori secara otomatis tanpa label awal (Zhang & Peng, 2024). Prinsip dasar pengelompokan foto dimulai dari proses representasi fitur, yaitu ekstraksi karakteristik visual dari setiap foto yang akan dikelompokkan. Fitur yang diekstrak bisa berupa warna, tekstur, bentuk, hingga fitur semantik seperti wajah manusia. Dalam pengelompokan foto berbasis wajah, fitur utama yang digunakan adalah *embedding* vektor dari citra wajah yang dihasilkan oleh model deep learning (Bjerge dkk., 2025). Setelah fitur berhasil diekstraksi, proses selanjutnya adalah pengelompokan (*clustering*) menggunakan algoritma DBSCAN. DBSCAN bekerja dengan melakukan pengelompokan berdasarkan kepadatan data dan unggul dalam menangani noise serta bentuk *cluster* yang tidak beraturan (Simovici, 2022).

Proses pengelompokan foto secara umum meliputi beberapa tahapan, yaitu pra-pemrosesan gambar, ekstraksi fitur, pemrosesan fitur, proses *clustering*, dan validasi hasil (Bjerge dkk., 2025). Tahap pra-pemrosesan bertujuan untuk menyiapkan foto agar fitur yang dihasilkan lebih optimal. Selanjutnya, fitur yang telah diekstraksi diproses dan digunakan sebagai dasar dalam proses *clustering* untuk menentukan kelompok foto berdasarkan kemiripan fitur yang dimiliki. Salah satu penerapan penting dari pengelompokan foto adalah sistem pengelompokan berbasis wajah, yang dalam beberapa tahun terakhir menjadi solusi efektif dalam mengelola koleksi foto besar, seperti pada galeri digital atau dokumentasi kegiatan organisasi. Sistem ini bekerja dengan mendeteksi wajah di setiap foto, mengekstrak fitur wajah, lalu mengelompokkan foto-foto berdasarkan kemiripan fitur tersebut menggunakan algoritma *clustering* DBSCAN (Zhang & Peng, 2024).

C. Face Recognition

Face recognition atau pengenalan wajah adalah proses identifikasi atau

verifikasi identitas seseorang berdasarkan fitur wajah yang terekstrak dari citra digital. Teknologi ini termasuk dalam kategori biometrik karena memanfaatkan karakteristik unik pada wajah manusia, seperti bentuk mata, hidung, mulut, dan struktur tulang pipi, untuk membedakan satu individu dari individu lainnya (Meena dkk., 2022).

Secara umum, sistem *face recognition* terdiri dari tiga tahap utama yaitu deteksi wajah (*face detection*), ekstraksi fitur (*feature extraction*), dan pengenalan wajah (*face recognition*) (Meena dkk., 2022). Pada tahap deteksi wajah, sistem mengidentifikasi area pada citra yang mengandung wajah. Selanjutnya, pada tahap ekstraksi fitur, sistem mengambil ciri-ciri penting dari wajah, baik secara global (keseluruhan wajah) maupun lokal (bagian-bagian seperti mata, hidung, mulut). Tahap akhir adalah pengenalan, di mana fitur yang telah diekstrak dibandingkan dengan *database* wajah yang telah ada untuk menentukan identitas atau menemukan kemiripan.

Metode *face recognition* dapat diklasifikasikan menjadi tiga kelompok utama, yaitu *appearance-based methods*, *feature-based methods*, dan *hybrid methods* (Meena dkk., 2022). *Appearance-based methods* menggunakan seluruh citra wajah sebagai input, di mana teknik seperti *Principal Component Analysis* (PCA) dan *Linear Discriminant Analysis* (LDA) sering digunakan untuk mengenali pola dari keseluruhan tampilan wajah. Sementara itu, *feature-based methods* berfokus pada ekstraksi fitur lokal yang khas dari wajah, seperti jarak antar mata, bentuk hidung, dan kontur wajah. Adapun *hybrid methods* merupakan pendekatan yang menggabungkan keunggulan dari kedua metode sebelumnya, yaitu *appearance-based* dan *feature-based*, guna meningkatkan tingkat akurasi dalam proses pengenalan wajah.

Perkembangan pesat dalam bidang deep learning, khususnya dengan

hadirnya *Convolutional Neural Networks* (CNN), telah merevolusi sistem *face recognition* dalam satu dekade terakhir (O'Toole & Castillo, 2021). CNN memungkinkan ekstraksi fitur wajah secara otomatis dan lebih *robust* terhadap variasi pose, pencahayaan, ekspresi, bahkan penggunaan masker. Pustaka seperti InsightFace telah menjadi standar dalam penelitian dan aplikasi pengenalan wajah modern, karena kemampuannya mengekstrak *embedding* fitur wajah yang sangat representatif dan dapat digunakan dalam berbagai kondisi nyata (Saragih & To, 2022).

Deep learning juga memungkinkan pengenalan wajah pada gambar "*in the wild*" (gambar dengan kondisi tidak terkontrol), yang sebelumnya menjadi tantangan besar bagi metode konvensional. Arsitektur *deep learning* dapat memetakan citra wajah ke dalam ruang vektor berdimensi tinggi, di mana wajah yang sama akan membentuk *cluster* yang berdekatan, sementara wajah yang berbeda akan terpisah jauh. Hal ini memungkinkan sistem untuk melakukan identifikasi dan verifikasi dengan tingkat akurasi yang sangat tinggi, bahkan melebihi kemampuan manusia dalam beberapa kasus (O'Toole & Castillo, 2021).

D. Feature Extraction

Feature extraction atau ekstraksi fitur merupakan proses fundamental dalam sistem pengolahan citra dan *machine learning*, termasuk dalam aplikasi pengenalan wajah dan pengelompokan foto otomatis. Proses ini bertujuan mengubah data mentah (*raw data*) menjadi representasi yang lebih terstruktur dan bermakna, sehingga memudahkan algoritma pembelajaran mesin untuk mengenali pola, membedakan objek, dan melakukan klasifikasi secara efisien (Kumar dkk., 2025).

Secara konseptual, fitur adalah karakteristik atau atribut penting yang

merepresentasikan informasi utama dari suatu data, baik berupa citra, suara, maupun teks (Upadhyay & Gupta, 2024). Pada citra wajah, fitur dapat berupa kontur, tekstur, sudut, jarak antar titik kunci (*landmarks*), hingga representasi vektor hasil transformasi non-linear dari model *deep learning*. Proses ekstraksi fitur tidak hanya mereduksi dimensi data, tetapi juga menghilangkan informasi yang tidak relevan atau redundant, sehingga mempercepat proses komputasi dan meningkatkan akurasi model (Datacamp, 2025).

Dalam beberapa tahun terakhir, ekstraksi fitur otomatis berbasis *deep learning*, khususnya menggunakan *Convolutional Neural Networks* (CNN), telah menjadi standar baru dalam pengolahan citra wajah (H. Liu dkk., 2024). CNN mampu secara otomatis mempelajari representasi fitur yang kompleks dan *robust* terhadap variasi pose, pencahayaan, maupun ekspresi wajah. Model seperti ArcFace telah terbukti efektif dalam menghasilkan *embedding* vektor berdimensi tetap yang merepresentasikan identitas unik setiap wajah (Nixon & Aguado, 2020). *Embedding* ini kemudian dapat digunakan untuk tugas pengenalan, verifikasi, maupun *clustering* wajah.

Dalam sistem pengenalan wajah modern, proses ekstraksi fitur memainkan peran yang sangat penting dan biasanya diawali dengan deteksi wajah pada citra, dilanjutkan dengan tahap normalisasi seperti penyesuaian ukuran dan pencahayaan. Setelah itu, citra wajah akan diproses oleh model *deep learning* untuk menghasilkan vektor fitur (*embedding*) yang merepresentasikan identitas wajah dalam ruang fitur multidimensi (Nixon & Aguado, 2020). *Embedding* ini memungkinkan wajah yang sama memiliki representasi yang berdekatan, sementara wajah yang berbeda akan berada jauh satu sama lain dalam ruang tersebut. Representasi fitur yang baik sangat krusial dalam proses *clustering*, karena kualitas dan kemampuan diskriminasi

fitur sangat menentukan keberhasilan pengelompokan wajah secara akurat (Geetha dkk., 2024).

Secara umum, *feature extraction* memiliki beberapa peran penting dalam sistem pengolahan citra dan *machine learning*, di antaranya adalah sebagai alat reduksi dimensi untuk menyederhanakan data agar lebih mudah dianalisis, peningkatan akurasi melalui representasi yang lebih informatif dan diskriminatif, efisiensi komputasi karena data yang telah diekstrak membutuhkan sumber daya lebih sedikit, serta membantu model dalam melakukan generalisasi yang lebih baik terhadap data baru, sehingga mengurangi risiko *overfitting* (Datacamp, 2025; Geetha dkk., 2024).

E. Density Based Spatial Clustering of Application With Noise

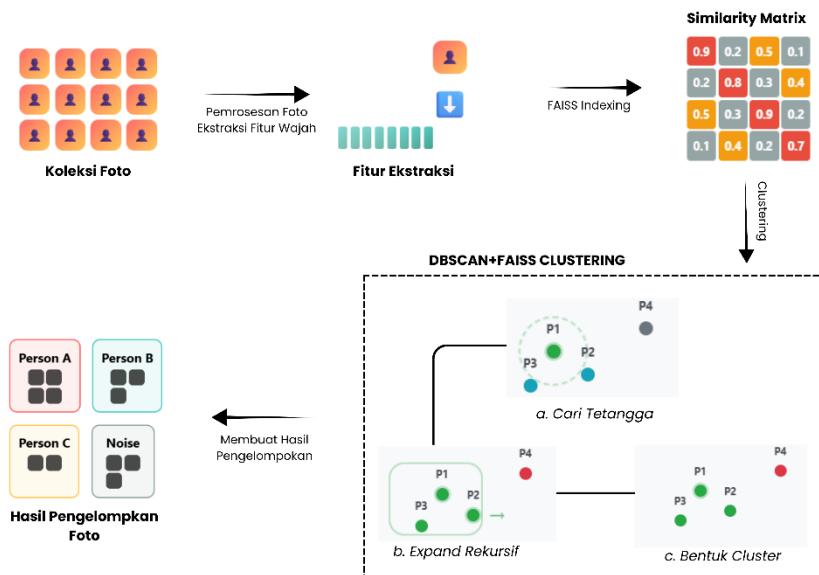
Density-Based Spatial Clustering of Applications with Noise (DBSCAN) adalah algoritma *clustering* berbasis kepadatan yang dikembangkan oleh Ester pada tahun 1996 (Daqiqil Id & Evfi Mahdiyah, 2021). DBSCAN dirancang untuk mengidentifikasi kelompok data (*cluster*) berdasarkan kepadatan titik-titik data dalam ruang fitur, serta memisahkan titik-titik yang tidak memenuhi kriteria kepadatan sebagai *noise* atau *outlier* (Indira & Kavitha Devi, 2020). Berbeda dari algoritma *clustering* konvensional seperti K-Means yang mengharuskan jumlah *cluster* ditentukan sejak awal dan cenderung mengasumsikan bentuk *cluster* bulat, DBSCAN mampu mengenali *cluster* dengan bentuk bebas serta jumlah yang tidak diketahui sebelumnya. Menggunakan pendekatan berbasis kepadatan, DBSCAN dapat membedakan titik berkepadatan rendah sebagai *noise*, sehingga memberikan hasil lebih akurat pada data yang kompleks dan mengandung derau. Keunggulan ini menjadikannya lebih tangguh daripada metode partisi, sekaligus efisien dalam

menangani data dengan beragam bentuk dan ukuran (Mustikasari & Salman, 2023).

DBSCAN bekerja dengan dua parameter utama, yaitu *epsilon* (ε) dan *minimum points* (MinPts). Parameter ε digunakan untuk menentukan radius maksimum di sekitar suatu titik dalam pencarian tetangga, sedangkan MinPts merupakan jumlah minimum titik yang harus ditemukan dalam radius tersebut agar titik tersebut dapat dikategorikan sebagai *core point*. Berdasarkan dua parameter ini, DBSCAN mengklasifikasikan titik-titik data ke dalam tiga kategori utama. Pertama, *core point*, yaitu titik yang memiliki paling sedikit MinPts tetangga dalam radius ε dan menjadi pusat dari sebuah *cluster*. Kedua, *border point*, yaitu titik yang berada dalam radius ε dari sebuah *core point* namun tidak memiliki cukup tetangga untuk menjadi *core point*. Ketiga, *noise point* atau *outlier*, yaitu titik yang tidak memenuhi kriteria sebagai *core point* maupun *border point*, sehingga tidak termasuk dalam *cluster* manapun (Indira & Kavitha Devi, 2020).

Dalam konteks pengelompokan foto berdasarkan identifikasi wajah, proses clustering DBSCAN diawali dengan tahap ekstraksi fitur wajah dari koleksi foto yang tersedia. Setiap foto yang mengandung wajah manusia diolah melalui pemrosesan foto untuk menghasilkan fitur ekstraksi berupa vektor numerik yang merepresentasikan karakteristik unik dari setiap wajah. Fitur-fitur ini kemudian digunakan untuk membangun *similarity matrix* yang menggambarkan tingkat kemiripan antar wajah dalam bentuk nilai numerik antara 0 hingga 1, dimana nilai yang mendekati 1 menunjukkan kemiripan yang tinggi antar wajah. Untuk meningkatkan efisiensi proses pencarian tetangga dalam *dataset* yang besar, implementasi DBSCAN menggunakan *Facebook AI Similarity Search* (FAISS) *indexing* sebagai mekanisme optimasi.

FAISS merupakan *library* yang dirancang khusus untuk melakukan pencarian kemiripan vektor secara efisien pada *dataset* berdimensi tinggi, sehingga dapat mempercepat proses identifikasi tetangga terdekat yang diperlukan dalam algoritma DBSCAN. Dengan menggunakan FAISS *indexing*, sistem dapat dengan cepat menentukan tetangga-tetangga dari setiap titik data dalam radius ϵ yang telah ditentukan, tanpa harus melakukan perhitungan jarak secara *brute force* terhadap seluruh *dataset*.



Gambar 2. 1 Alur Kerja *Clustering* DBSCAN dengan FAISS

Proses *clustering* kemudian dilakukan melalui tiga tahap utama dalam DBSCAN berdasarkan pada gambar 2.1. Tahap pertama adalah pencarian tetangga (*neighbor search*), dimana sistem menggunakan FAISS untuk mengidentifikasi semua titik yang berada dalam radius ϵ dari setiap titik data. Tahap kedua adalah ekspansi rekursif (*recursive expansion*), dimana sistem melakukan perluasan *cluster* secara rekursif dengan menambahkan tetangga-tetangga dari *core points* ke dalam *cluster* yang sama. Tahap ketiga adalah pembentukan *cluster* (*cluster formulation*), dimana sistem mengelompokkan titik-titik data berdasarkan keterhubungan mereka melalui *core points* dan

border points. Hasil akhir dari proses ini adalah pengelompokan foto yang menghasilkan beberapa kategori, yaitu *Person A*, *Person B*, dan *Person C* yang masing-masing berisi foto-foto dengan wajah yang sama, serta kategori *Noise* yang berisi foto-foto yang tidak dapat dikategorikan ke dalam kelompok manapun karena tidak memenuhi kriteria kepadatan yang ditetapkan. Pendekatan ini memungkinkan sistem untuk secara otomatis mengorganisir koleksi foto berdasarkan identitas wajah tanpa memerlukan informasi sebelumnya mengenai jumlah atau identitas orang yang terdapat dalam koleksi foto tersebut.

F. Facebook AI Similarity Search

Facebook AI Similarity Search atau FAISS adalah sebuah pustaka perangkat lunak *open-source* yang dikembangkan oleh Facebook AI Research untuk melakukan pencarian kemiripan vektor secara cepat dan efisien, khususnya dalam data berdimensi tinggi. Pustaka ini dirancang untuk menangani tugas-tugas pencarian tetangga terdekat (*nearest neighbor search*) dalam skala besar, baik pada CPU maupun GPU. FAISS sangat berguna dalam berbagai aplikasi berbasis *machine learning* dan kecerdasan buatan, terutama yang berkaitan dengan *image retrieval*, *recommendation system*, dan *face recognition* (Schwarz dkk., 2025), termasuk dalam proses pengelompokan wajah berbasis fitur vektor seperti yang digunakan dalam penelitian ini.

