P-ISSN: 2774-4574 ; E-ISSN: 363-4582
TRILOGI, 6(4), Okt-Desember 2025 (77-85)

DOI: 10.33650/trilogi.v6i4.13424

Keamanan Pengenalan Wajah Berbasis Deep Learning: Tinjauan Sistematis Serangan Adversarial dan Strategi Pertahanan (Systematic Literature Review)

Fahmy Syahputra, S.Kom., M.Kom
Universitas Negeri Medan, Indonesia
famybd@unimed.ac.id

Frans Jhonatan Saragi
Universitas Negeri Medan, Indonesia
franssaragi.5233151019@mhs.unimed.ac.id

Elsa Sabrina, S.Pd., M.Pd.T
Universitas Negeri Medan, Indonesia
elsasabrina@unimed.ac.id

Suci Nurrahma
Universitas Negeri Medan, Indonesia
sucinurahma15@mhs.unimed.ac.id

Andika Sitorus
Universitas Negeri Medan, Indonesia
andikasitorus.5233151021@mhs.unimed.ac.id

Novi Novanni Sinaga
Universitas Negeri Medan, Indonesia
novisinagaa.5233351032@mhs.unimed.ac.id

Khodijah May Nuri Lubis
Universitas Negeri Medan, Indonesia
khadijah.5231151020@mhs.unimed.ac.id

Abstract
Deep learning–based face recognition is widely adopted due to its strong performance, yet its susceptibility to attacks—particularly adversarial attacks—poses critical risks to the security and reliability of biometric systems. This study presents a **Systematic Literature Review (SLR)** to synthesize evidence on **performance, vulnerabilities, and defense strategies** in deep learning–based face recognition. The review follows **PRISMA** guidelines, including literature retrieval from reputable scholarly sources, deduplication, title/abstract screening, and full-text eligibility assessment based on predefined inclusion and exclusion criteria. Study quality is examined through critical appraisal, and findings are synthesized using **thematic analysis**, yielding four major themes: (1) model performance and factors influencing accuracy, (2) attack types and their impact on recognition outcomes, (3) defense mechanisms and their effectiveness, and (4) real-world deployment constraints (e.g., illumination, pose, image quality, and identity scale). The synthesis indicates that high accuracy does not necessarily imply high robustness; several defenses (e.g., adversarial training, attack detection, and robust learning) can improve resilience but may introduce trade-offs in computational cost and/or accuracy. This review provides a comparative synthesis and a conceptual model linking **accuracy–attacks–defenses**, and offers practical recommendations for model selection and security evaluation design. Limitations include heterogeneity in datasets and experimental protocols, inconsistent reporting metrics, and

potential publication bias.

Keywords: adversarial attacks; deep learning; face recognition; robustness; systematic literature review.

Abstrak

Pengenalan wajah berbasis deep learning banyak diadopsi karena performanya yang tinggi, namun kerentanannya terhadap serangan—terutama serangan adversarial—menimbulkan risiko pada keamanan dan keandalan sistem biometrik. Penelitian ini menyajikan tinjauan sistematis (Systematic Literature Review/SLR) untuk merangkum bukti ilmiah terkait kinerja, kerentanan, serta strategi pertahanan pada sistem pengenalan wajah berbasis deep learning. Proses kajian mengikuti pedoman PRISMA, meliputi pencarian literatur pada beberapa sumber ilmiah bereputasi, penghapusan duplikasi, penyaringan judul/abstrak, serta penilaian kelayakan teks penuh berdasarkan kriteria inklusi–eksklusi yang ditetapkan. Kualitas studi dievaluasi melalui penilaian kritis, dan sintesis dilakukan menggunakan analisis tematik yang mengelompokkan temuan ke dalam empat tema utama: (1) performa model dan faktor yang memengaruhi akurasi, (2) jenis serangan dan dampaknya terhadap performa pengenalan, (3) strategi pertahanan dan efektivitasnya, serta (4) tantangan penerapan real-world (mis. variasi pencahayaan, pose, kualitas citra, dan skala identitas). Hasil sintesis menunjukkan bahwa akurasi tinggi tidak selalu sejalan dengan ketahanan; sejumlah pendekatan pertahanan (mis. adversarial training, deteksi serangan, dan pembelajaran robust) dapat meningkatkan ketahanan namun berpotensi menimbulkan trade-off pada biaya komputasi dan/atau akurasi. Kajian ini merumuskan ringkasan komparatif serta model konseptual hubungan akurasi–serangan–pertahanan, dan memberikan rekomendasi praktis untuk pemilihan model serta rancangan evaluasi keamanan. Keterbatasan kajian mencakup heterogenitas dataset dan protokol eksperimen, variasi metrik pelaporan, serta potensi bias publikasi.

