Jurnal SimanteC
Vol. x, No. x Desember 2023
P-ISSN : 2088-2130
E-ISSN : 2502-4884

OPTIMASI SISTEM AUTENTIKASI PENGENALAN WAJAH MENGGUNAKAN INCEPTIONRESNETV1 DENGAN PRETRAINED VGGFACE2

OPTIMIZATION OF FACIAL RECOGNITION AUTHENTICATION SYSTEM USING INCEPTIONRESNETV1 WITH PRETRAINED VGGFACE2

[1]Ellexia Leonie Gunawan, [2] I Gede Susrama Mas Diyasa*, [3] Wahyu Syaifullah Jauharis Saputra

[1,3]Program Studi Sains Data, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional Veteran Jawa Timur
[2]Program Magister Teknologi Informasi, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional Veteran Jawa Timur
Jl. Rungkut Madya, Gn. Anyar, Kec. Gn. Anyar, Surabaya, Jawa Timur
*e-mail: igsusrama.if@upnjatim.ac.id
Corresponding author. Phone : *+62 000 000 0000*

(***dikirim*****:** ?, ***direvisi*****:** ?, ***diterima*****:** ? diisi oleh editor)

Abstrak

Pengenalan wajah sebagai metode autentikasi biometrik terus berkembang karena keamanannya yang tinggi dan kemudahan dalam penggunaan. Namun, pelatihan model dari awal menghadapi tantangan kebutuhan data besar dan komputasi tinggi. Penelitian ini bertujuan mengoptimalkan sistem autentikasi wajah menggunakan arsitektur InceptionResNetV1 dengan pendekatan *transfer learning* dari model *pretrained* VGGFace2, serta membandingkan kinerjanya dengan CASIA-WebFace. Deteksi wajah dilakukan menggunakan YOLOv8, *embedding* wajah dihasilkan oleh InceptionResNetV1, dan autentikasi dilakukan dengan menghitung jarak Euclidean antar *embedding*. Data wajah dikumpulkan dari mahasiswa dan dibagi menjadi data pelatihan serta pengujian. Evaluasi performa mencakup akurasi, precision, recall, F1-score, serta confusion matrix. Hasil menunjukkan bahwa model VGGFace2 menghasilkan akurasi 98,75%, recall 100%, F1-score 99,26%, tanpa False Negative, sedangkan CASIA-WebFace menghasilkan akurasi 86,25% dengan recall 85,07%. Penelitian ini berkontribusi dalam menunjukkan keunggulan model pretrained VGGFace2 dalam sistem autentikasi wajah serta menegaskan efektivitas *transfer learning* untuk implementasi sistem yang akurat dengan sumber daya terbatas.

Kata kunci: Pengenalan Wajah, Sistem Autentikasi, InceptionResNetV1, *Transfer Learning*, VGGFace2

Abstract

Facial recognition has become a prominent biometric authentication method due to its high security and user convenience. However, training models from scratch requires large datasets and high computational resources. This study aims to optimize a face authentication system using the InceptionResNetV1 architecture with a transfer learning approach from the pretrained VGGFace2 model, and to compare its performance with CASIA-WebFace. Face detection is performed using YOLOv8, facial embeddings are generated using InceptionResNetV1, and authentication is done by calculating Euclidean distances between embeddings. Face data were collected from students and split into training and testing sets. Evaluation metrics include accuracy, precision, recall, F1-score, and confusion matrix. Results show that the VGGFace2 model achieved 98.75% accuracy, 100% recall, and 99.26% F1-score, with zero False Negatives,

while CASIA-WebFace reached only 86.25% accuracy with 85.07% recall. This research contributes by demonstrating the superiority of the VGGFace2 pretrained model in facial authentication systems and highlighting the effectiveness of transfer learning in achieving high accuracy with limited resources..

Keywords: *Face Recognition, Authentication System, InceptionResNetV1, Transfer Learning, VGGFace2*

1 PENDAHULUAN

Autentikasi biometrik kini menjadi pendekatan yang semakin diandalkan dalam meningkatkan sistem keamanan, terutama karena keterbatasan metode tradisional seperti penggunaan kata sandi dan kartu identitas yang rentan terhadap pencurian, pemalsuan, dan penyalahgunaan [1][2]. Salah satu metode biometrik yang paling luas digunakan adalah pengenalan wajah. Teknologi ini memungkinkan identifikasi individu secara unik berdasarkan fitur wajah, sehingga banyak diadopsi dalam berbagai bidang seperti sistem keamanan [3], sistem pembayaran elektronik [4], serta sistem kehadiran digital [5].

Metode pengenalan wajah konvensional, seperti Eigenfaces [6] dan Fisherfaces [7] yang berbasis pada teknik Principal Component Analysis (PCA) dan Linear Discriminant Analysis (LDA), telah banyak digunakan di masa awal pengembangan sistem ini. Namun, pendekatan tersebut memiliki keterbatasan dalam menangani variasi pencahayaan, pose wajah, dan ekspresi, yang seringkali muncul dalam kondisi nyata [10][11]. Seiring berkembangnya teknologi kecerdasan buatan, penggunaan arsitektur *deep learning* terutama *Convolutional Neural Network* (CNN) menjadi solusi yang lebih adaptif. CNN memiliki kemampuan untuk mengekstraksi fitur wajah secara mendalam dan representatif, sehingga lebih akurat dan robust terhadap berbagai kondisi visual [10] [11]. Salah satu pendekatan yang semakin populer dalam pemanfaatan CNN adalah penggunaan pre-trained model, yaitu model yang sebelumnya telah dilatih pada dataset besar dan kompleks [12]. Model ini mampu membawa pengetahuan awal dari data latih sebelumnya, sehingga proses pelatihan menjadi lebih efisien dan hasil yang diperoleh lebih stabil serta akurat, meskipun diterapkan pada domain yang berbeda.

Meskipun CNN terbukti efektif, penerapannya memerlukan dataset besar [13] dan daya komputasi tinggi [14], yang menjadi tantangan tersendiri dalam implementasi sistem autentikasi berbasis wajah di lingkungan dengan keterbatasan perangkat keras [15]. *Transfer learning* hadir sebagai alternatif efisien untuk mengatasi masalah tersebut [16]. Dengan menggunakan bobot dari model yang telah dilatih pada dataset besar, seperti VGGFace2, proses pelatihan dapat dipercepat dan akurasi tetap terjaga meskipun menggunakan dataset yang lebih kecil [17]. Dataset VGGFace2 sendiri mencakup lebih dari 3 juta gambar wajah dari berbagai individu dengan keragaman pose dan pencahayaan, menjadikannya sangat representatif dalam pelatihan model pengenalan wajah [18].