Secara umum, tujuan utama dari FAISS adalah untuk mempercepat proses pencarian *approximate nearest neighbor* (ANN), yaitu pencarian tetangga terdekat yang tidak harus benar-benar akurat, tetapi cukup mendekati dengan akurasi tinggi dan waktu yang jauh lebih cepat dibanding pencarian eksak. Dalam ruang vektor berdimensi tinggi seperti *embedding* wajah dari model *deep learning*, pencarian eksak dengan metode *brute-force*

sangat tidak efisien karena memerlukan perhitungan jarak antara setiap pasangan vektor dalam *dataset*, yang jumlahnya bisa mencapai ribuan hingga jutaan. Di sinilah FAISS hadir sebagai solusi. FAISS mengimplementasikan berbagai algoritma ANN seperti *Inverted File* (IVF), *Product Quantization* (PQ), dan *Hierarchical Navigable Small World* (HNSW) yang memungkinkan pencarian dilakukan secara efisien tanpa mengorbankan akurasi secara signifikan (Y. Liu dkk., 2022).

FAISS bekerja dengan cara mengindeks seluruh vektor ke dalam struktur data tertentu untuk mempercepat pencarian (Schwarz dkk., 2025). Sebelum proses pencarian dilakukan, seluruh vektor *embedding* yang dihasilkan dari ekstraksi fitur wajah akan diindeks terlebih dahulu. Indeks ini akan digunakan untuk mempercepat pencarian tetangga terdekat dari setiap vektor wajah ketika dilakukan proses *clustering* menggunakan algoritma DBSCAN. Proses ini disebut *approximate* karena FAISS tidak mencari semua tetangga secara eksak, melainkan mengambil pendekatan heuristik dengan margin kesalahan kecil demi memperoleh kecepatan yang jauh lebih tinggi (Y. Liu dkk., 2022). Dalam praktiknya, pendekatan ini terbukti efektif, karena dalam banyak kasus, pencarian *approximate* menghasilkan hasil yang hampir identik dengan pencarian eksak, namun dengan waktu proses yang jauh lebih singkat.

Dalam konteks pengelompokan data berdimensi tinggi seperti *embedding* wajah, penggunaan FAISS menjadi sangat relevan (Schwarz dkk., 2025). *Embedding* wajah yang dihasilkan oleh model seperti ArcFace atau InsightFace memiliki dimensi yang tinggi (512 dimensi), sehingga pencarian tetangga terdekat antar vektor menjadi sangat mahal secara komputasi jika dilakukan tanpa indeks. Oleh karena itu, FAISS digunakan untuk mempercepat

proses *neighborhood search* dalam algoritma DBSCAN, yang memang mengandalkan perhitungan jarak antar titik untuk menentukan kepadatan lokal dan mengidentifikasi *cluster*. FAISS mempercepat tahap pencarian tetangga dalam radius ϵ yang merupakan bagian krusial dari algoritma DBSCAN, sehingga membuat pengelompokan menjadi jauh lebih efisien.

Selain kecepatan, keunggulan lain dari FAISS adalah skalabilitasnya. FAISS mampu menangani jutaan hingga miliaran vektor, baik di lingkungan *single-machine* maupun multi-GPU (Schwarz dkk., 2025). Hal ini sangat penting dalam sistem pengelompokan wajah otomatis yang menangani ribuan foto, karena memungkinkan sistem tetap berjalan optimal walau data bertambah besar. Dengan memanfaatkan FAISS, sistem yang dibangun dalam penelitian ini dapat mengoptimalkan performa *clustering* DBSCAN dan memberikan waktu respons yang cepat kepada pengguna, bahkan dalam pengolahan data besar secara *real-time*.

Secara keseluruhan, integrasi antara FAISS dan DBSCAN memberikan kontribusi penting dalam efisiensi dan efektivitas sistem pengelompokan foto berbasis wajah. Penggunaan FAISS menjawab tantangan utama dari proses *clustering* dalam ruang berdimensi tinggi, yaitu kecepatan pencarian tetangga. Tanpa FAISS, DBSCAN akan mengalami hambatan performa signifikan, terutama saat menangani *embedding* wajah dalam jumlah besar. Oleh karena itu, pemahaman terhadap prinsip kerja dan manfaat FAISS sangat penting untuk menjelaskan mengapa sistem yang dikembangkan dalam penelitian ini mampu mencapai efisiensi tinggi dalam proses pengelompokan wajah secara otomatis.

G. InsightFace

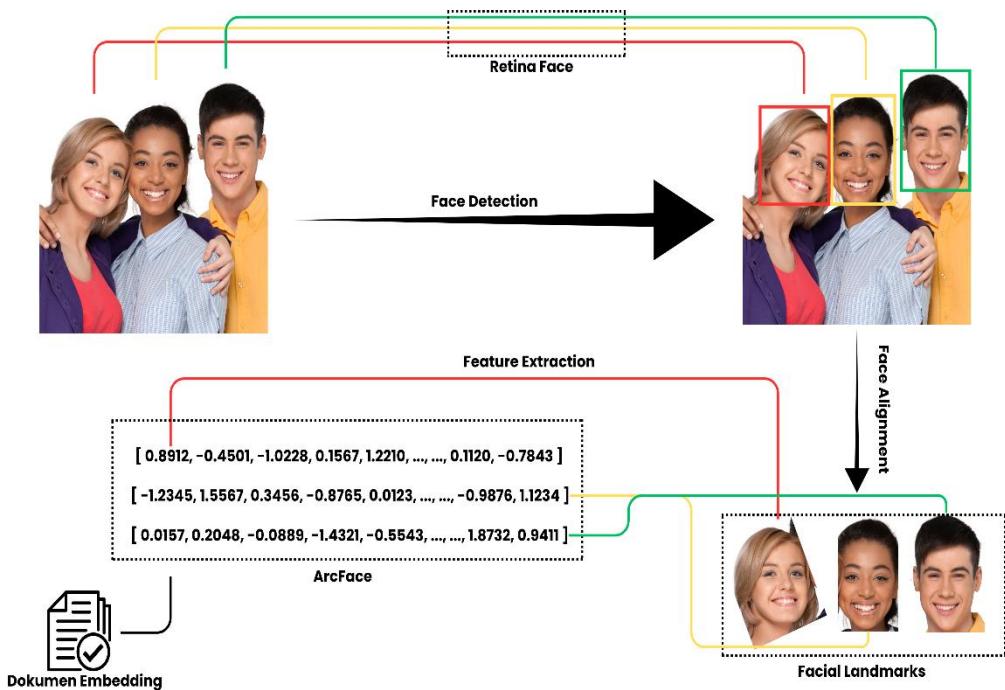
InsightFace adalah sebuah *toolbox* (perangkat) *open-source* yang

menyediakan implementasi algoritma-algoritma canggih (*state-of-the-art*) untuk analisis wajah 2D dan 3D. InsightFace bukanlah satu algoritma tunggal, melainkan sebuah pustaka yang berisi kumpulan model-model pra-pelatihan (*pre-trained*) yang sangat efisien untuk berbagai tugas terkait pengenalan wajah, seperti deteksi wajah, penyelarasan wajah, dan ekstraksi fitur. Karena performanya yang tinggi dan kemudahannya untuk diimplementasikan, InsightFace sering dijadikan pilihan utama dalam penelitian dan pengembangan aplikasi yang memerlukan identifikasi wajah secara akurat (Deng dkk., 2024). Dalam konteks penelitian mengenai pengelompokan foto otomatis, InsightFace berperan sebagai fondasi utama untuk mengubah data visual (gambar wajah) menjadi representasi numerik yang dapat diproses oleh algoritma *clustering*.

Proses identifikasi wajah hingga tahap ekstraksi fitur menggunakan *toolbox* InsightFace, sebagaimana diilustrasikan pada gambar 2.2, merupakan sebuah alur kerja yang sistematis. Proses ini dimulai dari gambar mentah yang berisi satu atau lebih individu, kemudian melalui serangkaian tahapan komputasi untuk menghasilkan vektor fitur (dikenal juga sebagai *embedding*) yang unik untuk setiap wajah. Vektor fitur inilah yang menjadi esensi dari identitas seseorang dalam bentuk data numerik, yang nantinya akan digunakan sebagai masukan untuk algoritma *clustering* seperti DBSCAN untuk melakukan pengelompokan.

Tahap pertama dalam alur kerja ini adalah Deteksi Wajah (*Face Detection*). Pada tahap ini, model yang sering digunakan dari InsightFace adalah RetinaFace. RetinaFace adalah sebuah algoritma deteksi wajah berbasis *deep learning* yang bertugas untuk memindai keseluruhan gambar dan melokalisasi keberadaan setiap wajah manusia di dalamnya. Hasil dari tahap

ini adalah serangkaian koordinat kotak pembatas (*bounding box*) yang secara presisi mengelilingi setiap wajah yang terdeteksi. Keunggulan RetinaFace terletak pada kemampuannya untuk mendeteksi wajah dengan berbagai skala, pose, dan bahkan dalam kondisi yang sedikit terhalang (*occlusion*).



Gambar 2. 2 Alur Kerja InsightFace

Setelah wajah-wajah berhasil dideteksi, tahap selanjutnya adalah Penyelarasan Wajah (*Face Alignment*). Proses ini sangat penting untuk menstandardisasi pose wajah sebelum dilakukan ekstraksi fitur. Tahap ini bekerja dengan cara mengidentifikasi titik-titik kunci pada wajah yang disebut Penanda Wajah (*Facial Landmarks*), seperti sudut mata, ujung hidung, dan sudut bibir. Berdasarkan posisi *landmarks* ini, setiap gambar wajah yang berada di dalam *bounding box* akan melalui transformasi geometris (seperti rotasi dan penskalaan) untuk memastikan wajah berada dalam posisi yang tegak dan terstandardisasi (misalnya, kedua mata berada pada garis horizontal yang sama). Proses penyelarasan ini bertujuan untuk mengurangi variasi pose

sehingga model ekstraksi fitur dapat bekerja lebih konsisten dan akurat.

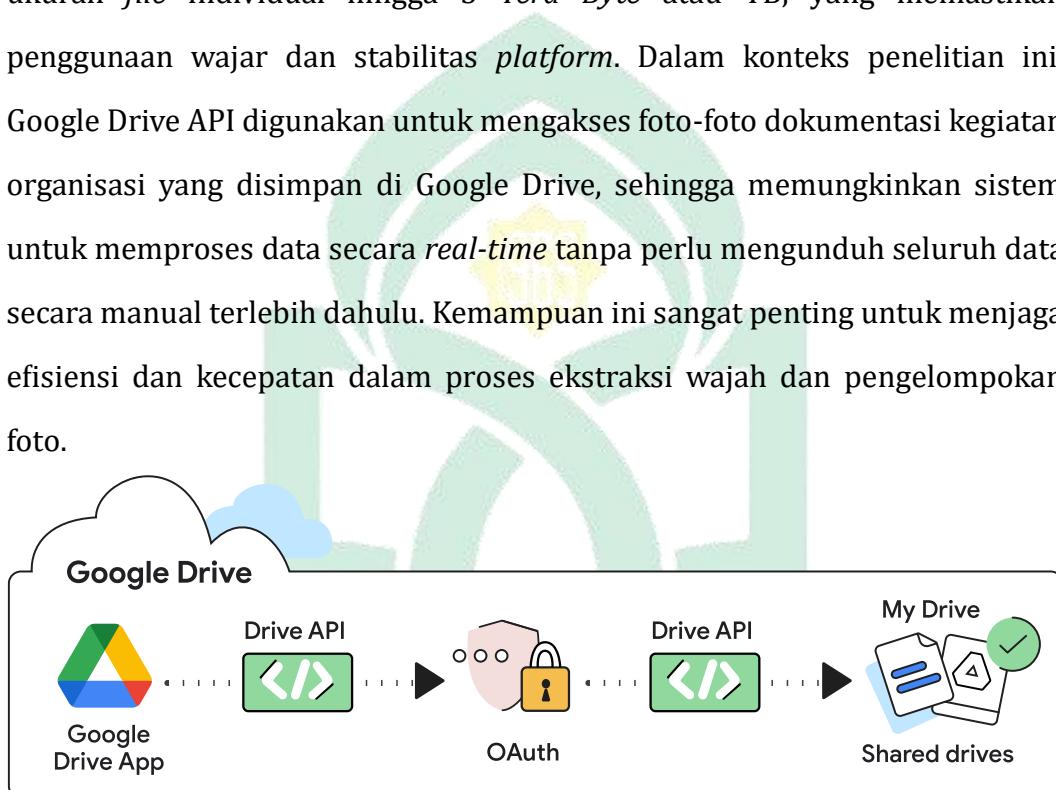
Tahap terakhir dalam proses ini adalah Ekstraksi Fitur (*Feature Extraction*). Gambar wajah yang telah diselaraskan kemudian dimasukkan ke dalam model jaringan syaraf tiruan konvolusional CNN yang mendalam, seperti ArcFace, yang juga merupakan bagian dari *toolbox* InsightFace. ArcFace bertugas untuk mengubah citra wajah 2D menjadi sebuah vektor numerik berdimensi tinggi (512 dimensi) yang disebut *embedding* Wajah. Vektor ini secara matematis merepresentasikan fitur-fitur unik dan diskriminatif dari wajah tersebut. Wajah dari orang yang sama akan menghasilkan vektor fitur yang memiliki jarak sangat dekat satu sama lain dalam ruang fitur, sementara wajah dari orang yang berbeda akan menghasilkan vektor yang berjauhan. Vektor-vektor fitur inilah yang kemudian menjadi representasi final dari setiap wajah, siap untuk dianalisis dan dikelompokkan pada tahap selanjutnya menggunakan algoritma DBSCAN.

H. Google Drive API

Google Drive merupakan layanan penyimpanan berbasis *cloud* yang disediakan oleh Google dan memungkinkan pengguna untuk menyimpan, mengakses, dan membagikan file secara daring. Fitur utama dari Google Drive adalah penyimpanan berbasis pengguna (*My Drive*) dan *folder* kolaboratif yang dikenal sebagai *Shared Drives* (Google Workspace, 2025). Kedua fitur ini memungkinkan pengguna untuk menyimpan ribuan foto hasil dokumentasi, yang dapat diakses oleh sistem pihak ketiga secara fleksibel menggunakan Google Drive API.

Google Drive API adalah sebuah antarmuka *RESTful* yang memungkinkan pengembang aplikasi untuk berinteraksi langsung dengan layanan Google Drive melalui permintaan HTTP standar. Melalui API ini, sistem

yang dikembangkan dapat mengakses file, *folder*, *metadata*, dan bahkan melakukan manipulasi seperti mengunggah, mengunduh, atau menghapus file yang tersimpan di akun pengguna (Sholva & Novriando, 2022). Meskipun sangat fleksibel, penting untuk memperhatikan batasan yang ada, seperti kuota unggahan harian sebesar 750 *Giga Byte* atau GB per pengguna dan batas ukuran *file* individual hingga 5 *Tera Byte* atau TB, yang memastikan penggunaan wajar dan stabilitas *platform*. Dalam konteks penelitian ini, Google Drive API digunakan untuk mengakses foto-foto dokumentasi kegiatan organisasi yang disimpan di Google Drive, sehingga memungkinkan sistem untuk memproses data secara *real-time* tanpa perlu mengunduh seluruh data secara manual terlebih dahulu. Kemampuan ini sangat penting untuk menjaga efisiensi dan kecepatan dalam proses ekstraksi wajah dan pengelompokan foto.



Gambar 2. 3 Google Drive API (Sumber: (Google Workspace, 2025))

Untuk mengakses data Google Drive secara aman, sistem perlu melakukan proses otentikasi dan otorisasi menggunakan protokol *OAuth* 2.0. Dalam proses ini, pengguna akan diminta memberikan izin akses kepada aplikasi pihak ketiga untuk membaca dan memanipulasi data mereka di Google Drive. Setelah berhasil diotorisasi, sistem akan menerima token akses yang dapat digunakan untuk melakukan permintaan API secara sah. Mekanisme ini menjaga keamanan data pengguna, sekaligus memungkinkan integrasi

langsung antara sistem dan akun Google Drive pengguna. Dengan menerapkan *OAuth*, sistem dapat mengakses hanya file yang relevan, seperti *folder* tertentu yang berisi dokumentasi kegiatan, tanpa perlu mengambil seluruh isi Google Drive pengguna (Google Workspace, 2025).