Katakunci: *deep learning; pengenalan wajah; robustness; serangan adversarial; tinjauan sistematis.*

1 Pendahuluan

Pengenalan wajah merupakan salah satu teknologi biometrik yang kini banyak dimanfaatkan dalam berbagai aplikasi keamanan, mulai dari sistem kontrol akses, autentikasi pembayaran digital, pengawasan keamanan publik, hingga verifikasi identitas di berbagai layanan daring (Al-dmour et al., 2023). Teknologi ini bekerja dengan mengidentifikasi atau memverifikasi identitas seseorang berdasarkan fitur-fitur unik yang terdapat pada wajah manusia. Dibandingkan dengan metode biometrik lainnya seperti *fingerprint* atau iris, pengenalan wajah memiliki keunggulan dalam hal kemudahan akuisisi(Maulidiansyah & Yaqin, 2023) data karena dapat dilakukan tanpa kontak fisik dan menggunakan perangkat kamera yang sudah tersedia secara luas pada *smartphone* dan sistem pengawasan (Chi et al., 2023).

Perkembangan teknologi *deep learning* (DL) telah membawa perubahan paradigma yang signifikan dalam peningkatan kinerja sistem pengenalan wajah (Deng et al., 2024). Arsitektur *Convolutional Neural Networks* (CNN) seperti VGG-16, VGG-19, ResNet-50, ResNet-101, dan *Inception* telah terbukti mampu mengekstraksi fitur-fitur wajah yang kompleks dan menghasilkan representasi yang lebih diskriminatif dibandingkan metode konvensional seperti *Principal Component Analysis* (PCA) atau *Linear Discriminant Analysis* (LDA) (Kristanto et al., 2023). Model *deep learning* modern seperti *FaceNet*, *DeepFace*, dan *VGGFace* telah mencapai akurasi yang mendekati atau bahkan melampaui kemampuan manusia dalam mengenali wajah pada kondisi ideal, dengan tingkat akurasi mencapai 99,33% pada berbagai *benchmark dataset* seperti *Labeled Faces in the Wild* (LFW) (Hangaragi et al., 2023; Yan et al., 2023).

Namun, kemajuan ini juga memunculkan tantangan baru yang signifikan, terutama terkait isu keamanan dan kerentanan sistem terhadap berbagai bentuk serangan. Berbagai penelitian terkini menunjukkan bahwa sistem pengenalan wajah berbasis *deep learning* rentan terhadap serangan *adversarial*, di mana penyerang dapat memanipulasi citra input dengan *noise* atau *patch* yang dirancang khusus untuk mengecoh model (Hung Hwang et al., 2023). Serangan *face*

morphing, yang menggabungkan fitur wajah dari dua individu berbeda, telah terbukti dapat melewati sistem verifikasi wajah pada paspor elektronik dan kontrol perbatasan (Long et al., 2022). Lebih lanjut, perkembangan teknologi *deepfake* yang memanfaatkan *Generative Adversarial Networks* (GAN) memungkinkan pembuatan video wajah sintetis yang sangat realistis, menimbulkan ancaman serius terhadap integritas sistem autentikasi dan verifikasi identitas (Ramachandran et al., 2021).

Tantangan keamanan ini semakin kompleks ketika sistem pengenalan wajah diimplementasikan pada lingkungan *real-world* yang tidak terkendali. Faktor-faktor seperti variasi pencahayaan, pose wajah yang tidak frontal, oklusi parsial, kualitas citra yang rendah, dan *noise* visual dapat menurunkan akurasi sistem secara drastis (Xing et al., 2025). Penelitian oleh Kristanto *et al.* (Kristanto et al., 2023) menunjukkan bahwa sistem berbasis PCA mengalami penurunan akurasi hingga 0% pada citra dengan permasalahan kecerahan ekstrem, meskipun dapat mencapai akurasi 80% pada citra dengan *blur*. Kondisi ini menciptakan *gap* antara performa sistem pada lingkungan laboratorium yang terkontrol dengan aplikasi di dunia nyata.