Berbagai penelitian telah memanfaatkan pretrained VGGFace2 pada arsitektur model yang berbeda dan menunjukkan hasil yang baik dalam tugas pengenalan wajah [19] [20] [21]. Namun, masih jarang ditemukan studi yang secara khusus menguji efektivitas arsitektur InceptionResNetV1 dengan pretrained VGGFace2 dalam konteks autentikasi pengguna berbasis wajah. Padahal, arsitektur InceptionResNetV1 menawarkan keunggulan dalam efisiensi pemrosesan dan kestabilan performa yang tinggi, menjadikannya kandidat ideal untuk sistem autentikasi *real-time*.

Penelitian ini bertujuan untuk mengeksplorasi dan mengoptimalkan sistem autentikasi wajah berbasis web dengan memanfaatkan arsitektur InceptionResNetV1 dan *transfer learning* dari VGGFace2. Tujuan spesifik dari penelitian ini meliputi: (1) Mengimplementasikan dan mengevaluasi model InceptionResNetV1 untuk autentikasi wajah, (2) Membandingkan performa model pretrained dari VGGFace2 dan CASIA-WebFace untuk mengukur efektivitas keduanya, dan (3) Melakukan evaluasi kinerja menggunakan confusion matrix sebagai indikator performa.

Kontribusi dari penelitian ini diharapkan mampu memberikan solusi autentikasi wajah yang tidak hanya akurat, tetapi juga efisien secara komputasi, serta dapat diimplementasikan secara luas pada sistem berbasis web. Selain itu, penelitian ini juga memberikan gambaran empiris terkait pemanfaatan pretrained model yang optimal untuk tugas autentikasi wajah, yang dapat

menjadi referensi untuk penelitian lanjutan maupun pengembangan sistem biometrik di masa depan.

2 TINJAUAN PUSTAKA

Penelitian mengenai sistem autentikasi berbasis wajah dengan dukungan teknologi *deep learning* telah mengalami kemajuan yang signifikan dalam beberapa tahun terakhir. Salah satu titik balik penting dalam bidang ini adalah diperkenalkannya model FaceNet oleh Schroff et al. [22], yang memperkenalkan pendekatan triplet loss untuk menghasilkan representasi wajah dalam bentuk embedding vektor. Pendekatan ini memungkinkan sistem untuk memetakan wajah ke dalam ruang vektor secara efisien, dengan menjaga jarak antar identitas yang konsisten. Hal ini menjadikan FaceNet sangat andal untuk keperluan autentikasi, di mana kemampuan membedakan individu secara akurat menjadi krusial. Perkembangan ini kemudian diperkuat oleh studi-studi lanjutan dari Anwarul dkk. [23], yang menunjukkan efektivitas *pendekatan convolutional neural network* (CNN) dalam membentuk *embedding* wajah dengan tingkat akurasi yang tinggi. Penerapan metode ini terbukti unggul dalam berbagai benchmark terkenal seperti Labeled Faces in the Wild (LFW) memperkuat posisi embedding berbasis CNN sebagai pendekatan utama dalam sistem pengenalan wajah modern.

CNN telah digunakan secara luas tidak hanya dalam identifikasi identitas, tetapi juga untuk klasifikasi ekspresi wajah. Sebagai contoh, Riyantoko dkk. [24]memanfaatkan kombinasi CNN dan Haar-Cascade untuk klasifikasi tujuh jenis ekspresi wajah pada dataset FER2013, menunjukkan bahwa CNN mampu mengekstraksi fitur visual secara efektif meskipun dalam kondisi pencahayaan dan ekspresi yang bervariasi. Pendekatan ini menunjukkan relevansi tinggi dalam konteks autentikasi wajah, yang juga memerlukan ketahanan terhadap variabilitas ekspresi.

Penelitian lain oleh Diyasa dkk. [25] mengembangkan sistem multi-face recognition untuk deteksi narapidana di ruang tahanan dengan memanfaatkan CNN dan Haar Cascade Classifier. Meskipun ditujukan untuk pengawasan keamanan, pendekatan ini relevan untuk pengenalan wajah secara real-time, terutama dalam hal deteksi multi-individu, pengelolaan sudut pandang, serta pencahayaan yang tidak merata. Temuan tersebut menegaskan pentingnya sistem yang mampu mempertahankan akurasi tinggi meskipun dalam kondisi lingkungan yang menantang.

Dataset pelatihan memegang peranan krusial dalam kinerja sistem pengenalan wajah. Salah satu dataset yang menonjol adalah VGGFace2, yang dikembangkan oleh Cao et al. [26]. VGGFace2 dikenal karena keragamannya yang luas, mencakup berbagai ekspresi wajah, kondisi pencahayaan, serta rentang usia subjek yang lebih bervariasi. Penelitian seperti yang dilakukan oleh Daoud et al. [27] menunjukkan bahwa VGGFace2 tidak hanya unggul dalam keragaman data, tetapi juga dalam mendukung pengembangan model pengenalan wajah yang lebih robust. Dataset ini memungkinkan pelatihan model deep learning untuk menghasilkan representasi wajah yang lebih akurat, bahkan dalam kondisi pencahayaan yang sulit atau variasi pose yang ekstrem. Dengan lebih dari 3,31 juta gambar dari 9.131 subjek, VGGFace2 memberikan fondasi yang kuat untuk meningkatkan kinerja sistem pengenalan wajah modern, menjadikannya salah satu dataset utama dalam penelitian di bidang ini.

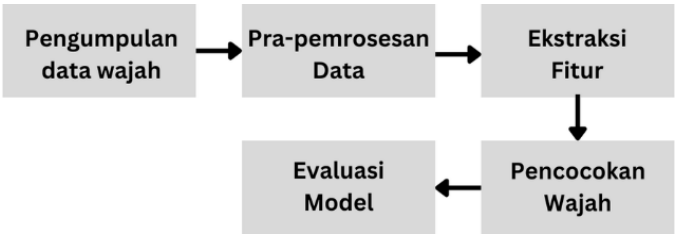
Model InceptionResNetV1, yang menggabungkan arsitektur Inception dengan ResNet, telah terbukti sangat efektif dalam menangani kompleksitas data wajah sekaligus menjaga efisiensi komputasi. Penelitian oleh Hidayati et al. [28] menunjukkan bahwa model Inception-ResNet mampu mencapai akurasi yang tinggi meskipun memiliki ukuran parameter yang lebih kecil dibandingkan dengan arsitektur lain seperti ResNet-101 atau VGG16. Keunggulan ini menjadikan Inception-ResNet pilihan yang menarik dalam pengenalan bentuk wajah, terutama untuk aplikasi yang membutuhkan keseimbangan antara efisiensi dan kinerja. Namun, penelitian ini lebih banyak berfokus pada pengujian konvensional seperti klasifikasi atau verifikasi bentuk wajah, yang menggunakan data uji terpisah. Celah penelitian yang diidentifikasi adalah kurangnya eksplorasi terhadap performa model dalam konteks dunia nyata, seperti autentikasi wajah, di mana pengguna unik direpresentasikan oleh satu atau lebih gambar wajah.