Integrasi Google Drive API dengan sistem pengelompokan foto juga memungkinkan penghematan sumber daya penyimpanan lokal. Karena file tidak perlu diunduh secara penuh dan disimpan di server, sistem hanya mengambil file secara selektif sesuai kebutuhan pemrosesan. Hal ini memberikan keuntungan dalam aspek efisiensi penyimpanan dan *bandwidth*, terutama bila sistem dijalankan pada perangkat dengan sumber daya terbatas. Selain itu, pendekatan ini juga meningkatkan kinerja sistem secara keseluruhan karena memungkinkan pemrosesan dilakukan secara terdistribusi, langsung dari sumber data utama di *cloud* (Sholva & Novriando, 2022).

Secara keseluruhan, penggunaan Google Drive API dalam sistem pengelompokan foto otomatis memberikan landasan teknologi yang kuat untuk menciptakan solusi yang *real-time*, efisien, dan terintegrasi. Dengan memanfaatkan fitur pencarian, akses selektif, serta otentikasi yang aman, sistem tidak hanya mampu mengakses data dokumentasi kegiatan secara otomatis, tetapi juga mampu beroperasi secara berkelanjutan seiring bertambahnya data baru. Oleh karena itu, Google Drive API tidak hanya berperan sebagai penyedia data, tetapi juga sebagai komponen kunci dalam arsitektur sistem yang memungkinkan fungsionalitas pintar dan dinamis dalam pengelolaan foto berbasis *cloud*.

BAB III

METODOLOGI PENELITIAN

A. Jenis Penelitian

Penelitian kuantitatif adalah pendekatan ilmiah yang berfokus pada pengumpulan dan analisis data numerik untuk menjelaskan, memprediksi, atau menguji hubungan antar variabel secara objektif (Wajdi dkk., 2024). Data yang dikumpulkan biasanya berupa angka atau statistik, dan dianalisis menggunakan metode statistik agar hasilnya dapat diuji serta digeneralisasi ke populasi yang lebih luas. Penelitian ini sering menggunakan instrumen seperti survei, kuisioner, atau eksperimen, dan menekankan pada objektivitas, keterukuran, serta penggunaan sampel yang mewakili populasi. (Rohman dkk., 2023).

Dalam konteks penelitian ini, jenis penelitian kuantitatif digunakan untuk mengukur kinerja algoritma DBSCAN dalam mengelompokkan foto berdasarkan identitas wajah. Data yang dianalisis berupa waktu komputasi dan kualitas *clustering*. Analisis statistik terhadap data tersebut memungkinkan penilaian objektif terhadap efektivitas dan efisiensi algoritma, sehingga solusi yang diusulkan dapat dievaluasi secara ilmiah dan terukur.

B. Pendekatan Penelitian

Penelitian ini menggunakan pendekatan kuantitatif untuk mengevaluasi kinerja sistem algoritma DBSCAN dalam mengelompokkan foto berdasarkan identitas wajah. Pendekatan ini mengutamakan pengumpulan dan analisis data numerik untuk memberikan gambaran yang objektif mengenai efektivitas dan efisiensi algoritma tersebut. Data yang diperoleh berupa hasil pengelompokan foto yang diuji menggunakan metrik performa

Silhouette Score dan *Davies Bouldin Index*. Dengan menggunakan metode statistik, hasil-hasil ini dapat dianalisis secara terukur, memungkinkan evaluasi yang lebih mendalam dan pengambilan keputusan berbasis data.

C. Sumber Data

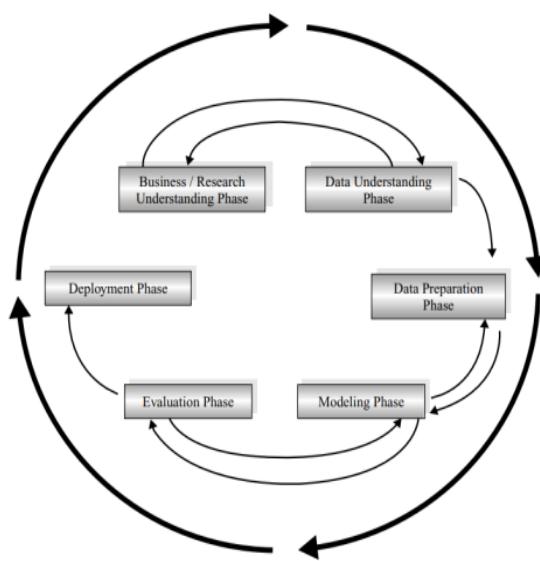
Sumber data yang digunakan dalam penelitian ini adalah kumpulan foto dokumentasi kegiatan yang diambil dari dua organisasi, yaitu Inready Workgroup dan Generasi Baru Indonesia (GenBI). Foto-foto ini disimpan dalam format digital seperti JPEG, JPG, PNG, dan HEIC, dan diakses melalui platform penyimpanan *cloud*, yaitu Google Drive. Penggunaan Google Drive sebagai sumber data memungkinkan akses yang mudah dan efisien terhadap koleksi foto yang besar, serta memfasilitasi kolaborasi antara anggota organisasi dalam mengelola dokumentasi kegiatan. Dengan memanfaatkan *link* Google Drive, peneliti dapat mengumpulkan data foto yang relevan untuk analisis lebih lanjut dalam pengelompokan otomatis berdasarkan identifikasi wajah.

D. Metode Pengumpulan dan Pengolahan Data

Metode pengumpulan data dalam penelitian ini dilakukan dengan cara mengumpulkan tautan Google Drive yang berisi foto-foto dokumentasi kegiatan dari kedua organisasi. Proses ini melibatkan pengidentifikasi dan pemilihan foto yang mengandung wajah manusia, yang kemudian akan digunakan dalam analisis.

Penelitian ini mengadopsi kerangka kerja *Cross-Industry Standard Process for Data Mining* atau CRISP-DM sebagai panduan sistematis dalam proses pengolahan data. CRISP-DM adalah kerangka kerja yang menawarkan pendekatan terstruktur, sistematik, dan terstandarisasi untuk menjalankan

proyek dari awal hingga akhir. Kerangka kerja ini dirancang untuk menjadi industri-agnostik, artinya dapat diterapkan di berbagai sektor. CRISP-DM terdiri dari enam tahap utama (*business understanding, data understanding, data preparation, modeling, evaluation, deployment*) (Binus University, 2025), yang relevan dengan alur kerja penelitian ini.



Gambar 3. 1 Kerangka Kerja CRISP-DM (Zebua dkk., 2022)

E. Instrumen Penelitian

1. Perangkat Keras

Perangkat keras (*Hardware*) sebagai alat yang digunakan untuk mengumpulkan dan mengembangkan data yang telah didapatkan adalah laptop Asus TUF Gaming A15 dengan spesifikasi sebagai berikut.

- Processor AMD Ryzen™ 5 4600H Mobile Processor (6 core/12-thread, 11MB Cache, 4.0 GHz max boost)*
- RAM DDR 8 GB
- SSD 512 GB

2. Perangkat Lunak

- a. Bahasa Pemrograman Python
- b. *Browser* Google Chrome
- c. *Code Editor Visual Studio Code*
- d. Sistem Operasi Windows

F. Teknik Pengolahan dan Analisis Data

1. Teknik Pengolahan Data

a. Business Understanding

Tahap ini terkait dengan pemahaman tujuan penelitian, yaitu mengembangkan sistem pengelompokan foto otomatis berbasis wajah. Fokusnya adalah mendefinisikan kebutuhan data (foto dokumentasi kegiatan organisasi) dan kriteria kualitas data (foto harus mengandung wajah manusia yang jelas).

b. Data Understanding

Pada tahap ini, mengidentifikasi sumber data (tautan Google Drive organisasi) dan melakukan eksplorasi awal terhadap karakteristik data. Proses pengumpulan tautan dan identifikasi foto yang mengandung wajah manusia termasuk dalam fase ini. Analisis awal bertujuan untuk memastikan kelayakan data (ketersediaan wajah, variasi pencahayaan, resolusi gambar).

c. Data Preparation

Persiapan data merupakan tahap di mana foto-foto yang telah dikumpulkan dipersiapkan untuk analisis lebih lanjut. Dalam hal ini, data diproses melalui beberapa langkah pra-pemrosesan, termasuk deteksi wajah dengan menggunakan InsightFace. Proses ini melibatkan identifikasi wajah pada gambar dan pemotongan untuk menyisakan hanya area wajah, guna

mengurangi *noise* dan meningkatkan akurasi analisis. Selain itu, data juga dipersiapkan dengan menyesuaikan ukuran dan pencahayaan foto agar memenuhi standar untuk diproses oleh model *deep learning*.

d. Modeling

Pada tahap *modeling*, gambar-gambar wajah yang telah dipotong dan dinormalisasi digunakan untuk ekstraksi fitur menggunakan *library deep learning*, dalam hal ini InsightFace. InsightFace akan menghasilkan vektor *embedding* yang merepresentasikan ciri khas dari setiap wajah. Vektor-vektor ini kemudian digunakan untuk mengelompokkan foto-foto berdasarkan kemiripan wajah menggunakan algoritma DBSCAN. DBSCAN akan mengelompokkan wajah yang memiliki kemiripan tinggi dalam satu kelompok, sehingga memudahkan identifikasi individu dalam gambar.

e. Evaluation

Evaluation adalah tahap untuk menilai sejauh mana model yang telah dibangun dapat menghasilkan hasil yang sesuai dengan tujuan analisis. Pada tahap ini, hasil pengelompokan wajah yang dilakukan oleh DBSCAN dievaluasi dengan memeriksa akurasi dan keakuratan pengelompokan. Ini termasuk memverifikasi apakah wajah yang sama dikelompokkan bersama, sementara wajah yang berbeda dipisahkan dengan benar. Evaluasi ini sangat penting untuk memastikan bahwa model dapat diterapkan secara efektif untuk pencarian foto berdasarkan individu.

f. Deployment

Deployment atau implementasi adalah langkah akhir di mana hasil analisis diterapkan pada konteks dunia nyata. Dalam hal ini, hasil pengelompokan wajah dapat digunakan untuk mempermudah pencarian foto berdasarkan identitas individu. Setelah proses pengelompokan yang efektif,

data foto dapat diorganisasi dalam sistem yang memungkinkan pengguna untuk dengan mudah mencari dan mengelompokkan foto berdasarkan wajah individu.

2. Analisis Data

Analisis data dilakukan dengan menerapkan algoritma DBSCAN pada vektor fitur wajah yang telah diekstraksi. Proses ini bertujuan untuk mengelompokkan foto berdasarkan kemiripan identitas wajah. Metrik performa *Silhouette Score* dan *Davies Bouldin Index* akan digunakan untuk mengevaluasi kinerja algoritma dalam mengelompokkan foto. Hasil analisis ini akan memberikan gambaran yang jelas mengenai efektivitas sistem yang dikembangkan dalam meningkatkan efisiensi pencarian foto personal di antara kumpulan foto.

G. Teknik Pengujian dan Evaluasi Model

1. Teknik Pengujian

Teknik pengujian yang diterapkan dalam penelitian ini yaitu teknik pengujian kuantitatif berbasis kinerja (*Performance-Based Quantitative Testing*). Pendekatan ini digunakan karena memungkinkan evaluasi sistem secara objektif tanpa melibatkan responden secara langsung. Teknik ini mengandalkan metrik proksi, yaitu alat ukur kuantitatif tidak langsung untuk menilai kualitas sistem yang bersifat konseptual seperti kemudahan dan akurasi (Junqing Tang, 2023). Metrik proksi utama yang digunakan dalam penelitian ini meliputi, waktu proses sebagai proksi untuk kecepatan dan efisiensi sistem dari perspektif pengguna, kualitas hasil algoritma yang diukur menggunakan *Silhouette Score* dan *Davies Bouldin Index* sebagai proksi untuk ketepatan dan akurasi pengelompokan, serta keberhasilan pemrosesan data sebagai proksi untuk keandalan sistem.

2. Evaluasi Model

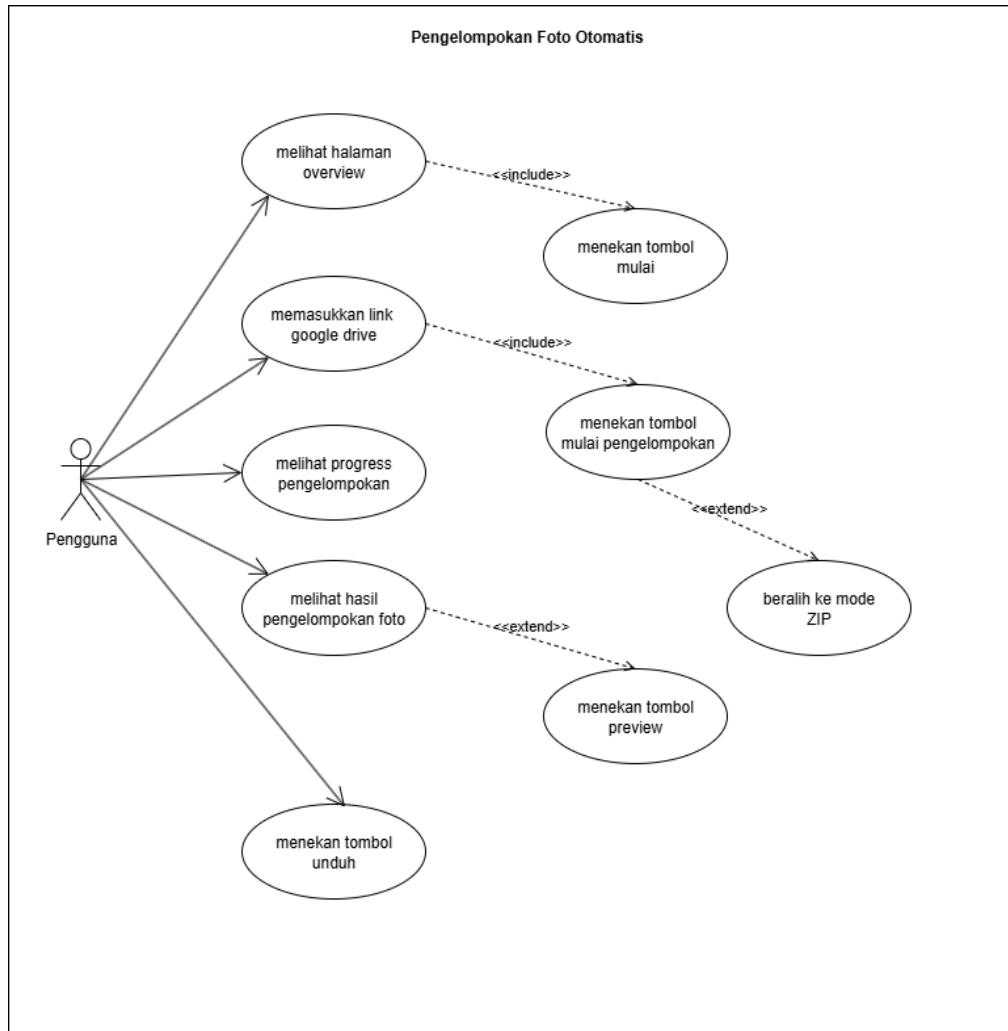
Evaluasi model dalam penelitian ini dilakukan dengan menggunakan metrik performa yang relevan untuk menilai efektivitas sistem dalam mengelompokkan foto. Metrik utama yang digunakan adalah *Silhouette Score* dan *Davies Bouldin Index* (DBI) untuk menilai kualitas *cluster* yang dihasilkan oleh algoritma DBSCAN.