Berbagai penelitian telah dilakukan untuk mengatasi tantangan ini melalui berbagai pendekatan. Al-Dmour *et al.* (Al-dmour et al., 2023) mengembangkan sistem deteksi dan pengenalan wajah yang dapat berfungsi pada kondisi wajah tertutup masker, memanfaatkan arsitektur *deep learning* yang dioptimalkan untuk mengekstraksi fitur dari area wajah yang tersisa. Hwang *et al.* (Hung Hwang et al., 2023) mengusulkan mekanisme pertahanan berbasis GAN untuk menangkal serangan *adversarial patch*, yang mampu mendeteksi hampir semua serangan *dodging* dan lebih dari setengah serangan *impersonation*. Sementara itu, penelitian tentang deteksi *face morphing* menggunakan fitur tingkat *patch* dan jaringan *lightweight* menunjukkan hasil yang menjanjikan dalam mengidentifikasi citra wajah yang telah dimanipulasi, dengan kemampuan untuk beroperasi pada perangkat dengan sumber daya komputasi terbatas (Long et al., 2022).

Meskipun berbagai solusi telah diusulkan, masih terdapat kesenjangan dalam pemahaman komprehensif mengenai dampak keseluruhan algoritma *deep learning* terhadap keamanan sistem pengenalan wajah, terutama dalam konteks ancaman yang terus berkembang (Wang & Deng, 2021). Penelitian-penelitian sebelumnya cenderung fokus pada aspek spesifik seperti peningkatan akurasi atau pertahanan terhadap jenis serangan tertentu, namun kurang memberikan analisis integratif yang menghubungkan berbagai dimensi permasalahan.

Kesenjangan tersebut mengindikasikan perlunya satu kajian yang tidak hanya merangkum performa model, tetapi juga menilai ketahanan dan implikasi keamanan secara lintas penelitian. Dalam konteks pengenalan wajah, ukuran "baik" tidak cukup dimaknai sebagai akurasi tinggi pada dataset benchmark, melainkan juga mencakup robustitas ketika sistem menghadapi gangguan, manipulasi, dan variasi lingkungan yang realistis. Di sisi lain, perkembangan ancaman yang bergerak cepat—mulai dari adversarial examples, adversarial patch, face morphing, hingga deepfake—membuat literatur berkembang secara terfragmentasi: sebagian studi mengutamakan peningkatan akurasi, sebagian lain fokus pada skenario serangan tertentu, sementara studi pertahanan sering menggunakan asumsi dan protokol evaluasi yang berbeda-beda. Akibatnya, sulit untuk menarik kesimpulan yang konsisten terkait pertanyaan kunci seperti: kapan akurasi tinggi berbanding lurus dengan ketahanan; bagaimana karakteristik serangan tertentu memengaruhi hasil pengenalan (pengenalan identitas, verifikasi, atau autentikasi); serta sejauh mana strategi pertahanan meningkatkan keamanan tanpa menurunkan kegunaan sistem. Oleh karena itu, penelitian ini dirancang sebagai tinjauan yang bersifat integratif untuk menyatukan temuan-temuan tersebut ke dalam kerangka analisis yang seragam, sehingga pembaca dapat memahami peta riset keamanan pengenalan wajah berbasis deep learning secara komprehensif, bukan sekadar potongan hasil dari studi-studi yang berdiri sendiri.

Untuk menjawab kebutuhan tersebut, penelitian ini menerapkan **Systematic Literature Review (SLR)** dengan pedoman **PRISMA** guna memastikan proses seleksi studi berlangsung transparan, dapat ditelusuri, dan dapat direplikasi. Melalui tahapan identifikasi, penghapusan duplikasi, penyaringan judul/abstrak, serta telaah kelayakan teks penuh berdasarkan kriteria inklusi–eksklusi yang ditetapkan, kajian ini menetapkan **sebanyak 25 studi** sebagai korpus utama analisis. Fokus SLR ini diarahkan pada tiga dimensi yang saling terkait: (1) **kinerja/akurasi** model pengenalan wajah berbasis deep learning beserta faktor yang memengaruhinya (misalnya arsitektur, strategi ekstraksi fitur, dan kondisi data), (2) **kerentanan dan karakter serangan** (misalnya mekanisme manipulasi input, skenario ancaman, serta dampak terhadap metrik

performa), dan (3) **strategi pertahanan** (misalnya adversarial training, deteksi serangan, robust learning, atau pendekatan berbasis generatif) beserta konsekuensi implementasinya. Dengan ruang lingkup ini, SLR tidak berhenti pada "model mana paling akurat", melainkan menekankan "model mana yang paling layak diadopsi jika ancaman dan kondisi lapangan diperhitungkan". Selain itu, penelitian ini juga membatasi pembahasan pada konteks pengenalan wajah berbasis deep learning, sehingga kesimpulan yang dihasilkan tetap tajam dan tidak melebar ke domain biometrik lain yang memiliki karakter ancaman dan metrik berbeda.