Berdasarkan celah tersebut, penelitian ini berfokus pada penerapan model InceptionResNetV1 yang dilatih menggunakan dua dataset besar, yaitu VGGFace2 dan CASIA-WebFace, untuk menghasilkan *embedding* wajah yang digunakan dalam proses autentikasi.

Artikel ini secara khusus mengeksplorasi pengaruh dataset pretraining terhadap kualitas embedding dengan pendekatan pengujian menggunakan metrik *Euclidean distance*. Selain mengevaluasi akurasi, penelitian ini juga mempertimbangkan efisiensi identifikasi dalam kondisi data terbatas, serta ketahanan terhadap variasi wajah yang umum terjadi dalam skenario nyata. Dengan demikian, studi ini mengisi kekosongan yang belum banyak dibahas dalam literatur sebelumnya terkait optimalisasi sistem autentikasi wajah berbasis *embedding*.

3 METODE PENELITIAN

Proses pengenalan wajah dalam penelitian ini dijelaskan secara sistematis pada Gambar 1, yang mencakup tahapan utama mulai dari pengumpulan data, pra-pemrosesan, ekstraksi fitur, hingga pencocokan dan evaluasi model.



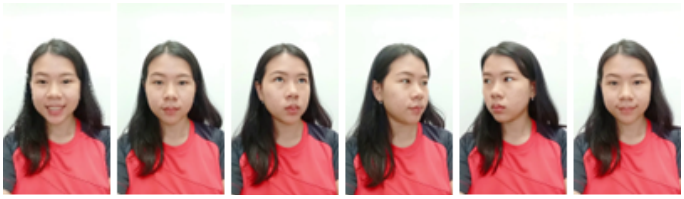
Gambar 1. Alur penelitian secara umum

a. Pengumpulan Data

Penelitian ini menggunakan data wajah dari 67 mahasiswa Universitas Pembangunan Nasional Veteran Jawa Timur. Proses pengumpulan data dilakukan melalui perekaman video wajah selama 5–10 detik per individu, mencakup berbagai sudut pandang seperti frontal, samping kiri, samping kanan, atas, dan bawah. Pendekatan ini bertujuan untuk menangkap variasi ekspresi dan sudut pandang guna meningkatkan robustnes model terhadap perubahan pose wajah [29].

Perekaman dilakukan menggunakan kamera depan ponsel Oppo Reno 3 dengan resolusi 44 MP, yang memungkinkan pengambilan citra wajah dengan detail tinggi dalam berbagai kondisi pencahayaan alami. Setelah proses perekaman, video diubah menjadi kumpulan frame citra statis dengan interval setiap 10 frame. Dari masing-masing video, diperoleh sekitar 15 hingga 25 citra per individu, tergantung pada durasi dan variasi ekspresi yang terekam. Selanjutnya dilakukan seleksi kualitas citra, dengan menghilangkan gambar yang buram atau mengandung noise signifikan. Proses ini bertujuan untuk memastikan bahwa hanya gambar dengan tampilan wajah yang jelas yang digunakan dalam tahap pelatihan dan pengujian model.

Sebagai bagian dari evaluasi sistem, peneliti juga menambahkan 13 citra wajah dari individu yang tidak termasuk dalam database pelatihan, atau sekitar 20% dari total subjek. Citra tambahan ini digunakan sebagai data negatif untuk menguji kemampuan sistem dalam mengenali wajah yang tidak terdaftar, yang penting dalam konteks autentikasi dan keamanan [30]. Contoh hasil pengumpulan data wajah yang digunakan dalam penelitian ini dapat dilihat pada Gambar 2.



Gambar 2. Contoh data wajah

b. Pra-pemrosesan Data

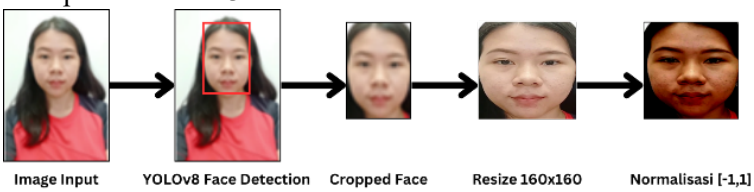
Setelah data diperoleh dalam bentuk gambar, langkah selanjutnya adalah membagi dataset menjadi training set dan testing set. Proses pembagian dilakukan secara manual, di mana satu gambar dari setiap individu dipilih sebagai data uji (*testing set*), sedangkan gambar lainnya digunakan untuk pelatihan model (*training set*). Strategi ini bertujuan untuk memastikan bahwa setiap individu memiliki representasi dalam proses evaluasi, sehingga memungkinkan pengujian performa model secara individual [31].

Setiap gambar wajah kemudian melalui tahapan pra-pemrosesan guna memastikan kualitas data optimal sebelum digunakan dalam proses pelatihan dan pengujian. Tahapan pertama adalah deteksi dan pemotongan wajah (*face cropping*) yang dilakukan secara otomatis menggunakan model YOLOv8 Face Detection. Proses ini mendeteksi lokasi wajah dalam citra dan memotong area relevan, sehingga hanya bagian wajah yang digunakan dalam proses berikutnya. Pemotongan ini penting untuk mengeliminasi latar belakang yang tidak diperlukan dan meningkatkan fokus ekstraksi fitur [32]. Setelah dipotong, seluruh gambar diubah ukurannya menjadi 160×160 piksel.