Silhouette Score mengukur seberapa baik setiap foto dikelompokkan dengan mempertimbangkan kedekatannya dengan foto dalam *cluster* yang sama (kohesi) dan seberapa jauh foto tersebut dari *cluster* terdekat (separasi). Nilai *Silhouette Score* yang lebih tinggi menunjukkan bahwa objek lebih cocok dengan *clusternya* sendiri dibandingkan dengan *cluster* lain, yang menandakan kualitas pengelompokan yang baik. Di sisi lain, DBI mengevaluasi kualitas pengelompokan berdasarkan rasio antara jarak internal dalam *cluster* (*intra-cluster*) dengan jarak antar *cluster* (*inter-cluster*). Berbeda dengan *Silhouette Score*, nilai DBI yang lebih rendah menunjukkan performa *clustering* yang lebih baik. Hal ini karena nilai yang rendah mengindikasikan bahwa *cluster* yang terbentuk lebih padat (kompak) dan terpisah dengan baik satu sama lain. Penggunaan kedua metrik ini memberikan penilaian yang lebih komprehensif dan *robust* terhadap efektivitas model.

H. Perancangan Sistem

1. Use Case Diagram

Diagram *Use Case* merupakan bagian dari *Unified Modeling Language* (UML) yang berfungsi untuk memvisualisasikan interaksi antara sistem dengan penggunanya (aktor). Diagram ini menunjukkan fungsi apa saja yang dimiliki sistem dan bagaimana aktor dapat menggunakan fungsi-fungsi tersebut.



Gambar 3. 2 Use Case Diagram Sistem yang Diusulkan

Dalam penelitian ini, sistem yang dibangun berinteraksi dengan satu jenis pengguna. Pengguna dapat melakukan pengelompokan foto secara otomatis berdasarkan identifikasi wajah menggunakan algoritma DBSCAN melalui sistem ini. Terdapat beberapa alur kerja utama dalam sistem ini yaitu pertama, pengguna menginisiasi proses dengan memasukkan tautan Google Drive yang berisi kumpulan foto yang akan dianalisis. Kedua, pengguna dapat melihat hasil akhir berupa foto-foto yang telah dikelompokkan berdasarkan wajah yang teridentifikasi. Sistem juga menyediakan fungsi pendukung seperti memantau progres pengelompokan, melakukan pratinjau (*preview*) hasil,

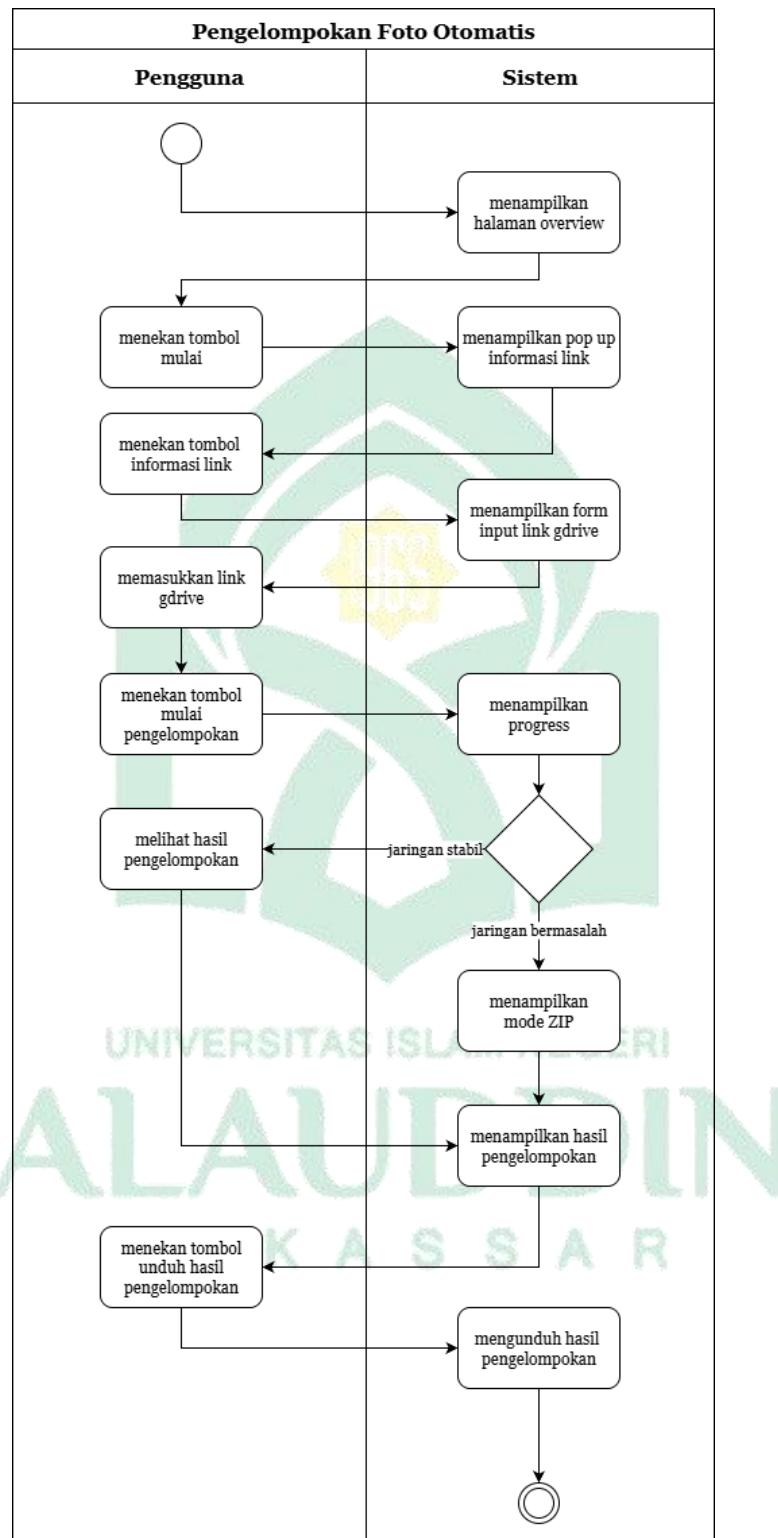
serta fitur adaptif yang menawarkan untuk beralih ke mode ZIP jika jaringan pengguna melambat untuk mempermudah proses pengunduhan. Terakhir, pengguna dapat mengunduh seluruh hasil pengelompokan tersebut.

2. Activity Diagram

Activity Diagram adalah bagian dari UML yang berfungsi memvisualisasikan alur kerja sebuah sistem. Diagram ini memetakan urutan langkah, titik keputusan, dan keterkaitan antar aktivitas dalam suatu proses.

Diagram aktivitas pada gambar 3.3, menguraikan secara rinci alur kerja sistem pengelompokan foto otomatis, yang dimulai saat pengguna masuk dan sistem menampilkan halaman *overview*. Pengguna kemudian menginisiasi proses dengan menekan tombol mulai, yang mengarahkan sistem untuk menampilkan *form* input guna meminta tautan Google Drive. Setelah pengguna memasukkan tautan dan menekan tombol mulai, sistem akan mulai bekerja dan memberikan umpan balik berupa tampilan *progress* (kemajuan). Pada tahap ini, terdapat sebuah titik keputusan kritis di mana sistem secara cerdas memeriksa stabilitas jaringan. Jika koneksi jaringan terdeteksi stabil, proses berlanjut secara normal di mana pengguna dapat langsung melihat hasil pengelompokan yang ditampilkan. Sebaliknya, jika jaringan teridentifikasi bermasalah, sistem secara adaptif akan mengaktifkan mode ZIP, sebuah mekanisme untuk mengemas hasil ke dalam satu *file* arsip demi menjamin keandalan saat diunduh, sebelum kemudian tetap menampilkan hasil pengelompokan tersebut. Terlepas dari kondisi jaringan sebelumnya, setelah hasil ditampilkan, pengguna memiliki opsi untuk menekan tombol unduh untuk menyimpan hasil pengelompokan ke perangkatnya, yang sekaligus menjadi aktivitas terakhir sebelum alur kerja ini selesai. Diagram menggambarkan proses yang lengkap dan berurutan terhadap interaksi

pengguna dan sistem dalam mengakses sistem pengelompokan foto otomatis.

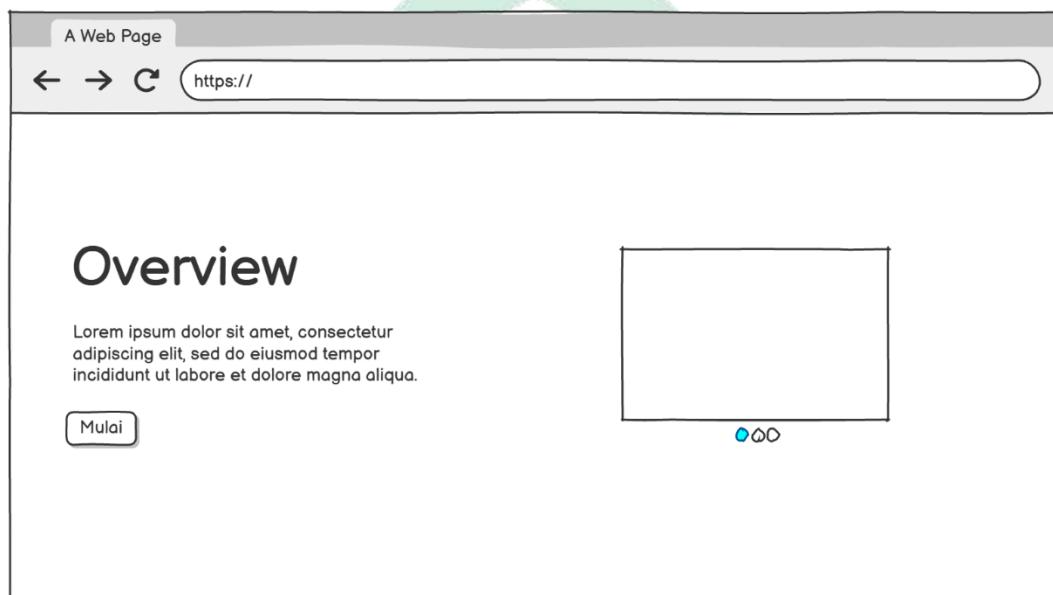


Gambar 3. 3 Activity Diagram Sistem yang Diusulkan

I. Perancangan Antarmuka

1. Halaman Utama (*Overview*)

Halaman utama merupakan gerbang awal interaksi pengguna dengan sistem. Tampilan ini dirancang untuk memberikan gambaran singkat mengenai fungsi utama sistem secara jelas dan ringkas. Terdapat tombol Mulai yang berfungsi sebagai titik awal bagi pengguna untuk memulai proses pengelompokan foto.



Gambar 3. 4 Wireframe Halaman Utama

2. Halaman Informasi Link Google Drive

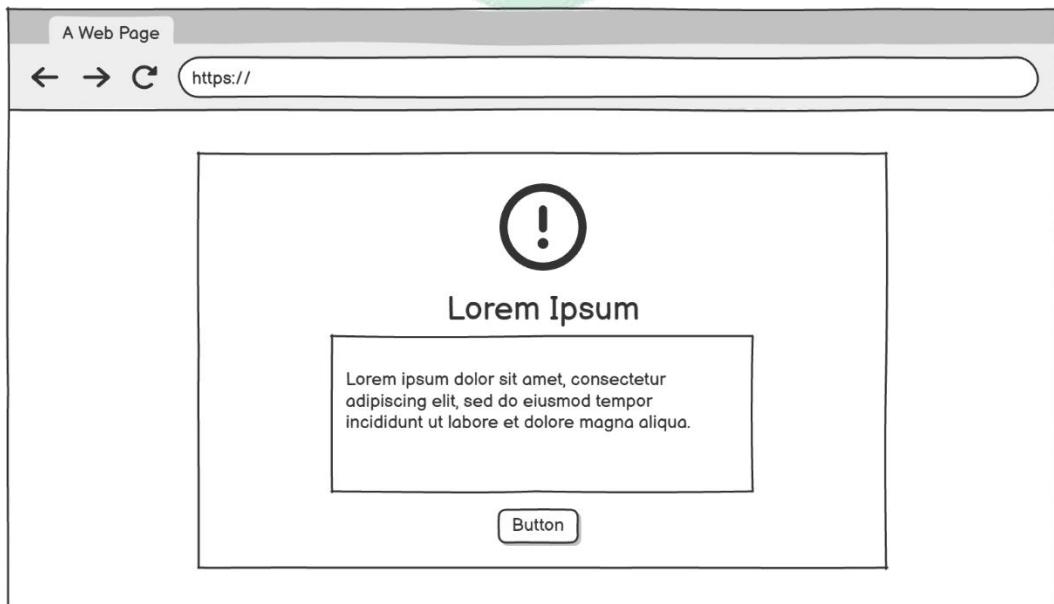
Halaman informasi link Google Drive adalah bagian yang akan memberikan informasi ke pengguna terkait status link Google Drive yang bisa diproses oleh sistem untuk melakukan pengelompokan foto otomatis. Terdapat dua tombol untuk memberikan informasi kepada pengguna. Tombol pertama sebelah kanan untuk melanjutkan aksi ketika link pengguna dalam keadaan yang sudah sesuai dengan yang tertera pada sistem. Tombol kedua sebelah kiri untuk informasi yang lebih detail terkait link Google Drive.



Gambar 3. 5 Wireframe Informasi Link Google Drive

3. Halaman Informasi Lanjutan Link Google Drive

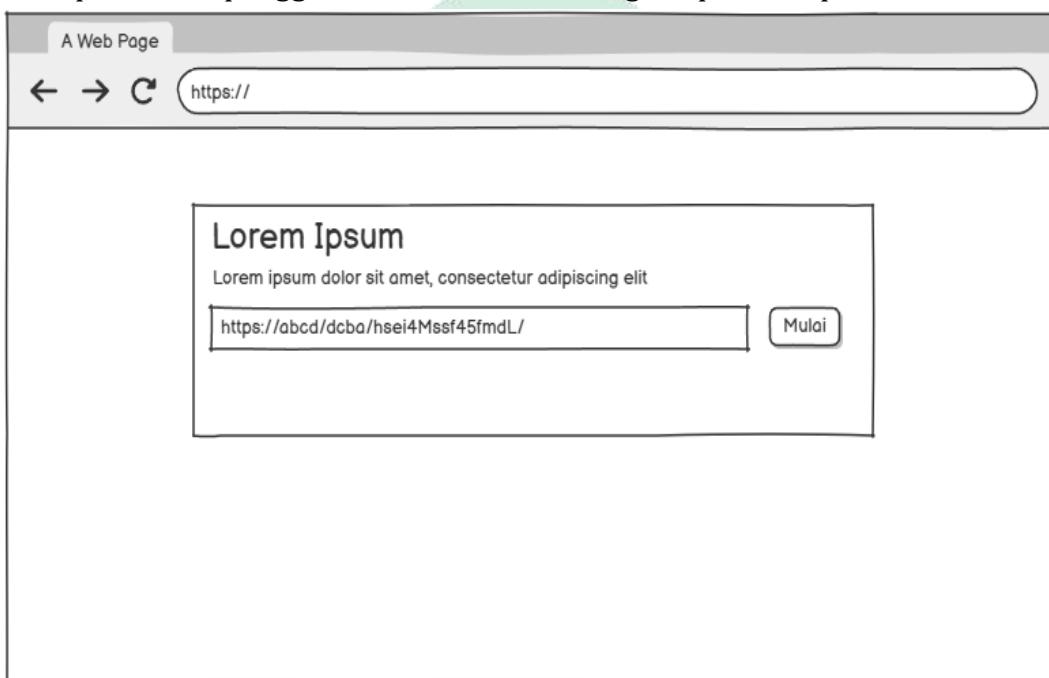
Halaman ini adalah bagian yang menjelaskan secara detail terkait ciri link yang bisa diproses oleh sistem dan ciri link yang tidak bisa diproses. Halaman ini berupa informasi dan peringatan kepada pengguna terkait link Google Drive yang akan mereka masukkan ke inputan pada sistem.