Selanjutnya, untuk menghindari rangkuman yang hanya bersifat naratif per-studi, penelitian ini menggunakan **analisis tematik** guna menyintesiskan temuan secara lintas penelitian, sehingga pola dan tren dapat terlihat lebih jelas. Temuan akan dikelompokkan ke dalam tema-tema utama yang konsisten: performa model dan determinannya, tipologi serangan dan kondisi keberhasilannya, efektivitas pertahanan serta trade-off yang muncul, dan tantangan real-world yang menjembatani perbedaan antara evaluasi laboratorium dan kebutuhan implementasi. Melalui sintesis tersebut, penelitian ini menargetkan beberapa kontribusi praktis: pertama, menyajikan ringkasan komparatif yang memudahkan pembaca menilai keterkaitan antara akurasi dan robustitas pada berbagai kondisi; kedua, merumuskan model konseptual yang menjelaskan relasi **akurasi–serangan–pertahanan** sehingga pembahasan tidak terputus-putus; dan ketiga, memberikan rekomendasi evaluasi keamanan yang lebih terstandar, misalnya perlunya pelaporan skenario ancaman (white-box/black-box), karakter data uji, serta metrik yang relevan agar hasil antar studi lebih mudah dibandingkan. Dengan kata lain, kontribusi SLR ini tidak hanya bersifat akademik untuk memetakan literatur, tetapi juga operasional untuk mendukung pengambil keputusan dalam memilih model dan menyusun prosedur uji yang sesuai kebutuhan keamanan.

Agar keseluruhan naskah tersaji runtut dan tidak "melompat-lompat" (sebagaimana kritik reviewer tentang alur yang tidak konsisten), penelitian ini disusun dengan alur yang bergerak dari masalah menuju jawaban secara sistematis. Setelah pendahuluan menetapkan konteks, ancaman, dan kesenjangan penelitian, bagian metode menjelaskan tahapan PRISMA, kriteria seleksi, serta pendekatan sintesis tematik sehingga pembaca memahami dari mana kesimpulan berasal. Bagian hasil dan pembahasan kemudian menampilkan ringkasan temuan terstruktur per tema—bukan sekadar daftar studi—dengan penekanan pada keterkaitan antara performa, serangan, dan pertahanan, termasuk kondisi real-world yang sering luput dalam evaluasi. Terakhir, kesimpulan merangkum jawaban atas tujuan penelitian sekaligus menegaskan implikasi praktis, keterbatasan SLR (misalnya heterogenitas dataset/protokol, variasi metrik, dan potensi bias publikasi), serta peluang riset lanjutan. Dengan struktur seperti ini, pembaca tidak akan menemukan urutan yang "hasil → masalah → contoh → hasil lagi", karena urutan pembahasan dijaga tetap linear: mulai dari latar dan gap, menuju metode seleksi, lalu sintesis temuan, dan ditutup dengan implikasi serta keterbatasan secara proporsional. Alur yang sama juga sebaiknya tercermin di abstrak: dimulai dari latar masalah dan tujuan, diikuti metode SLR (termasuk **jumlah studi 25**), lalu temuan utama (ringkas per tema), dan ditutup dengan implikasi serta keterbatasan.

Penelitian ini bertujuan untuk mengkaji secara komprehensif dampak algoritma *deep learning* terhadap akurasi dan keamanan sistem pengenalan wajah melalui analisis sistematis terhadap literatur terkini. Secara khusus, penelitian ini akan: (1) menganalisis berbagai arsitektur *deep learning* dan performa mereka dalam pengenalan wajah, (2) mengidentifikasi dan mengklasifikasikan berbagai bentuk serangan terhadap sistem pengenalan wajah, (3) mengevaluasi efektivitas mekanisme pertahanan yang telah dikembangkan, dan (4) merumuskan rekomendasi untuk pengembangan sistem pengenalan wajah yang lebih robust dan aman (Fakrogha et al., 2025).