Untuk meningkatkan kestabilan model dalam menangani variasi pencahayaan dan perbedaan kontras, dilakukan normalisasi piksel pada setiap gambar. Normalisasi dilakukan dengan rumus (1) berikut:

$$N = \frac{image - 0.5}{0.5} \quad (1)$$

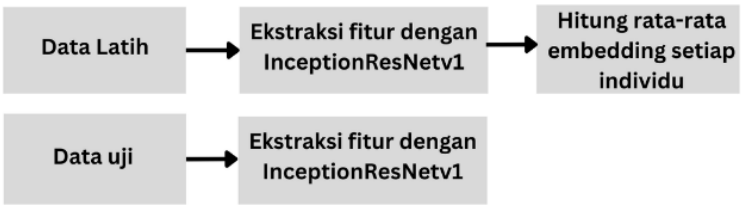
Rumus ini mengubah rentang nilai piksel gambar dari [0, 1] menjadi [-1, 1], yang merupakan prasyarat umum dalam pelatihan jaringan saraf untuk mempercepat konvergensi dan meningkatkan akurasi model [33]. Dengan menerapkan tahapan pra-pemrosesan ini, data wajah menjadi lebih terstruktur, bersih, dan konsisten, sehingga siap digunakan dalam tahap pelatihan dan evaluasi sistem autentikasi berbasis pengenalan wajah. Visualisasi dari hasil pra-pemrosesan tersebut dapat dilihat pada Gambar 3.



Gambar 3. Pra-proses Data

c. Ekstraksi Fitur

Tahap ekstraksi fitur merupakan proses utama dalam sistem pengenalan wajah yang bertujuan untuk memperoleh representasi numerik dari setiap wajah dalam bentuk embedding [34]. Pada penelitian ini, proses ekstraksi dilakukan menggunakan arsitektur InceptionResNetV1 yang telah dilatih sebelumnya (*pretrained*) pada dataset berskala besar. Proses ekstraksi fitur yang diterapkan ditampilkan pada Gambar 4.



Gambar 4. Proses ekstraksi fitur

Dalam tahap ini, terdapat perbedaan perlakuan antara data latih dan data uji. Pada data latih, setelah fitur wajah diekstraksi, semua embedding dari satu individu dirata-ratakan untuk memperoleh satu vektor representatif yang stabil. Pendekatan ini bertujuan untuk mengurangi variabilitas akibat pencahayaan, sudut pandang, atau ekspresi wajah dalam gambar pelatihan. Sebaliknya, pada data uji, proses ekstraksi dilakukan secara independen untuk setiap gambar tanpa dirata-ratakan. Hal ini mensimulasikan kondisi autentikasi di dunia nyata, di mana sistem harus mampu mengenali seseorang dari satu gambar saja.

1) Arsitektur InceptionResNetV1

InceptionResNetV1 merupakan arsitektur jaringan saraf tiruan yang menggabungkan dua pendekatan canggih, yakni Inception dan Residual Network (ResNet). Model ini dikembangkan untuk mengatasi tantangan dalam ekstraksi fitur kompleks dari citra, khususnya dalam tugas pengenalan wajah [35]. Arsitektur Inception pertama kali diperkenalkan melalui GoogLeNet (InceptionV1) dengan konsep utama penggunaan beberapa jenis filter konvolusi secara paralel dalam satu modul. Pendekatan ini memungkinkan model untuk menangkap informasi dari berbagai skala dalam gambar, sehingga meningkatkan

efisiensi dalam hal memori dan daya komputasi. Dengan menggunakan kombinasi beberapa ukuran filter, seperti 1x1, 3x3, dan 5x5 secara bersamaan, arsitektur ini mampu mengekstrak fitur-fitur kompleks dari citra pada berbagai tingkat representasi.

Sementara itu, ResNet dikembangkan untuk mengatasi permasalahan degradasi performa yang sering terjadi ketika jaringan saraf menjadi semakin dalam. Dengan memperkenalkan mekanisme skip connections atau shortcut connections, ResNet memungkinkan gradien dapat mengalir lebih mudah selama proses propagasi mundur, sehingga mengatasi masalah vanishing gradient yang sering muncul dalam pelatihan jaringan dalam. Dengan adanya shortcut connections ini, model dapat belajar lebih cepat dan mempertahankan informasi penting dari lapisan sebelumnya.

InceptionResNetV1 menggabungkan arsitektur Inception dengan mekanisme skip connection dari ResNet untuk meningkatkan efisiensi dan stabilitas dalam pelatihan. Setiap modul dalam InceptionResNetV1 terdiri dari beberapa cabang konvolusi dengan ukuran filter yang berbeda, yang bertujuan untuk menangkap fitur multi-skala. Selain itu, skip connection dalam arsitektur ini memungkinkan output dari satu lapisan untuk dilewatkan langsung ke lapisan yang lebih jauh di jaringan tanpa melalui semua lapisan perantara, sehingga meminimalkan kehilangan informasi selama propagasi mundur. Dengan kombinasi ini, model mampu mempertahankan keunggulan Inception dalam menangkap fitur multi-skala, sekaligus memastikan kestabilan aliran gradien melalui mekanisme residual learning dari ResNet.

Secara lebih rinci, arsitektur InceptionResNetV1 terdiri dari beberapa tahap utama, yaitu layer konvolusi awal, blok residual Inception (Block35, Block17, Block8), lapisan reduksi (Mixed_6a dan Mixed_7a), serta lapisan akhir yang menghasilkan embedding wajah berdimensi 512. Proses ekstraksi fitur diawali dengan serangkaian lapisan konvolusi dan pooling untuk menangkap fitur dasar dari citra wajah. Selanjutnya, fitur tersebut diproses melalui blok residual Inception, yang terdiri dari beberapa cabang konvolusi yang berjalan secara paralel dan dikombinasikan dengan mekanisme skip connection. Kemudian, lapisan reduksi digunakan untuk menyesuaikan dimensi fitur sebelum masuk ke blok Inception berikutnya. Pada tahap akhir, model menerapkan fully connected layer, batch normalization, dan dropout sebelum menghasilkan embedding wajah berdimensi 512 yang dapat digunakan untuk tugas klasifikasi atau verifikasi identitas wajah. Setiap lapisan pada InceptionResNetV1 dijabarkan secara terperinci dalam Tabel 1.

Tabel 1. Arsitektur Lapisan InceptionResNetV1

Layer	Details
Conv2D_1a	3x3, 32 filters, stride 2
Conv2D_2a	3x3, 32 filters, stride 1
Conv2D_2b	3x3, 64 filters, stride 1, padding 1
MaxPool_3a	3x3, stride 2
Conv2D_3b	1x1, 80 filters, stride 1
Conv2D_4a	3x3, 192 filters, stride 1
Conv2D_4b	3x3, 256 filters, stride 2
Block35 (x5)	Scale=0.17
Mixed_6a	Reduction block
Block17 (x10)	Scale=0.10
Mixed_7a	Reduction block
Block8 (x5)	Scale=0.20
Final Conv2D	1x1, 1792 filters
AvgPool	8x8
Dropout	p=0.6
Linear	512 units
BatchNorm	512 features

Dengan arsitektur ini, model mampu menghasilkan representasi wajah yang lebih kaya dan robust, serta telah terbukti efektif dalam berbagai aplikasi pengenalan wajah.