Gambar 3. 6 Wireframe Informasi Lanjutan Link Google Drive

4. Halaman Input Tautan Google Drive

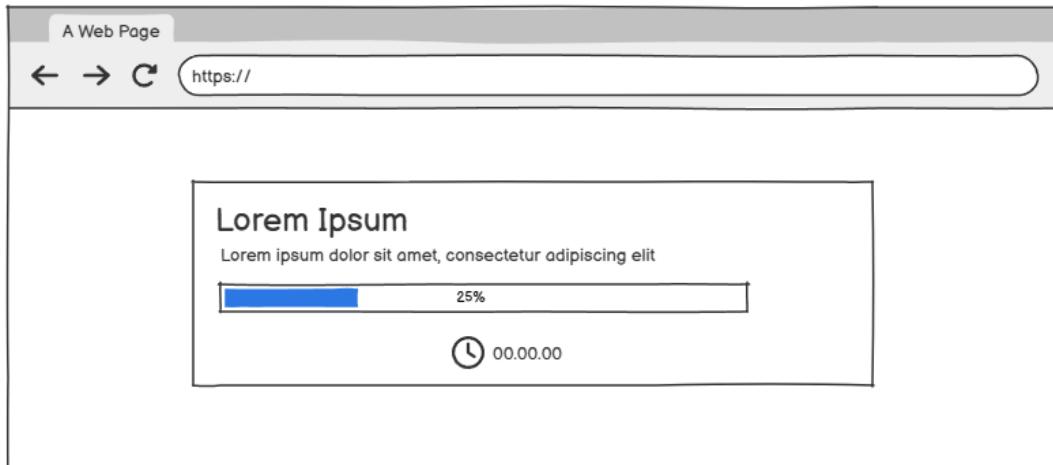
Setelah menekan tombol Mulai dan mendapatkan informasi link Google Drive, pengguna akan diarahkan ke halaman input. Pada halaman ini, terdapat sebuah kolom input yang dirancang secara khusus untuk memasukkan tautan Google Drive yang berisi kumpulan foto yang akan diproses. Desain yang terfokus pada satu tindakan ini bertujuan untuk menghilangkan distraksi dan mempermudah pengguna dalam memulai langkah pertama pemrosesan.



Gambar 3. 7 Wireframe Input Tautan Google Drive

5. Tampilan *Progress* Pengelompokan Foto

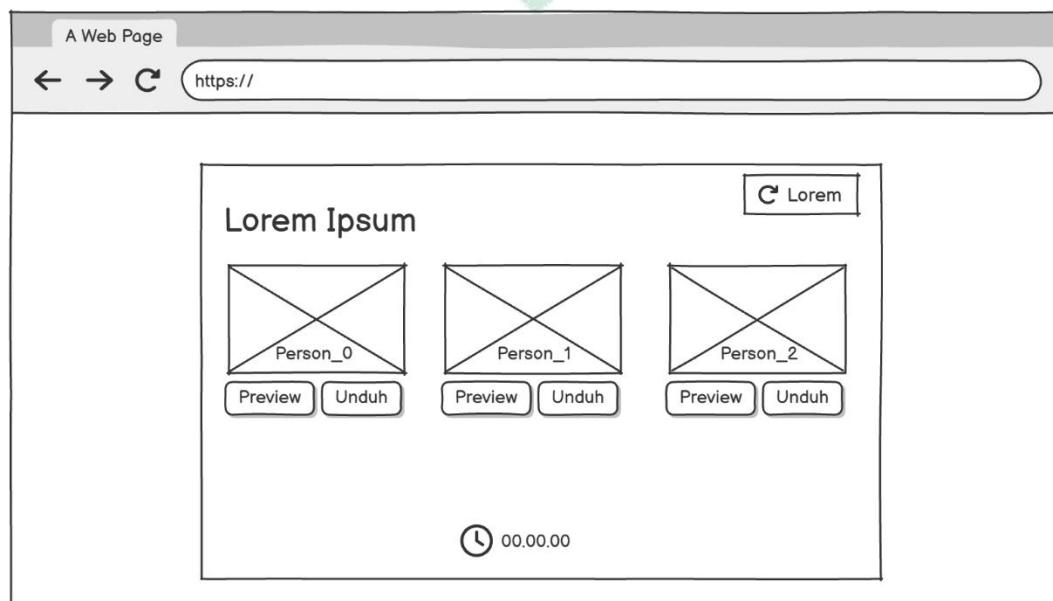
Untuk memberikan umpan balik yang informatif selama sistem bekerja, antarmuka akan menampilkan halaman status proses. Halaman ini berisi *progress bar* yang menunjukkan persentase kemajuan, status tahapan yang sedang berjalan (misalnya mengunduh data atau mengelompokkan wajah), dan estimasi waktu berjalan. Tampilan ini penting untuk mengelola ekspektasi pengguna dan memberikan kepastian bahwa proses sedang berlangsung.



Gambar 3. 8 Wireframe Progress Pengelompokan Foto

6. Tampilan Hasil Pengelompokan Foto

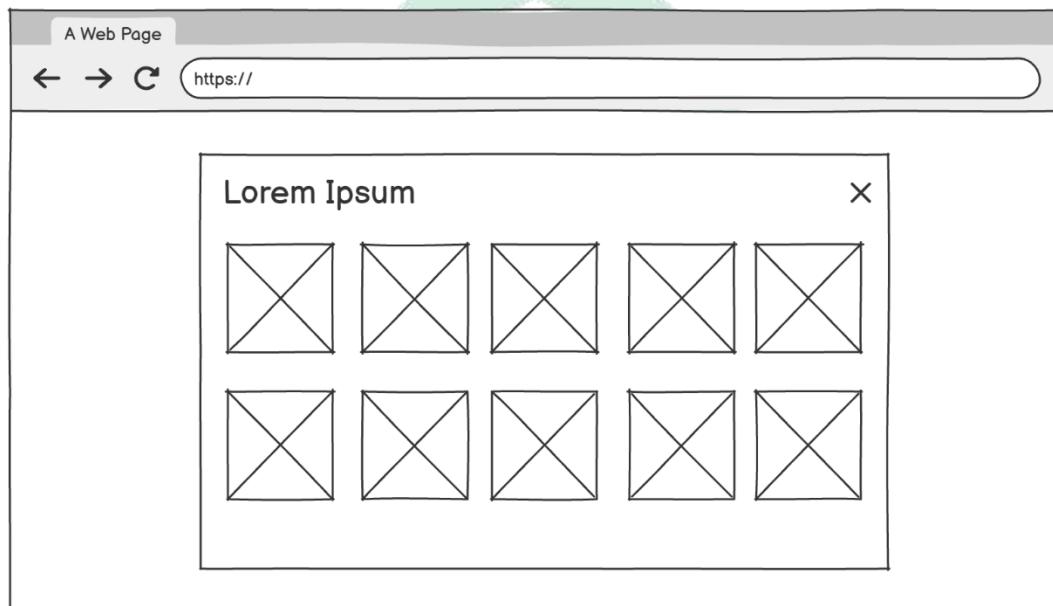
Setelah proses selesai, sistem akan menyajikan hasil pengelompokan dalam format galeri yang mudah dipahami. Setiap kelompok (*cluster*) akan merepresentasikan satu individu yang berhasil diidentifikasi, ditampilkan bersama jumlah foto yang ditemukan. Di bawah setiap *cluster*, tersedia dua tombol aksi yaitu *Preview* untuk melihat pratinjau foto di dalam *cluster* tersebut, dan Unduh untuk menyimpan semua foto dalam *cluster* itu sebagai file ZIP.



Gambar 3. 9 Wireframe Hasil Pengelompokan Foto

7. Tampilan Pratinjau (*Preview*) Cluster

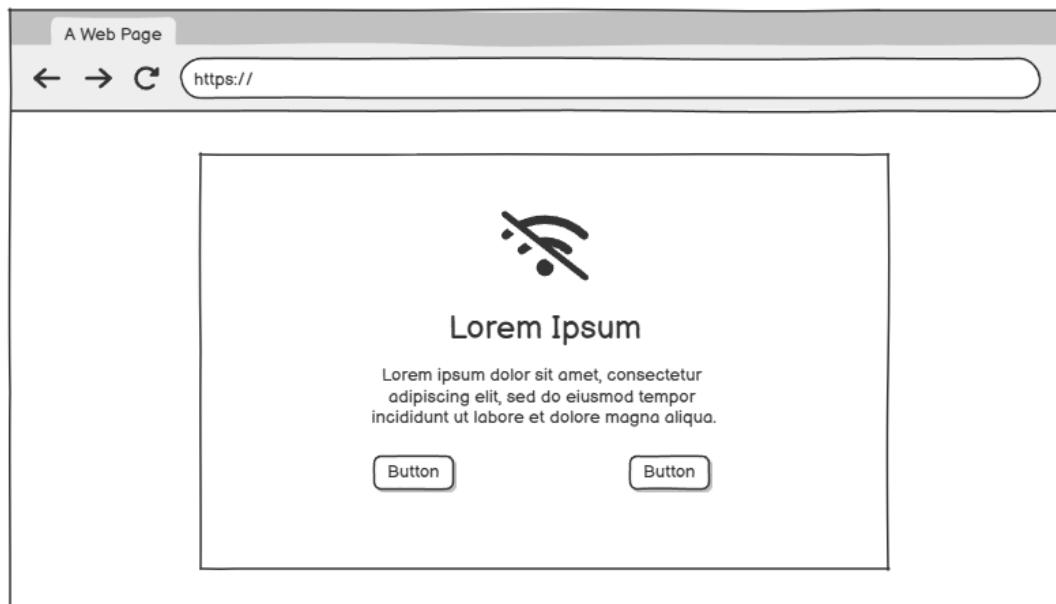
Fitur pratinjau dirancang untuk memungkinkan pengguna memeriksa isi dari setiap *cluster* tanpa harus mengunduhnya terlebih dahulu. Ketika tombol *Preview* ditekan, sebuah jendela modal akan muncul menampilkan seluruh gambar yang termasuk dalam *cluster* yang dipilih. Hal ini memberikan fleksibilitas kepada pengguna untuk memverifikasi keakuratan hasil sebelum melakukan pengunduhan.



Gambar 3. 10 Wireframe Preview Cluster

8. Tampilan Notifikasi Adaptif Mode ZIP

Sistem ini dirancang dengan fitur cerdas untuk menangani kendala jaringan. Jika sistem mendeteksi koneksi ke server Google Drive melambat, sebuah notifikasi akan muncul secara otomatis. Pengguna diberikan opsi untuk tetap menunggu atau beralih ke mode yang disarankan. Halaman ini memberikan opsi kepada pengguna untuk menggunakan dua fitur pada sistem pengelompokan foto otomatis. Untuk itu, dirancang dua tombol untuk melakukan aksi beralih atau tidak beralih mode.



Gambar 3. 11 Wireframe Notifikasi Adaptif Mode ZIP



BAB IV

HASIL DAN PEMBAHASAN

A. Hasil Penelitian

1. Implementasi Antarmuka Sistem

Sistem dikembangkan dengan antarmuka pengguna berbasis web yang dirancang untuk kemudahan penggunaan. Pengguna dapat berinteraksi dengan sistem melalui beberapa tampilan utama yang memandu proses dari awal hingga akhir.

a. Halaman Utama (*Overview*)

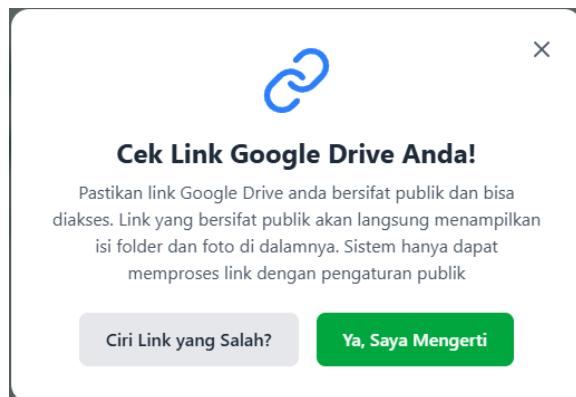
Halaman awal menyajikan gambaran umum tentang fungsionalitas sistem. Pengguna dapat memahami kapabilitas utama sistem sebelum memulai proses pengelompokan.



Gambar 4. 1 Tampilan Halaman Utama Sistem

b. Halaman Informasi *Link Google Drive*

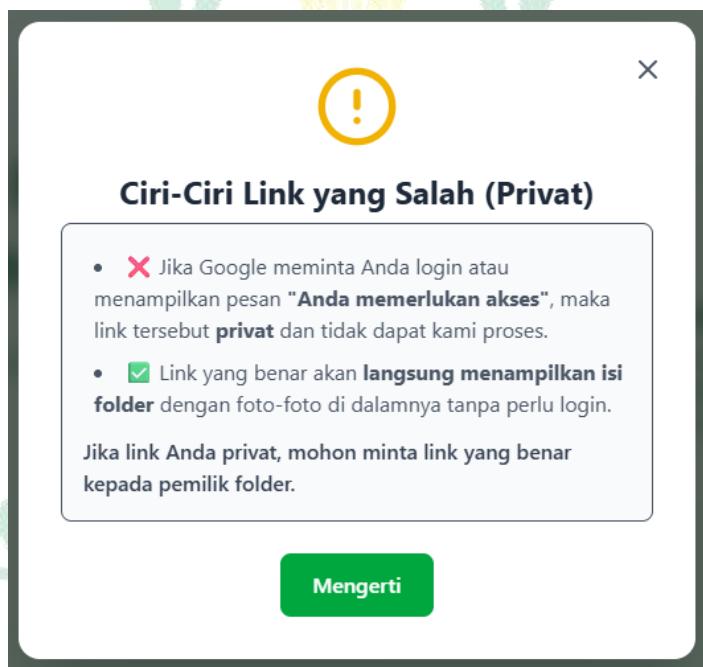
Halaman ini menampilkan informasi dan penjelasan terkait dengan *link* Google Drive yang bisa diproses oleh sistem dan juga menampilkan informasi lanjutan terkait ciri dari *link* Google Drive yang salah dengan menekan tombol sebelah kiri.



Gambar 4. 2 Halaman Informasi Link Google Drive

c. Halaman Informasi Lanjutan *Link Google Drive*

Halaman yang menampilkan ciri ciri detail link Google Drive yang salah atau link yang tidak bisa diproses oleh sistem.

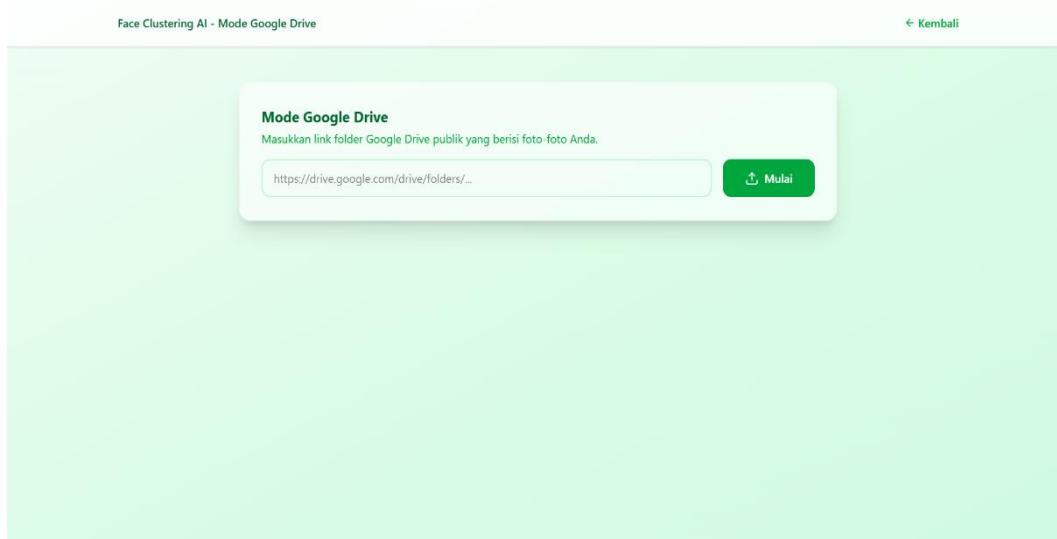


Gambar 4. 3 Halaman Informasi Lanjutan Link Google Drive

d. Halaman Proses Utama

Pada halaman ini, pengguna memulai proses dengan memasukkan tautan Google Drive yang berisi kumpulan foto. Antarmuka ini juga menyediakan opsi untuk beralih ke mode unggah ZIP jika diperlukan. Terdapat informasi yang didapatkan berupa *placeholder* terkait struktural *link* Google

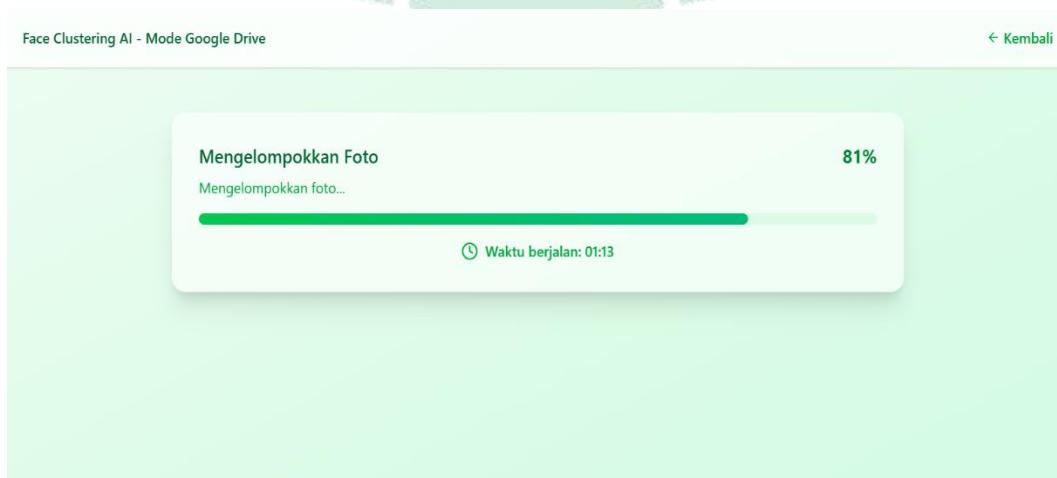
Drive untuk memudahkan pengguna memasukkan *link* yang sesuai.