2 Metode

Penelitian ini menggunakan metode studi literatur sistematis (*Systematic Literature Review*) untuk mengkaji dampak algoritma *deep learning* (Sya'roni et al., 2025)terhadap akurasi dan keamanan sistem pengenalan wajah. Pendekatan ini dipilih karena memungkinkan analisis komprehensif terhadap berbagai penelitian yang telah dipublikasikan, sehingga dapat memberikan gambaran menyeluruh tentang perkembangan teknologi, tantangan, dan solusi yang ada dalam domain pengenalan wajah berbasis *deep learning*.

Sumber data primer penelitian ini adalah artikel jurnal ilmiah yang dipublikasikan pada rentang tahun 2021 hingga 2025, dengan fokus pada publikasi yang membahas algoritma *deep learning* untuk pengenalan wajah, keamanan sistem biometrik, dan serangan terhadap sistem

pengenalan wajah. *Database* yang digunakan mencakup IEEE Xplore, ACM Digital Library, ScienceDirect, Springer, dan MDPI. Kata kunci pencarian yang digunakan meliputi "*deep learning*", "*face recognition*", "*adversarial attack*", "*face morphing*", "*deepfake detection*", "CNN", "VGG", "ResNet", dan kombinasinya.

Kriteria inklusi untuk seleksi artikel meliputi: (1) publikasi dalam jurnal atau prosiding konferensi yang *peer-reviewed*, (2) membahas penerapan *deep learning* dalam pengenalan wajah atau keamanan sistem biometrik, (3) menyajikan data eksperimental atau analisis yang terukur, dan (4) ditulis dalam bahasa Inggris atau bahasa Indonesia. Kriteria eksklusi mencakup artikel yang hanya bersifat *review* tanpa analisis baru, publikasi yang tidak dapat diakses *full-text*, dan artikel yang tidak relevan dengan fokus penelitian.

Analisis data dilakukan melalui pendekatan tematik dengan mengkategorikan temuan berdasarkan beberapa dimensi utama: (1) arsitektur dan performa algoritma *deep learning*, (2) jenis dan karakteristik serangan terhadap sistem pengenalan wajah, (3) mekanisme pertahanan dan mitigasi, dan (4) tantangan implementasi di lingkungan *real-world*. Setiap artikel yang terpilih dianalisis untuk mengekstraksi informasi mengenai metodologi yang digunakan, *dataset* eksperimen, metrik evaluasi, hasil akurasi, dan kesimpulan utama.

3 Hasil dan Diskusi

Hasil analisis literatur menunjukkan bahwa algoritma *deep learning* telah memberikan dampak transformatif terhadap akurasi sistem pengenalan wajah. Arsitektur CNN modern seperti VGG-16, ResNet-50, dan *FaceNet* secara konsisten mengungguli metode konvensional berbasis *handcrafted features* (Deng et al., 2024; Singh et al., 2023). Penelitian oleh Kristanto *et al.* (Kristanto et al., 2023) mendemonstrasikan bahwa metode PCA dapat mencapai akurasi 63,33% pada *dataset* LFW dan 46,66% pada *dataset* Face94, dengan performa terbaik pada citra yang mengalami *blur* (80-100%) namun mengalami penurunan drastis pada citra dengan permasalahan kecerahan.

Tabel 1. Ringkasan Studi dan Temuan Utama (Hasil SLR)

Studi	Fokus	Metode/Model	Temuan Utama
Kristanto et al. (2023)	Baseline & gangguan citra	PCA (konvensional)	Akurasi 63,33% (LFW) dan 46,66% (Face94); kecerahan ekstrem menurunkan performa drastis.
Iqbal et al. (2021)	Surveillance real-time	Faster R-CNN	Deteksi ancaman keamanan AP 79% pada sistem quadcopter surveillance.
Al-Dmour et al. (2023)	Real-world (masker)	CNN adaptif	Adaptasi fitur area mata/dahi untuk pengenalan wajah bermasker.
Hung Hwang et al. (2023)	Serangan adversarial	Adversarial patch (GAN)	Dodging success rate 39,94% (DB 10) → 81,77% (DB 22).
Long et al. (2022)	Morphing & edge	Lightweight + patch-level	Deteksi morphing hingga 95% dan inferensi 3–4× lebih cepat.
Ramachandran et al. (2021)	Deepfake	Deteksi berbasis CNN	Akurasi hingga 98% pada dataset tertentu; turun signifikan pada lintas-domain.
Zhang et al. (2022)	Defense	Adversarial training	Menurunkan keberhasilan serangan sekitar 50–70%

			(tergantung skenario).
Chi et al. (2023)	Mobile/ edge	Model lightweight	Menekankan efisiensi komputasi untuk deployment perangkat terbatas.
Hangaragi et al. (2023); Singh et al. (2023)	Multimodal	VGGFace/ResNet	Multimodal meningkatkan akurasi 7–12% dibanding unimodal.