2) Dataset Pretrained: VGGFace2 dan CASIA-WebFace

Pada penelitian ini menggunakan bobot pretrained dari VGGFace2 dan CASIA-WebFace untuk menerapkan transfer learning dalam menghasilkan embedding wajah. VGGFace2 terdiri dari 3.310.000 gambar dari 9.131 individu, sedangkan CASIA-WebFace mencakup 494.414 gambar dari 10.575 individu. Kedua dataset ini sering digunakan dalam penelitian pengenalan wajah karena cakupan dan variasinya yang luas.

Penelitian ini membandingkan pengaruh transfer learning dari kedua dataset terhadap akurasi model yang dihasilkan. Evaluasi dilakukan untuk mengetahui sejauh mana perbedaan dataset pretrained memengaruhi kualitas embedding wajah dan performa autentikasi yang diperoleh.

3) Perhitungan Rata-rata Embedding

Untuk mengatasi hal ini, dilakukan perhitungan rata-rata embedding bagi setiap individu. Caranya adalah dengan menghitung rata-rata dari seluruh embedding yang dimiliki oleh individu tersebut dalam data latih. Dengan metode ini, setiap individu memiliki satu vektor representasi yang lebih stabil, yang dapat mengurangi pengaruh variasi seperti perubahan pencahayaan dan sudut wajah.

Sebaliknya, pada data uji, embedding dihitung langsung dari masing-masing gambar tanpa perhitungan rata-rata, sehingga model diuji dalam kondisi yang lebih realistis. Dengan pendekatan ini, sistem pengenalan wajah dapat melakukan identifikasi dengan lebih konsisten dan akurat, karena setiap individu memiliki satu embedding yang mencerminkan ciri khasnya secara menyeluruh.

d. Pencocokan Wajah

Dalam sistem pengenalan wajah, pencocokan identitas dilakukan dengan menghitung jarak antara dua vektor embedding menggunakan Euclidean Distance [36]. Pendekatan ini mengukur tingkat kesamaan antara dua fitur wajah yang telah diekstrak menggunakan model deep learning, seperti InceptionResNetV1.

1) Euclidean Distance

Euclidean Distance digunakan untuk mengukur jarak antara dua vektor dalam ruang berdimensi tinggi [37]. Semakin kecil jaraknya, semakin tinggi kemiripannya. Secara matematis, Euclidean Distance antara dua vektor A dan B dengan dimensi n dapat dihitung menggunakan rumus (2) sebagai berikut:

$$d(A,B) = \sqrt{\sum_{i=1}^{n}(A_i - B_i)^2} \quad (2)$$

di mana:

A: vektor embedding wajah yang ingin dikenali,

B: vektor embedding wajah yang telah terdaftar dalam database,

n: jumlah dimensi embedding.

Jika nilai Euclidean Distance kecil, maka wajah baru dianggap mirip dengan wajah dalam database. Sebaliknya, jika jarak Euclidean besar, wajah dianggap berbeda [38].

2) Proses Pencocokan Wajah

Proses pengenalan wajah dilakukan dengan mencocokkan embedding wajah yang diuji dengan embedding yang telah tersimpan dalam database. Pencocokan ini dilakukan dengan menghitung jarak Euclidean antara embedding wajah baru dan embedding yang telah terdaftar sebelumnya.

Setelah embedding wajah diperoleh, sistem membandingkannya dengan seluruh embedding dalam database menggunakan Euclidean Distance. Perhitungan ini mengukur seberapa mirip wajah yang diuji dengan data yang ada. Jika jarak terkecil yang ditemukan lebih kecil dari ambang batas tertentu, sistem menganggap wajah tersebut cocok dengan salah satu identitas dalam

database. Sebaliknya, jika tidak ada wajah dengan jarak di bawah ambang batas, sistem akan mengembalikan hasil "No match found", menandakan bahwa tidak ada kecocokan yang ditemukan.

e. Evaluasi Model

Setelah proses pencocokan wajah selesai, langkah berikutnya adalah mengevaluasi kinerja model untuk memastikan efektivitasnya dalam mengidentifikasi wajah dengan akurat. Evaluasi ini dilakukan dengan menghitung sejumlah metrik yang mencerminkan performa sistem dalam mengklasifikasikan wajah dengan benar maupun salah.

Kinerja model diukur menggunakan empat kategori utama dalam face matching, yaitu:

a) True Positive (TP): Model berhasil mengidentifikasi wajah yang benar-benar ada dalam database sebagai kecocokan.
b) True Negative (TN): Model dengan benar mengidentifikasi bahwa wajah yang tidak ada dalam database memang tidak memiliki kecocokan.
c) False Positive (FP): Model salah mengidentifikasi wajah yang sebenarnya tidak ada dalam database sebagai kecocokan.
d) False Negative (FN): Model gagal mengidentifikasi wajah yang sebenarnya ada dalam database.

Berdasarkan hasil klasifikasi tersebut, dihitung beberapa metrik evaluasi yang digunakan untuk menggambarkan performa sistem pengenalan wajah. Metrik-metrik tersebut meliputi akurasi, precision, recall, dan F1-score, yang masing-masing dirumuskan pada rumus 3 hingga rumus (6) :