Gambar 4. 4 Tampilan Input *Link* Google Drive

e. Tampilan Status Proses

Setelah proses dimulai, sistem akan menampilkan *progress bar* beserta status dan estimasi waktu berjalan. Ini memberikan umpan balik visual kepada pengguna mengenai tahapan yang sedang berlangsung, mulai dari pengunduhan hingga pengelompokan.

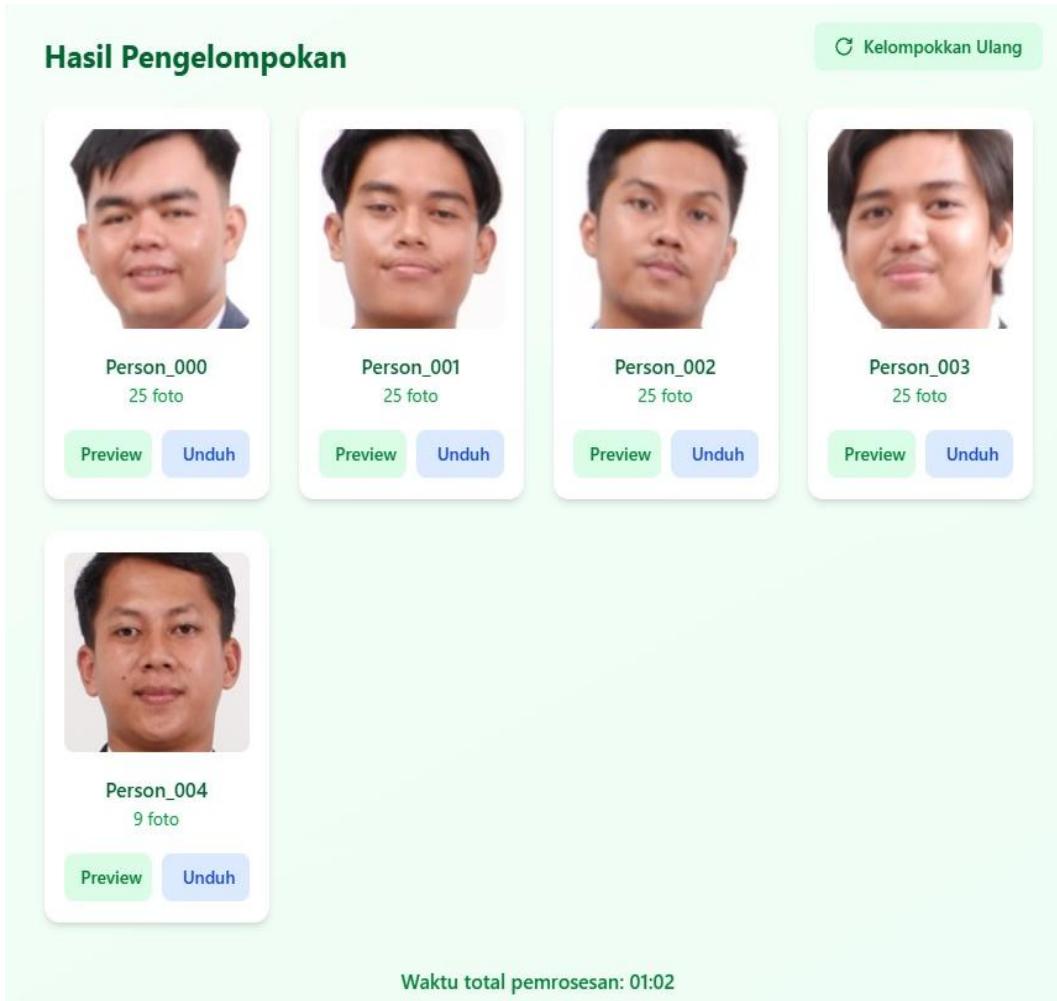


Gambar 4. 5 Tampilan Status Proses dan Waktu Berjalan

f. Tampilan Hasil Pengelompokan Foto

Ketika proses selesai, hasil akan ditampilkan dalam bentuk galeri *cluster*. Setiap *cluster* mewakili satu individu yang teridentifikasi, lengkap

dengan jumlah foto dan gambar pratinjau. Pengguna dapat mengunduh semua foto dalam satu *cluster* sebagai file ZIP atau melihat pratinjaunya.



Gambar 4. 6 Tampilan Hasil Pengelompokan Foto

g. Modal Pratinjau *Cluster*

Fitur pratinjau memungkinkan pengguna untuk melihat seluruh foto yang termasuk dalam satu *cluster* dalam sebuah modal tanpa harus mengunduhnya terlebih dahulu. Tampilan modal pratinjau *cluster* dapat dilihat pada gambar 4. 5.

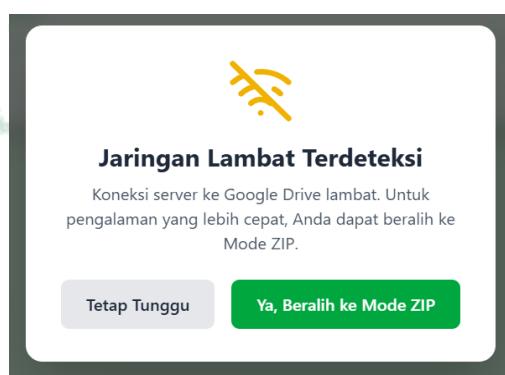
h. Fitur Beralih ke Mode ZIP

Sistem secara cerdas mendeteksi jika koneksi jaringan untuk

mengunduh dari Google Drive lambat. Sebuah notifikasi akan muncul dan menyarankan pengguna untuk beralih ke Mode ZIP untuk proses yang lebih cepat dan andal. Tampilan beralih ke mode ZIP dapat dilihat pada gambar 4. 6.



Gambar 4. 7 Tampilan Pratinjau Gambar dalam Satu *Cluster*



Gambar 4. 8 Notifikasi Jaringan Lambat dan Opsi Beralih ke Mode ZIP

2. Hasil Pengujian Sistem Melalui Metrik Proksi

Pengujian kuantitatif dilakukan dalam beberapa skenario untuk mengukur kinerja sistem melalui metrik proksi yang telah ditentukan.

a. Pengujian Skalabilitas Berdasarkan Kuantitas Data

Pengujian ini bertujuan untuk mengukur bagaimana sistem menangani volume data yang berbeda. Tiga set data dengan jumlah foto 25, 159, dan 1067 digunakan.

Tabel 4.1 Hasil Pengujian Berdasarkan Kuantitas Data

Skenario Pengujian	Jumlah Foto	Jumlah Wajah Terdeteksi	Jumlah Cluster	Waktu Proses (Detik)	Silhouette Score	DBI
Link 1	25	109	5	6,99	0,817	0,616
Link 2	159	2482	39	141,87	0,724	0,855
Link 3	1067	10927	173	1205,62	0,613	1,417

b. Pengujian Ketahanan Berdasarkan Kualitas Data

Pengujian ini bertujuan mengevaluasi ketahanan sistem terhadap variasi kualitas gambar yang umum ditemui.

Tabel 4.2 Hasil Pengujian Berdasarkan Kualitas Data

c. Pengujian Keandalan Berdasarkan Metode Input

Pengujian ini membandingkan dua metode input untuk mengukur dampak metode terhadap keandalan dan kinerja.

Tabel 4.3 Hasil Perbandingan Metode Input

Metode Input	Jumlah Foto Awal	Foto Berhasil Diproses	Waktu Proses (Detik)	Jumlah Cluster	Silhouette Score	DBI
<i>Link</i>						
Google	25	25	68	5	0,817	0,616
Drive						
ZIP File	25	25	8	5	0,817	0,616

B. Pembahasan

Bagian ini akan menguraikan dan menganalisis hasil penelitian yang telah disajikan untuk menjawab rumusan masalah yang telah ditetapkan. Analisis difokuskan pada interpretasi data dari pengujian kuantitatif berbasis kinerja, menggunakan metrik proksi seperti waktu proses, *Silhouette Score*, dan DBI untuk mengevaluasi aspek kemudahan, akurasi, dan keandalan sistem.

1. Sistem Pengelompokan Foto Otomatis sebagai Solusi Akses Foto Personal

Sistem pengelompokan foto otomatis dapat membantu pengguna dalam mengakses foto personal mereka berdasarkan identitas wajah dengan mudah. Hasil penelitian secara kuantitatif menunjukkan bahwa sistem ini memberikan solusi nyata terhadap tantangan tersebut melalui dua aspek utama yaitu efisiensi dan akurasi.

a. Efisiensi sebagai Proksi Kemudahan

Kemudahan akses secara langsung berkaitan dengan minimalnya usaha dan waktu yang dibutuhkan pengguna. Sistem ini mengotomatisasi tugas yang secara manual akan sangat melelahkan dan memakan waktu, yaitu menyortir ratusan atau ribuan foto. Tabel 4.1 menunjukkan bahwa sistem mampu memproses 1.067 foto dan mendekripsi 10.927 wajah dalam waktu sekitar 20 menit (1205,62 detik). Meskipun waktu proses meningkat seiring jumlah data, proses otomatis ini jauh lebih efisien dibandingkan jika pengguna harus meninjau dan mengelompokkan setiap foto secara manual. Dengan demikian, sistem secara signifikan mengurangi beban kognitif dan waktu pengguna, yang merupakan manifestasi dari kemudahan.

b. Akurasi Pengelompokan sebagai Proksi Akses yang Tepat Sasaran

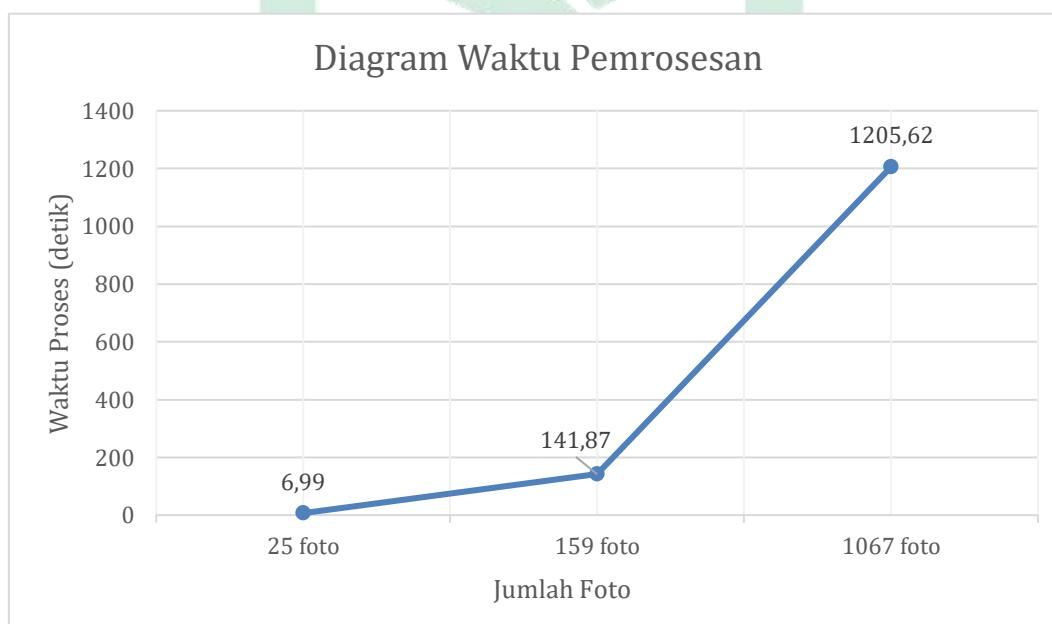
Kemudahan akses tidak hanya tentang kecepatan, tetapi juga tentang ketepatan. Pengguna harus dapat percaya bahwa kelompok foto yang dihasilkan benar-benar berisi orang yang sama. Di sinilah metrik kualitas *cluster* berperan. Pada skenario ideal (Tabel 4.1, *Link 1*), sistem mencapai *Silhouette Score* yang sangat tinggi (0,817) dan DBI yang rendah (0,616). Nilai *Silhouette Score* yang mendekati 1 menunjukkan bahwa *cluster* yang terbentuk sangat padat (anggota *cluster* sangat mirip satu sama lain) dan terpisah dengan baik (berbeda dari *cluster* lain). Ini membuktikan bahwa sistem dapat secara akurat mengidentifikasi dan memisahkan individu yang berbeda ke dalam kelompok-kelompok yang relevan. Ketika pengguna mengklik sebuah *cluster*, mereka dapat dengan yakin menemukan foto-foto dari orang yang mereka cari. Kemampuan untuk menyajikan hasil yang akurat inilah yang membuat akses menjadi mudah dan efektif.

2. Otomatisasi Pengelompokan Foto dengan Penerapan Algoritma DBSCAN

Algoritma DBSCAN dapat diterapkan untuk proses otomatisasi ini. Hasil pengujian memberikan wawasan mendalam tentang kinerja, skalabilitas, dan ketahanan DBSCAN dalam konteks pengelompokan wajah.

a. Penerapan DBSCAN pada Skala Berbeda (Skalabilitas)

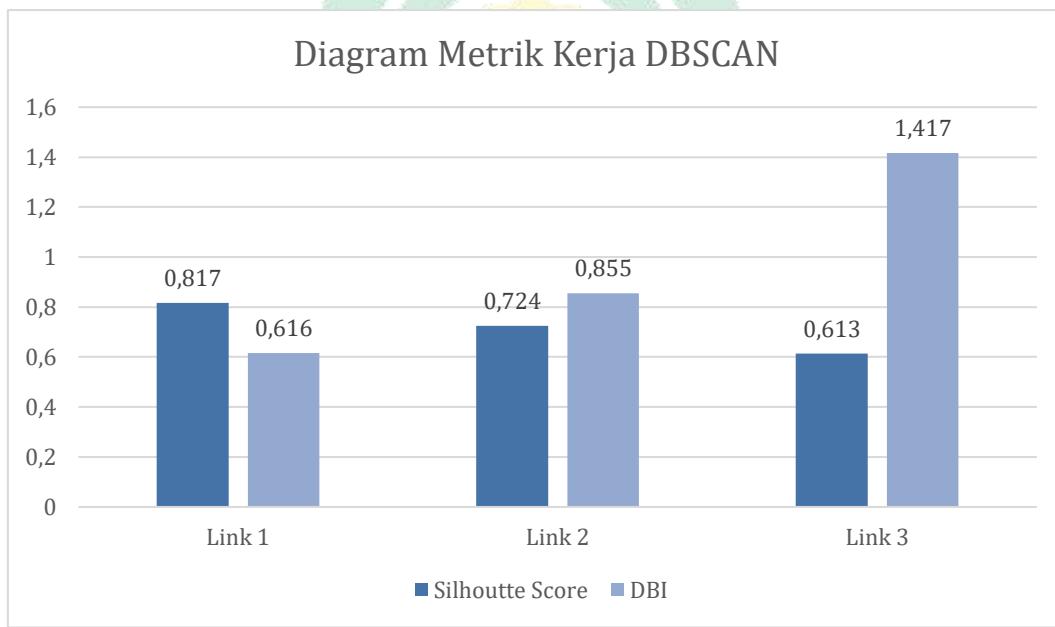
Hasil pengujian menunjukkan bagaimana kinerja DBSCAN berubah seiring dengan peningkatan volume data. Pada gambar 4.9 Terlihat adanya peningkatan waktu proses yang signifikan, dari 6,99 detik untuk 25 foto menjadi 1205,62 detik untuk 1067 foto. Ini menunjukkan bahwa penerapan DBSCAN memiliki biaya komputasi yang skalanya meningkat seiring bertambahnya jumlah titik data.



Gambar 4. 9 Diagram Waktu Pemrosesan

Selain kinerja DBSCAN yang berubah seiring dengan peningkatan volume data, terdapat penurunan kualitas pengelompokan pada set data yang lebih besar. Gambar 4.10 menunjukkan bahwa *Silhouette Score* menurun dari

0,817 menjadi 0,613, sementara DBI meningkat dari 0,616 menjadi 1,417. Penurunan ini mengindikasikan bahwa semakin banyak wajah yang diproses, semakin besar kemungkinan beberapa *cluster* menjadi kurang padat atau saling tumpang tindih. Hal ini wajar terjadi karena peningkatan variasi pose, pencahayaan, dan ekspresi wajah dalam set data yang besar, sehingga menyulitkan algoritma untuk mempertahankan pemisahan *cluster* yang sempurna. Ini adalah pertimbangan penting dalam penerapan praktis DBSCAN, yaitu terdapat *trade-off* antara jumlah data yang diproses dan kualitas ideal dari *cluster* yang dihasilkan.

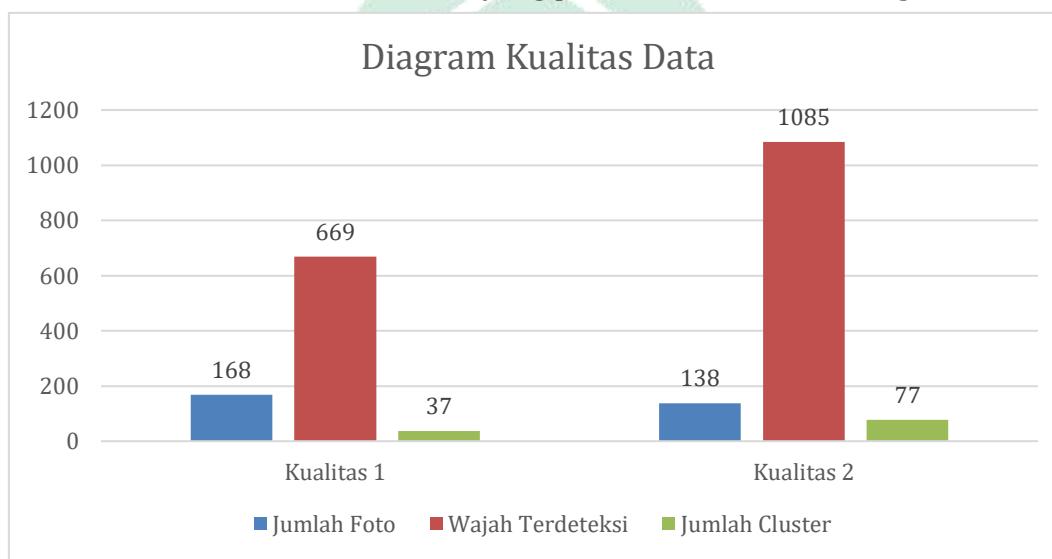


Gambar 4. 10 Diagram Metrik Kerja DBSCAN

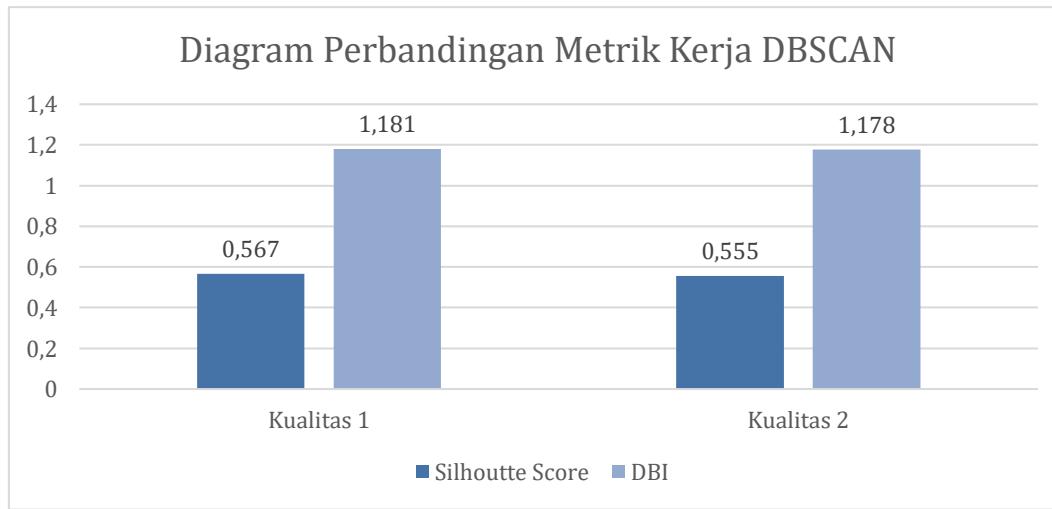
b. Penerapan DBSCAN pada Kualitas Data Berbeda (Ketahanan)

Hasil pengujian yang menguji ketahanan algoritma terhadap kualitas data input yang divisualisasikan dalam bentuk diagram pada gambar 4.11 dan gambar 4.12. Hasilnya menunjukkan bahwa kualitas gambar sangat memengaruhi kinerja DBSCAN. Dibandingkan dengan skenario data di Tabel 4.1 (*Silhouette Score* 0,817), skenario dengan kualitas gambar yang lebih

rendah atau bervariasi (*outdoor, indoor*, kamera HP) menghasilkan *Silhouette Score* yang jauh lebih rendah (0,567 dan 0,555) dan DBI yang lebih tinggi (1,181 dan 1,178). Ini membuktikan bahwa penerapan DBSCAN sangat bergantung pada kualitas *face embeddings* (representasi numerik dari wajah) yang dihasilkan pada tahap sebelumnya. Gambar dengan kualitas buruk (buram, pencahayaan tidak merata, atau sudut wajah yang ekstrem) menghasilkan *embeddings* yang kurang konsisten, sehingga mempersulit DBSCAN untuk membentuk *cluster* yang padat dan terdefinisi dengan baik.



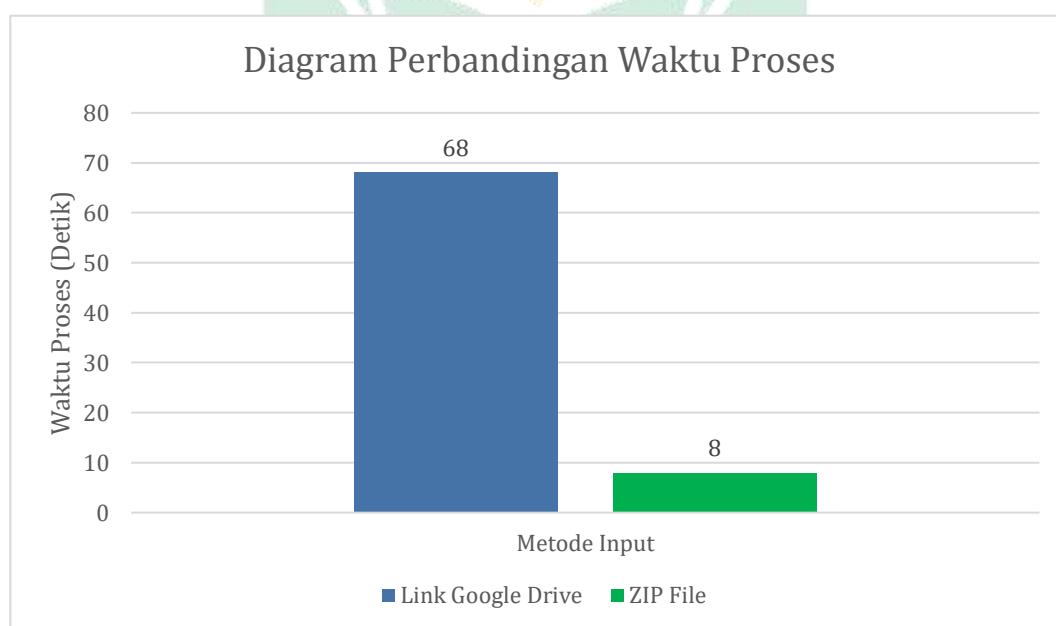
Gambar 4. 11 Diagram Kualitas Data



Gambar 4. 12 Diagram Perbandingan Metrik Kerja DBSCAN

c. Keandalan Sistem Berdasarkan Metode Input

Hasil pengujian menunjukkan temuan menarik terkait penerapan sistem secara keseluruhan. Algoritma DBSCAN itu sendiri menghasilkan *output* yang identik (jumlah *cluster*, *Silhouette Score*, dan DBI yang sama) terlepas dari apakah foto diunggah melalui Google Drive atau *file ZIP*. Ini menunjukkan bahwa algoritma inti bekerja secara konsisten dan andal. Namun, kinerja sistem secara keseluruhan (terutama Waktu Proses) sangat dipengaruhi oleh metode input. Gambar 4.13 menunjukkan bahwa proses melalui *file ZIP* lokal (8 detik) jauh lebih cepat daripada melalui tautan Google Drive (68 detik), karena adanya latensi jaringan untuk mengunduh *file* dari *cloud*.



Gambar 4. 13 Diagram Perbandingan Waktu Proses

d. Analisis Visual *Unclustered Faces* sebagai Bukti Efektivitas DBSCAN dalam Pemisahan Noise

Analisis visual terhadap kategori *noise* dari hasil pengelompokan menunjukkan keunggulan utama DBSCAN, yaitu kemampuannya untuk secara

eksplisit mengidentifikasi dan memisahkan *noise* atau *outlier*. Sistem tidak hanya menghasilkan *cluster* wajah yang koheren, tetapi juga sebuah kategori khusus untuk wajah-wajah yang tidak dapat dikelompokkan atau *unclustered faces*.



Gambar 4. 12 *Unclustered Faces*

Berdasarkan hasil pengelompokan foto, ditemukan adanya sejumlah foto yang tergolong dalam kategori *unclustered faces* atau *noise*. Kondisi ini menunjukkan bahwa DBSCAN mampu secara efektif memisahkan data yang tidak memiliki kesamaan signifikan dengan kelompok manapun, sehingga mencegah terjadinya kesalahan pengelompokan. Analisis visual terhadap *unclustered faces* memperlihatkan bahwa wajah-wajah dalam foto tersebut memang memiliki karakteristik yang berbeda, seperti sudut pengambilan gambar, pencahayaan yang rendah, atau wajah yang hanya terlihat sebagian.

Sebagai ilustrasi, foto-foto yang ditampilkan pada Gambar 4.12 merupakan contoh *unclustered faces* yang dihasilkan oleh proses *clustering*. Terlihat bahwa beberapa wajah hanya tampak sebagian atau terhalang objek lain, sementara yang lain memiliki perbedaan signifikan dalam ekspresi wajah atau pencahayaan. Fenomena ini menegaskan keunggulan DBSCAN dalam mengidentifikasi *outlier* secara otomatis tanpa memerlukan jumlah *cluster* yang telah ditentukan sebelumnya. Dengan demikian, pemisahan wajah-wajah tersebut dari *cluster* utama menjadi bukti bahwa DBSCAN optimal dalam mendeteksi dan memisahkan *noise*, sehingga kualitas hasil pengelompokan dapat dipertahankan pada tingkat yang tinggi.

C. Analisis Optimasi Parameter DBSCAN

Proses optimasi parameter DBSCAN dilakukan untuk meningkatkan kualitas hasil pengelompokan foto yang telah dievaluasi sebelumnya menggunakan metrik *Silhouette Score* dan DBI. Optimasi ini dilakukan melalui dua tahap utama. Pertama, penerapan *filter* kualitas yang lebih ketat dengan menaikkan nilai *quality threshold*, sehingga jumlah wajah berkualitas rendah dapat dikurangi dan rata-rata kualitas *embedding* meningkat. Kedua, dilakukan pencarian parameter *granular* dengan memanfaatkan fungsi *find optimal eps*

yang telah dimodifikasi menggunakan rentang nilai *eps* yang lebih halus dan variasi *min samples* yang lebih beragam. Langkah ini membutuhkan waktu komputasi yang lebih lama, tetapi penting untuk menemukan kombinasi parameter terbaik.

Pengujian dilakukan pada dua skenario, yaitu skalabilitas pada *Link 1* dan ketahanan pada Kualitas 1. Hasil pengujian sebelumnya atau pengujian awal dan hasil pengujian akhir atau pengujian terbaru setelah optimasi ditunjukkan pada Tabel 4.4 dan Tabel 4.5.

Tabel 4.4 Hasil Pengujian Awal

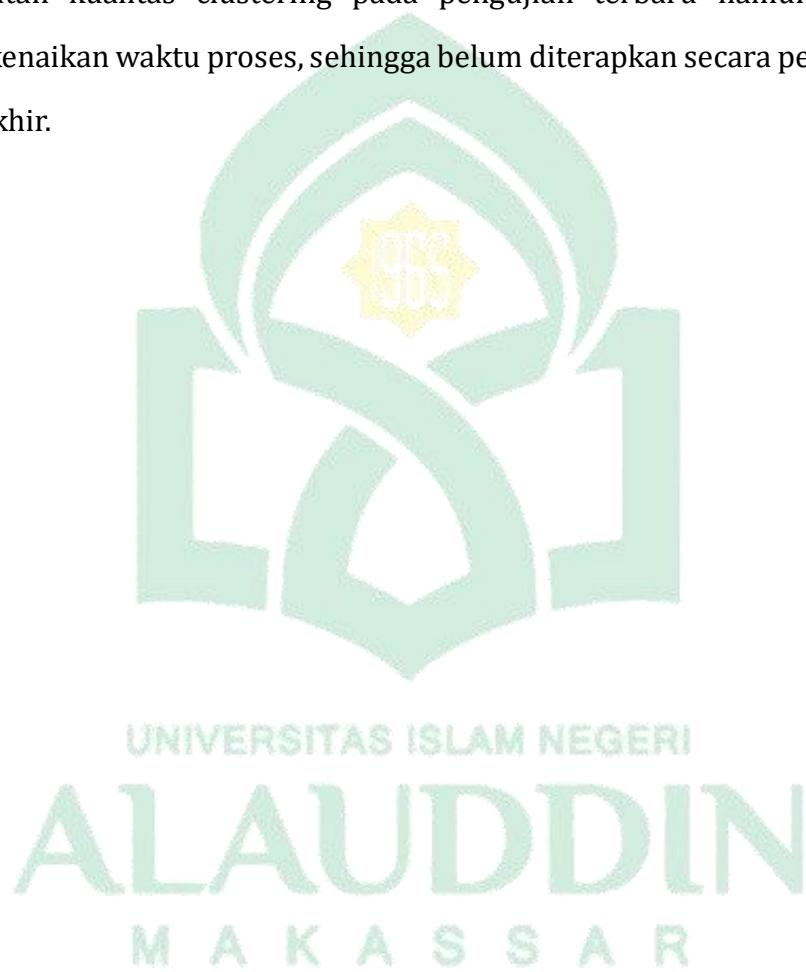
Skenario Pengujian	<i>Silhouette Score</i>	DBI	Waktu Proses (Detik)
<i>Link 1</i>	0,817	0,616	6,99
Kualitas 1	0,567	1,181	

Tabel 4.5 Hasil Pengujian Akhir / Terbaru

Skenario Pengujian	<i>Silhouette Score</i>	DBI	Waktu Proses (Detik)
<i>Link 1</i>	0,829	0,577	8,30
Kualitas 1	0,607	1,031	

Berdasarkan hasil pengujian pada kedua tabel, optimasi parameter DBSCAN menghasilkan peningkatan nilai *Silhouette Score* dan penurunan DBI pada kedua skenario. Pada skenario *Link 1*, *Silhouette Score* meningkat dari 0,817 menjadi 0,8289 dan DBI turun dari 0,616 menjadi 0,577, meskipun waktu proses bertambah dari 6,99 detik menjadi 8,30 detik. Sementara itu, pada skenario Kualitas 1, *Silhouette Score* meningkat dari 0,567 menjadi 0,6074 dan DBI menurun dari 1,181 menjadi 1,031. Hasil ini menunjukkan bahwa penerapan filter kualitas yang lebih ketat dan pencarian parameter granular berhasil menghasilkan *cluster* yang lebih terpisah dengan kualitas yang lebih baik.