a) Akurasi

$$Akurasi = \frac{TP + TN}{TP + TN + FP + FN} \quad (3)$$

b) Precision

$$Precision = \frac{TP}{TP + FP} \quad (4)$$

c) Recall

$$Recall = \frac{TP}{TP + FN} \quad (5)$$

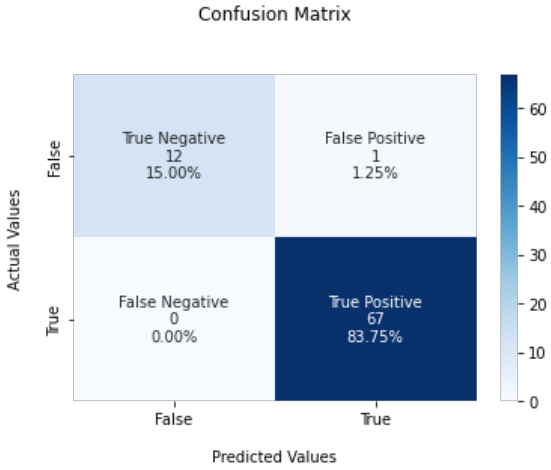
d) F1-Score

$$F1 = \frac{2 \times precision \times recall}{precision + recall} \quad (6)$$

4 HASIL DAN PEMBAHASAN

a. Evaluasi Model dengan Pretrained VGGFace2

Hasil evaluasi model yang menggunakan pretrained VGGFace2 menunjukkan performa yang optimal dalam mengenali wajah yang terdaftar di dalam database. Berdasarkan confusion matrix pada Gambar 5, model menghasilkan 67 True Positives (TP), yang berarti 67 wajah dikenali dengan benar sebagai individu yang ada dalam database. Tidak terdapat kesalahan klasifikasi dalam bentuk False Negative (FN = 0), yang menunjukkan bahwa model ini tidak melewatkan satu pun wajah yang seharusnya dikenali.

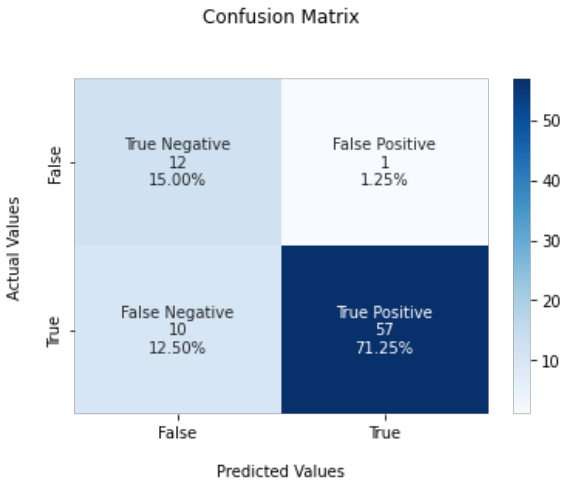


Gambar 5. Hasil confusion matrix dengan menggunakan pretrained model VGGFace2

Selain itu, model juga mampu mengidentifikasi wajah yang tidak ada dalam database dengan baik, sebagaimana ditunjukkan oleh 12 True Negatives (TN). Namun, terdapat 1 False Positive (FP), yang berarti model keliru mengidentifikasi satu wajah yang seharusnya tidak ada dalam database sebagai wajah yang cocok. Dengan akurasi tinggi dan tidak adanya False Negative, pretrained VGGFace2 menunjukkan keunggulan dalam memastikan bahwa wajah yang dikenali benar-benar sesuai dengan identitas yang ada. Hal ini mengindikasikan bahwa model memiliki sensitivitas tinggi dan tidak melewatkan wajah yang seharusnya dikenali.

b. Evaluasi Model dengan Pretrained CASIA-WebFace

Sementara itu, model yang dilatih menggunakan pretrained CASIA-WebFace menunjukkan performa yang lebih rendah dibandingkan VGGFace2. Hasil evaluasi menunjukkan bahwa model berhasil mengidentifikasi 57 wajah dengan benar (TP = 57), tetapi terdapat 10 kasus False Negative (FN = 10), yang berarti model gagal mengenali 10 wajah yang sebenarnya ada dalam database. Berdasarkan confusion matrix yang dapat dilihat pada Gambar 6, model ini juga menunjukkan performa yang lebih rendah dalam mendeteksi wajah yang terdaftar.



Gambar 6. Hasil confusion matrix dengan menggunakan pretrained model Casia-WebFace

Dari sisi deteksi wajah yang tidak terdaftar dalam database, model memiliki performa yang sama dengan pretrained VGGFace2, yaitu 12 True Negatives (TN) dan 1 False Positive (FP). Namun, keberadaan False Negative yang lebih tinggi menunjukkan bahwa model dengan pretrained CASIA-WebFace kurang optimal dalam mengenali wajah yang seharusnya dapat dikenali. Perbedaan utama antara kedua model ini terletak pada kemampuan mendeteksi wajah yang terdaftar dalam database. Model dengan pretrained VGGFace2 memiliki tingkat sensitivitas yang lebih tinggi dibandingkan CASIA-WebFace, sebagaimana ditunjukkan oleh jumlah False Negative yang lebih rendah.

c. Perbandingan Performa Model

Untuk mendapatkan gambaran yang lebih jelas mengenai performa kedua model, dilakukan perhitungan metrik evaluasi utama, seperti akurasi, precision, recall, specificity, dan F1-score. Table 2 menyajikan hasil evaluasi model berdasarkan metrik yang telah dihitung:

Tabel 2. Perbandingan performa dari kedua model

	InceptionResNetV1-VGGFace2	InceptionResNetV1-CASIA-WebFace
Akurasi	0.9875	0.8625
Precision	0.9853	0.9828
Recall	1	0.8507
F1 Score	0.9926	0.9119

Berdasarkan hasil di atas, pretrained VGGFace2 menunjukkan performa yang lebih unggul dibandingkan CASIA-WebFace, terutama dalam hal recall yang mencapai 100%, menandakan model tidak melewatkan satu pun wajah yang seharusnya dikenali. Sebaliknya, model CASIA-WebFace memiliki recall yang lebih rendah (85,1%), yang berarti masih terdapat wajah yang tidak dapat dikenali dengan baik.

Dari aspek precision dan specificity, kedua model memiliki performa yang hampir serupa, menunjukkan bahwa keduanya cukup andal dalam mengenali wajah yang tidak ada dalam database. Namun, perbedaan signifikan dalam recall mengindikasikan bahwa VGGFace2 lebih cocok untuk aplikasi yang memerlukan sensitivitas tinggi, seperti sistem autentikasi berbasis wajah.

5 KESIMPULAN

Penelitian ini bertujuan untuk mengoptimalkan sistem autentikasi berbasis pengenalan wajah dengan memanfaatkan model InceptionResNetV1 yang dilatih ulang menggunakan dataset VGGFace2. Sistem ini dirancang untuk meningkatkan keamanan proses login dengan mengidentifikasi pengguna terdaftar secara akurat melalui fitur wajah. Hasil evaluasi menunjukkan bahwa sistem yang diusulkan memiliki performa yang sangat baik, dengan perolehan 67 True Positives (TP), 0 False Negatives (FN), 12 True Negatives (TN), dan hanya 1 False Positive (FP), sehingga menghasilkan tingkat akurasi sebesar 98,75%. Tidak adanya False Negative menunjukkan sensitivitas yang tinggi serta keandalan sistem dalam mengenali pengguna yang memang terdaftar. Temuan ini memperkuat bahwa model VGGFace2 pretrained mampu menghasilkan embedding wajah yang kuat dan diskriminatif, yang berkontribusi besar terhadap optimasi proses pengenalan wajah.