Sebagai konklusi dari analisis optimasi parameter DBSCAN yang diambil oleh peneliti, meskipun hasil pengujian terbaru menunjukkan perbaikan pada *Silhouette Score* dan DBI, penelitian ini tetap mempertahankan hasil pengujian awal atau hasil pengujian sebelumnya karena fokus utama pada penelitian ini adalah pada efisiensi pengelompokan foto otomatis. Peningkatan kualitas clustering pada pengujian terbaru namun disertai dengan kenaikan waktu proses, sehingga belum diterapkan secara penuh pada sistem akhir.



BAB V

PENUTUP

A. Kesimpulan

Hasil analisis dari penelitian yang telah dilakukan dapat ditarik kesimpulan bahwa:

1. Sistem pengelompokan foto otomatis berhasil membantu pengguna mengatasi kesulitan dalam mengakses foto personal mereka dengan menyediakan solusi yang efisien dan akurat. Secara efisiensi, sistem ini secara signifikan mengurangi beban kerja dan waktu pengguna dengan mengotomatisasi proses pengelompokan foto, di mana pengujian menunjukkan kemampuan untuk memproses 1.067 foto dalam waktu sekitar 20 menit. Dari segi akurasi, yang merupakan kunci untuk kemudahan akses yang tepat sasaran, sistem ini menunjukkan kinerja yang sangat baik. Dengan pencapaian *Silhouette Score* setinggi 0,817, sistem terbukti mampu membentuk kelompok-kelompok foto yang padat dan terpisah dengan baik, sehingga pengguna dapat dengan percaya diri menemukan foto individu yang dicari dalam kelompok yang relevan
2. Algoritma DBSCAN dapat diterapkan secara efektif untuk proses otomatisasi pengelompokan foto berdasarkan wajah, namun kinerjanya dipengaruhi oleh skalabilitas dan kualitas data. Hasil pengujian menunjukkan penurunan nilai *Silhouette Score* dengan persentase 24,98% pada data yang lebih besar dan skor yang lebih rendah pada data berkualitas beragam menggarisbawahi tantangan ini. Namun, analisis terhadap kategori *noise* mengungkap kekuatan fundamental DBSCAN

yang menjadikannya sangat cocok untuk tugas ini. Kemampuannya untuk secara cerdas mengisolasi *outlier* atau wajah wajah yang tidak dapat dikelompokkan, membuktikan kemampuan DBSCAN dalam menangani data yang tidak sempurna memberikan keandalan dan akurasi praktis yang tinggi, memastikan bahwa hanya wajah yang benar-benar relevan yang dikelompokkan bersama.

B. Saran

Pengembangan lebih lanjut diperlukan guna meningkatkan efektivitas dan kapabilitas sistem. Beberapa saran yang dapat menjadi fokus penelitian selanjutnya antara lain:

1. Penelitian selanjutnya disarankan untuk fokus pada optimasi kinerja guna menangani data skala besar, misalnya dengan mengeksplorasi metode clustering lain seperti HDBSCAN dan teknik *indexing* yang lebih maju serta mengoptimalkan proses pengambilan data dari *cloud*.
2. Integrasi modul pra-pemrosesan berbasis AI guna memperbaiki kualitas gambar secara otomatis sebelum analisis, sehingga meningkatkan akurasi sistem pada foto berkualitas rendah sangat disarankan untuk penelitian selanjutnya
3. Disarankan untuk penelitian selanjutnya, pengembangan sistem mampu menganalisis frekuensi kemunculan wajah secara bersamaan untuk membangun *social graph*, sehingga dapat memberikan wawasan analitis mengenai interaksi sosial dalam sebuah organisasi, tidak hanya sebatas pengelompokan foto.

DAFTAR PUSTAKA

- Al-Bukhari, M. ibn I. (1997). *Shahih Al-Bukhari Vol. 8* (Vol. 8). Darussalam.
- Al-Sheikh, A. bin M. bin A. bin I. (2004). *Tafsir Ibnu Katsir Jilid 7* (M. Y. Harun, F. Okbah, T. S. Al-Katsiri, A. I. Al-Atsari, Arman, & B. Salam, Ed.; Vol. 7). Pustaka Imam asy-Syafi'i.
- Binus University. (2025). *Cross-Industry Standard Process for Data Mining (CRISP-DM)* - MMSI BINUS University. <https://mmsi.binus.ac.id/2020/09/18/cross-industry-standard-process-for-data-mining-crisp-dm/>
- Bjerge, K., Bodesheim, P., & Karstoft, H. (2025). Deep Image Clustering with Model-Agnostic Meta-Learning. *Proceedings of the International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*, 2, 286–297. <https://doi.org/10.5220/0013114600003912>
- Cheung, M. (2023). *Face Clustering for Connection Discovery from Event Images*. 2022-March. <https://doi.org/10.1145/nnnnnnnn.nnnnnnnn>
- Choo, S.-O., Lee, S.-Y., Seok, J.-H., Lee, G.-M., Lee, T.-S., & Yoo, H. (2024). *Automatic Photo Classification System Based on Face Feature Extraction and Clustering*. <https://en.wikipedia.org/wiki/DBSCAN>
- Daqiqil Id, I., & Evfi Mahdiyah, dan. (2021). Modifikasi DBSCAN (Density-Based Spatial Clustering With Noise) pada Objek 3 Dimensi. Dalam *Jurnal Komputer Terapan* (Vol. 3, Nomor 1). <http://jurnal.pcr.ac.id>
- Datacamp. (2025). *Feature Extraction in Machine Learning: A Complete Guide*. <https://www.datacamp.com/tutorial/feature-extraction-machine-learning>
- Deng, J., Zhu, Z., & Zafeiriou, S. (2024). *Masked Face Recognition Challenge: The InsightFace Track Report*. <https://ibug.doc.ic.ac.uk/resources/>
- Destiana, N., Aisyah, N., Sebayang, A., Yuniasih, I., Studi Sistem Informasi, P., Ilmu Komputer, F., Veteran Jakarta, U., Bsi, U., & Rs Fatmawati, J. (2024). *Implementasi Deep Learning Berbasis Tensorflow Untuk Pengenalan Wajah*.
- Fauzan, M., & Zainarti. (2025). Analisis Hadist "Manajemen Waktu." *Jurnal Penelitian Ekonomi Manajemen dan Bisnis*, 4(1), 269–277. <https://doi.org/10.55606/jekombis.v4i1.4794>
- Geetha, A. V., Mala, T., Priyanka, D., & Uma, E. (2024). Multimodal Emotion Recognition with Deep Learning: Advancements, challenges, and future directions. *Information Fusion*, 105. <https://doi.org/10.1016/j.inffus.2023.102218>
- Google Workspace. (2025). *Ringkasan Google Drive API | Google for Developers*. <https://developers.google.com/workspace/drive/api/guides/about-sdk?hl=id>
- IDC. (2022). *Equalizing Time Spent on Data Management vs. Analytics / IDC Blog*. IDC. <https://blogs.idc.com/2018/08/23/time-crunch-equalizing-time-spent-on-data-management-vs->

- analytics/?_gl=1*_fu2m9r*_gcl_au*MjU1MTgwMDQ4LjE3NDIzMzUwM
DE.*_ga*ODM4Mzc1ODQzMjE3NDIzMzUwMDI.*_ga_Y7CNRMFF6J*MTc
0MjMzNTAwMS4xLjEuMTc0MjMzNTE3MC4yNi4wLjA.*_ga_541ENG1F
9X*MTc0MjMzNTAwMS4xLjEuMTc0MjMzNTE3MC4yNi4wLjA.
- Indira, K., & Kavitha Devi, M. K. (2020). Multi Cloud Based Service Recommendation System Using DBSCAN Algorithm. *Wireless Personal Communications*, 115(2), 1019–1034. <https://doi.org/10.1007/S11277-020-07609-3>
- Junqing Tang, S. H. J. W. B. H. J. P. (2023). A Comparative Analysis of Performance-Based Resilience Metrics via a Quantitative-Qualitative Combined Approach: Are We Measuring the Same Thing? *International Journal of Disaster Risk Science*.
- Kementrian Agama. (2022a). <https://quran.kemenag.go.id/quran/per-ayat/surah/17?from=26&to=26>
- Kementrian Agama. (2022b). <https://quran.kemenag.go.id/quran/per-ayat/surah/103?from=1&to=3>
- Kementrian Agama. (2022c). <https://quran.kemenag.go.id/quran/per-ayat/surah/45?from=13&to=13>
- Kumar, S. N., Nirmalkumar, P., & H, K. (2025). *Feature Extraction And Classification In Context Of Machine Learning*. <https://store.iipbooks.com/book/feature-extraction-and-classification-in-context-of-machine-learning/>
- Liu, H., Li, I., Liang, Y., Sun, D., Yang, Y., & Yang, H. (2024). *Research on Deep Learning Model of Feature Extraction Based on Convolutional Neural Network*.
- Liu, Y., Pan, Z., Wang, L., & Wang, Y. (2022). A new fast inverted file-based algorithm for approximate nearest neighbor search without accuracy reduction. *Information Sciences*, 608, 613–629. <https://doi.org/10.1016/j.ins.2022.06.086>
- Meena, S., Vats, K., & Kumar, D. (2022). Review of factors affecting facial recognition algorithms performance. *International journal of health sciences*, 9528–9541. <https://doi.org/10.53730/ijhs.v6ns3.8253>
- Mustikasari, & Salman, N. (2023). *Analisis Klaster Berbasis Kepadatan dengan DBSCAN dan OPTICS* (Vol. 8, Nomor 1). <http://journal.uinalauddin.ac.id/index.php/insypro>
- Mutiah, S., Hasnataeni, Y., Fitrianto, A., Erfiani, E., & Jumansyah, L. M. R. D. (2024). Perbandingan Metode Klustering K-Means dan DBSCAN dalam Identifikasi Kelompok Rumah Tangga Berdasarkan Fasilitas Sosial Ekonomi di Jawa Barat. *Teorema: Teori dan Riset Matematika*, 9(2), 247. <https://doi.org/10.25157/teorema.v9i2.16290>
- Nixon, M. S., & Aguado, A. S. . (2020). *Feature extraction and image processing for computer vision*.
- O'Toole, A. J., & Castillo, C. D. (2021). Face Recognition by Humans and Machines: Three Fundamental Advances from Deep Learning. Dalam *Annual Review of Vision Science* (Vol. 7, hlm. 543–570). Annual Reviews Inc. <https://doi.org/10.1146/annurev-vision-093019-111701>

- Razec, I. (2022). *The Concept of Identity in the Technological Era*. 572–589. <https://doi.org/10.18662/wlc2021/58>
- Rohman, M. M., Bani, M. D., Antonny, G., Aryasatya, B., & Muri, D. (2023). *Metode Penelitian Kuantitatif Dan Kualitatif: Teori Dan Praktik*. Get Press Indonesia. <https://www.researchgate.net/publication/377329440>
- Saragih, R. E., & To, Q. H. (2022). A Survey of Face Recognition Based on Convolutional Neural Network. Dalam *Indonesian Journal of Information Systems (IJIS)* (Vol. 4, Nomor 2).
- Schwarz, M., Chapman, K. A., & Häussler, B. (2025). *Multilingual Medical Entity Recognition and Cross-lingual Zero-Shot Linking with Facebook AI Similarity Search*.
- Shihab, M. Q. (2021). *Tafsir Al-Misbah Pesan, Kesan dan Keserasian Al-Qur'an*.
- Sholichin, M. F. M. (2022). *Peran Seni Fotografi Dalam Menumbuhkan Kemandirian Ekonomi Fotografer di Kudus*. <https://ojs.petra.ac.id/ojsnew/index.php/dkv/>
- Sholva, Y., & Novriando, H. (2022). Aplikasi Manajemen Berkas Usulan dan Penilaian Sidang Skripsi menggunakan Layanan Google Drive API. *JEPIN (Jurnal Edukasi dan Penelitian Informatika)*, 8(2), 388. <https://doi.org/10.26418/jp.v8i2.54571>
- Simovici, D. A. . (2022). *Clustering : theoretical and practical aspects*. World Scientific Publishing Co. Pte. Ltd.
- Statista. (2022). *BeReal - statistics & facts* / Statista. <https://www.statista.com/topics/10096/bereal/>
- Sujatmiko, B. (2023). *Analisis Metode Clustering Dengan Algoritma Density-Based Spatial Clustering of Application With Noise (DBSCAN) Untuk Pengelompokan Cyberbullying di Instagram*.
- Upadhyay, U., & Gupta, S. (2024). A Survey on Image Feature Extraction Techniques. Dalam *International Journal of Scientific Research & Engineering Trends* (Vol. 10, Nomor 3).
- Varun, A., Rohit, T., Maheshwari, K. M. U., Chaitanya, A. N. V., & Prateek, T. S. S. (2024). Face Clustering with Advanced Techniques. *AIP Conference Proceedings*, 3075(1). <https://doi.org/10.1063/5.0217137/3305155>
- Wajdi, F., Seplyana, D., Rumahlewang, E., Nour Halisa, N., Rusmalinda, S., Kristiana, R., Fathun Niam, M., Wahyuning Purwanti, E., Melinasari, S., & Kusumaningrum, R. (2024). *Metode Penelitian Kuantitatif*. www.freepik.com
- Yunianto, I. (2021). *Teknik Fotografi / Belajar Dari Basic Hingga Professional*.
- Zebua, R., Heroza, R., Adrian, M., & Atrinawati, L. (2022). Determination of Discounts Using K-Means Clustering with RFM Models in Retail Business. *Jurnal CoreIT: Jurnal Hasil Penelitian Ilmu Komputer dan Teknologi Informasi*, 8, 1. <https://doi.org/10.24014/coreit.v8i1.14695>
- Zhang, H., & Peng, Y. (2024). Image Clustering: An Unsupervised Approach to Categorize Visual Data in Social Science Research. *Sociological Methods and Research*, 53(3), 1534–1587. <https://doi.org/10.1177/00491241221082603>

DAFTAR RIWAYAT HIDUP



Rafiul Muiz.K Merupakan Nama Penulis Skripsi ini. Penulis merupakan anak dari pasangan Kasmadi dan Harming, S.Pd.I, anak ketiga dari tiga bersaudara. Penulis dilahirkan di Mattoanging pada tanggal 21 Februari 2003. Penulis menempuh pendidikan dimulai dari TK ABA Mattoanging , SD Inpres 5/81 Mattoanging, MTS An Nur Nusa, kemudian MA An Nur Nusa dan sekarang berkuliah di Universitas Islam Negeri Alauddin Makassar, Fakultas Sains dan Teknologi, Jurusan Teknik Informatika. Motivasi hidup “Jika saya bisa menjadi developer, mengapa saya harus jadi player?”.

Semoga dalam penulisan tugas akhir skripsi ini dapat dan mampu memberikan hal positif atau kontribusi positif bagi masyarakat dan mahasiswa di dunia pendidikan terkhusus bidang teknologi informasi.

Akhir kata penulis mengucapkan rasa syukur yang sedalam-dalam dan sebesar-besarnya dan tak lupa mengucapkan terimakasih kepada orang-orang yang sudah terlibat dan diri sendiri atas terselesaiannya skripsi yang berjudul **“PENGELOMPOKAN FOTO OTOMATIS BERDASARKAN IDENTIFIKASI WAJAH MENGGUNAKAN ALGORITMA DENSITY BASED SPATIAL CLUSTERING OF APPLICATION WITH NOISE”**