Untuk pengembangan selanjutnya, sistem ini dapat diuji lebih lanjut dalam skenario waktu nyata (real-time) dengan menggunakan input kamera secara langsung guna mengukur performanya dalam kondisi dinamis. Selain itu, pengujian terhadap variasi kondisi seperti penggunaan masker, kacamata, pencahayaan berbeda, dan ekspresi wajah dapat dilakukan untuk meningkatkan ketahanan sistem. Implementasi sistem ini dalam platform mobile maupun berbasis web juga merupakan langkah penting menuju penerapan nyata. Penelitian di masa depan juga dapat mengeksplorasi arsitektur deep learning lain atau dataset pretrained berskala besar untuk lebih meningkatkan akurasi dan generalisasi sistem terhadap populasi yang lebih luas.

DAFTAR PUSTAKA

[1] P. S. B. Bele, G. S. Band, S. S. Kawade, A. D. Udimkar, and R. N. Khandare, "The Role of Biometric and Authentication in Security," *Int. J. Res. Appl. Sci. Eng. Technol.*, vol. 12, no. November, pp. 1997–1999, 2024, [Online]. Available: https://www.ijraset.com/best-journal/the-role-of-biometric-and-authentication-in-security

[2] Divya Khandekar and Shailesh Bendale, "Biometrics Security System," *Int. J. Adv. Res. Sci. Commun. Technol.*, vol. 3, no. 1, pp. 219–224, Mar. 2023, doi: 10.48175/IJARSCT-8588.

[3] D. Tribuana, Hazriani, and A. L. Arda, "Face recognition for smart door security access with convolutional neural network method," *Telkomnika (Telecommunication Comput. Electron. Control.*, vol. 22, no. 3, pp. 702–710, 2024, doi: 10.12928/TELKOMNIKA.v22i3.25946.

[4] W. Hutamaputra and F. Utaminingrum, "Implementasi Facial Landmark dalam Pengenalan Wajah pada Sistem Pembayaran Elektronik," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 5, no. 5, pp. 2058–2064, 2021, [Online]. Available: http://j-ptiik.ub.ac.id

[5] K. Alhanaee, M. Alhammadi, N. Almenhali, and M. Shatnawi, "Face recognition smart attendance system using deep transfer learning," *Procedia Comput. Sci.*, vol. 192, pp. 4093–4102, 2021, doi: 10.1016/j.procs.2021.09.184.

[6] M. A. Turk and A. P. Pentland, "Face recognition using eigenfaces," in *Proceedings. 1991 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, IEEE Comput. Sco. Press, 1991, pp. 586–591. doi: 10.1109/CVPR.1991.139758.

[7] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. Fisherfaces: recognition using class specific linear projection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 7, pp. 711–720, Jul. 1997, doi: 10.1109/34.598228.

[8] S. Swapna Rani, "Enhancing Facial Recognition Accuracy in Low-Light Conditions Using Convolutional Neural Networks," *J. Electr. Syst.*, vol. 20, no. 5s, pp. 2140–2148, 2024, doi: 10.52783/jes.2559.

[9] T. Q. Chung, H. C. Huyen, and D. V. Sang, "A Novel Generative Model to Synthesize Face Images for Pose-invariant Face Recognition," in *2020 International Conference on Multimedia Analysis and Pattern Recognition (MAPR)*, IEEE, Oct. 2020, pp. 1–6. doi: 10.1109/MAPR49794.2020.9237763.

[10] T. A. Kadhim, N. S. Zghal, W. Hariri, and D. Ben Aissa, "Face Recognition in Multiple Variations Using Deep Learning and Convolutional Neural Networks," in *2022 IEEE 9th International Conference on Sciences of Electronics, Technologies of Information and Telecommunications (SETIT)*, IEEE, May 2022, pp. 305–311. doi: 10.1109/SETIT54465.2022.9875530.

[11] J. J. Sundi, H. Kumar, Bharti, and R. Bedi, "Real-Time Facial Expression Recognition Using Convolutional Neural Networks for Adaptive User Interfaces," in *2024 5th International Conference for Emerging Technology (INCET)*, IEEE, May 2024, pp. 1–6. doi: 10.1109/INCET61516.2024.10593062.

[12] T. M. Fahrudin and I. Z. A. Illah, "SkinMate: Mobile-Based Application for Detecting Multi-Class Skin Diseases Classification Using Pre-Trained MobileNetV2 on CNN Architecture," in *2023 IEEE 9th Information Technology International Seminar (ITIS)*, IEEE, Oct. 2023, pp. 1–6. doi: 10.1109/ITIS59651.2023.10420370.

[13] R. Liu, J. Jia, Y. Zhou, Y. Zhou, and Y. Liu, "Training Deep Neural Networks with Large-scale Datasets on Sunway High Performance Computer," in *2022 IEEE International Conference on Artificial Intelligence and Computer Applications (ICAICA)*, IEEE, Jun.

2022, pp. 466–471. doi: 10.1109/ICAICA54878.2022.9844581.

[14] T. Li, B. He, and Y. Zheng, "Research and Implementation of High Computational Power for Training and Inference of Convolutional Neural Networks," *Appl. Sci.*, vol. 13, no. 2, 2023, doi: 10.3390/app13021003.

[15] S. Ansari, "Computer Vision Modeling on the Cloud," in *Building Computer Vision Applications Using Artificial Neural Networks*, Berkeley, CA: Apress, 2023, pp. 457–512. doi: 10.1007/978-1-4842-9866-4_10.

[16] W. S. J. Saputra, E. Y. Puspaningrum, W. F. Syahputra, A. P. Sari, Y. V. Via, and M. Idhom, "Car Classification Based on Image Using Transfer Learning Convolutional Neural Network," in *2022 IEEE 8th Information Technology International Seminar (ITIS)*, IEEE, Oct. 2022, pp. 324–327. doi: 10.1109/ITIS57155.2022.10010073.

[17] A. H. Ali, M. G. Yaseen, M. Aljanabi, S. A. Abed, and C. GPT, "Transfer Learning: A New Promising Techniques," *Mesopotamian J. Big Data*, no. February, pp. 29–30, 2023, doi: 10.58496/mjbd/2023/004.

[18] L. Yu, N. Jiang, H. Xu, and Z. Qi, "Facial Expression Recognition Based on Improved VGG-face Model and Transfer Learning," in *Proceedings of the 2023 International Conference on Computer, Vision and Intelligent Technology*, New York, NY, USA: ACM, Aug. 2023, pp. 1–7. doi: 10.1145/3627341.3630376.

[19] M. Usgan, R. Ferdiana, and I. Ardiyanto, "Deep learning pre-trained model as feature extraction in facial recognition for identification of electronic identity cards by considering age progressing," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 1115, no. 1, p. 012009, Mar. 2021, doi: 10.1088/1757-899X/1115/1/012009.

[20] T. H. Tsai and P. T. Chi, "A single-stage face detection and face recognition deep neural network based on feature pyramid and triplet loss," 2022. doi: 10.1049/ipr2.12479.

[21] B. W. Mugalu, R. C. Wamala, J. Serugunda, and A. Katumba, "Face Recognition as a Method of Authentication in a Web-Based System," 2021, [Online]. Available: http://arxiv.org/abs/2103.15144

[22] F. Schroff, D. Kalenichenko, and J. Philbin, "FaceNet: A unified embedding for face recognition and clustering," in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Jun. 2015, pp. 815–823. doi: 10.1109/CVPR.2015.7298682.

[23] S. Anwarul, T. Choudhury, and S. Dahiya, "A novel hybrid ensemble convolutional neural network for face recognition by optimizing hyperparameters," *Nonlinear Eng.*, vol. 12, no. 1, Jun. 2023, doi: 10.1515/nleng-2022-0290.

[24] P. A. Riyantoko, Sugiarto, and K. M. Hindrayani, "Facial Emotion Detection Using Haar-Cascade Classifier and Convolutional Neural Networks," *J. Phys. Conf. Ser.*, vol. 1844, no. 1, p. 012004, Mar. 2021, doi: 10.1088/1742-6596/1844/1/012004.

[25] I. G. S. M. Diyasa, A. Fauzi, M. Idhom, and A. Setiawan, "Multi-face Recognition for the Detection of Prisoners in Jail using a Modified Cascade Classifier and CNN," *J. Phys. Conf. Ser.*, vol. 1844, no. 1, p. 012005, Mar. 2021, doi: 10.1088/1742-6596/1844/1/012005.

[26] Q. Cao, L. Shen, W. Xie, O. M. Parkhi, and A. Zisserman, "VGGFace2: A Dataset for Recognising Faces across Pose and Age," in *2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018)*, IEEE, May 2018, pp. 67–74. doi: 10.1109/FG.2018.00020.

[27] E. Al Daoud and G. Samara, "Improving the Face Recognition Performance Using Gabor

and VGGFace2 Features Concatenation," in *2022 6th International Conference on Information Technology (InCIT)*, IEEE, Nov. 2022, pp. 187–190. doi: 10.1109/InCIT56086.2022.10067669.

[28] S. C. Hidayati, J. Tasyanita, C. Fatichah, and Y. Anistyasari, "Understanding Human Face Shape via Inception-ResNet Neural Network Architecture," in *2023 International Conference on Advanced Mechatronics, Intelligent Manufacture and Industrial Automation (ICAMIMIA)*, IEEE, Nov. 2023, pp. 631–636. doi: 10.1109/ICAMIMIA60881.2023.10427897.

[29] O. Elharrouss, N. Almaadeed, and S. Al-Maadeed, "LFR face dataset:Left-Front-Right dataset for pose-invariant face recognition in the wild," in *2020 IEEE International Conference on Informatics, IoT, and Enabling Technologies (ICIoT)*, IEEE, Feb. 2020, pp. 124–130. doi: 10.1109/ICIoT48696.2020.9089530.

[30] Shahjahan, N. Sheikh, A. Rehman, and M. Anjum, "Development of Face Recognition System Using Support Vector Machine Algorithm," *Eur. J. Appl. Sci. Eng. Technol.*, vol. 2, no. 3, pp. 214–227, May 2024, doi: 10.59324/ejaset.2024.2(3).20.

[31] H. R. Maier *et al.*, "On how data are partitioned in model development and evaluation: Confronting the elephant in the room to enhance model generalization," *Environ. Model. Softw.*, vol. 167, p. 105779, Sep. 2023, doi: 10.1016/j.envsoft.2023.105779.

[32] N. S. Vemulapalli, P. Paladugula, G. S. Prabhat, S. Abhishek, and A. T, "Face Detection with Landmark using YOLOv8," in *2023 3rd International Conference on Emerging Frontiers in Electrical and Electronic Technologies (ICEFEET)*, IEEE, Dec. 2023, pp. 1–5. doi: 10.1109/ICEFEET59656.2023.10452204.

[33] L. Huang, J. Qin, Y. Zhou, F. Zhu, L. Liu, and L. Shao, "Normalization Techniques in Training DNNs: Methodology, Analysis and Application," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 45, no. 8, pp. 10173–10196, Aug. 2023, doi: 10.1109/TPAMI.2023.3250241.

[34] I. Y. Susrama, I. G. S. M. D., Putra, A. H., Ariefwan, M. R. M., Atnanda, P. A., Trianggraeni, F., & Purbasari, "Feature Extraction for Face Recognition Using Haar Cascade Classifier," 2022. doi: 10.11594/nstp.2022.2432.

[35] K. Cao, "Novel coronavirus pneumonia CT image classification based on inception ResNet," *IET Conf. Proc.*, vol. 2024, no. 21, pp. 134–138, Jan. 2025, doi: 10.1049/icp.2024.4214.

[36] N. Mardiana, R. D. Dana, Faisal, I. Farida, A. G. Azwar, and Nurwathi, "Similarity Measures Implementation on Face Authentication using Indonesian Citizen ID Card," in *2023 17th International Conference on Telecommunication Systems, Services, and Applications (TSSA)*, IEEE, Oct. 2023, pp. 1–5. doi: 10.1109/TSSA59948.2023.10366880.

[37] F. Tessari and N. Hogan, "Surpassing Cosine Similarity for Multidimensional Comparisons: Dimension Insensitive Euclidean Metric (DIEM)," 2024, [Online]. Available: http://arxiv.org/abs/2407.08623

[38] S. M. Sami, J. McCauley, S. Soleymani, N. Nasrabadi, and J. Dawson, "Benchmarking human face similarity using identical twins," *IET Biometrics*, vol. 11, no. 5, pp. 459–484, 2022, doi: 10.1049/bme2.12090.