Implementasi sistem pengenalan wajah berbasis *deep learning* pada aplikasi *real-world* menghadapi berbagai tantangan teknis (Prabha et al., 2025). Dalam konteks surveillance keamanan, Iqbal *et al.* (Iqbal et al., 2021) mengembangkan sistem surveillance real-time berbasis quadcopter yang mengintegrasikan FasterRCNN untuk deteksi multi-objek termasuk wajah, senjata, penjaga, dan penyusup, mencapai average precision 79% untuk deteksi ancaman keamanan. Al-Dmour *et al.* (Al-dmour et al., 2023) mengidentifikasi bahwa penggunaan masker wajah selama pandemi COVID-19 mengharuskan adaptasi arsitektur CNN untuk fokus pada ekstraksi fitur dari area mata dan dahi. Penelitian mereka mengembangkan sistem yang dapat mendeteksi dan mengenali wajah bermasker dengan memodifikasi layer konvolusi awal untuk memperbesar *receptive field* pada area mata.

Analisis terhadap kerentanan sistem pengenalan wajah mengungkapkan berbagai jenis serangan yang mengancam keamanan sistem (Xing et al., 2025). Hwang *et al.* (Hung Hwang et al., 2023) melakukan studi komprehensif tentang serangan *adversarial patch* menggunakan GAN, menemukan bahwa serangan *dodging* (menghindari deteksi) memiliki tingkat keberhasilan rata-rata 39,94% pada *database* 10 orang dan meningkat menjadi 81,77% pada *database* 22 orang.

Serangan *face morphing* merepresentasikan ancaman serius terhadap sistem verifikasi identitas, terutama pada aplikasi paspor elektronik dan kontrol perbatasan (Long et al., 2022). Long *et al.* (Long et al., 2022) mengembangkan metode deteksi berbasis fitur tingkat *patch* dan jaringan *lightweight* yang dirancang khusus untuk lingkungan *mobile* dan kondisi dengan sumber daya terbatas. Pendekatan ini menggunakan modul perhatian kanal yang efisien (*Efficient Channel Attention*) dan mekanisme *Squeeze-and-Excitation* untuk mengurangi jumlah parameter sambil mempertahankan akurasi deteksi yang tinggi. Hasil eksperimen menunjukkan bahwa metode ini dapat mencapai akurasi deteksi hingga 95% dengan jumlah parameter yang jauh lebih sedikit dibandingkan arsitektur CNN konvensional, menjadikannya ideal untuk implementasi pada perangkat *edge computing* dan sistem keamanan perbatasan yang memerlukan respons cepat.

Teknologi *deepfake* yang memanfaatkan arsitektur GAN menghadirkan tantangan deteksi yang semakin kompleks (Ramachandran et al., 2021). Studi literatur sistematis menunjukkan bahwa metode deteksi *deepfake* berbasis CNN dapat mencapai akurasi hingga 98% pada *dataset* tertentu, namun mengalami penurunan performa signifikan ketika diuji pada *dataset* lintas-domain atau video yang dihasilkan oleh generator berbeda.

Mekanisme pertahanan terhadap serangan *adversarial* telah menjadi area penelitian yang intensif. *Adversarial training*, yang melibatkan pelatihan model dengan contoh *adversarial* yang dihasilkan secara sintetis, terbukti meningkatkan robustness sistem terhadap serangan *white-box* dan *black-box* (Hung Hwang et al., 2023; Zhang et al., 2022). Penelitian menunjukkan bahwa model yang dilatih dengan *adversarial training* dapat mengurangi tingkat keberhasilan serangan hingga 50-70%.

Implementasi sistem pengenalan wajah pada perangkat *mobile* dan *edge computing* menghadapi tantangan tambahan terkait keterbatasan komputasi dan memori (Chi et al., 2023; Long et al., 2022). Arsitektur *lightweight* seperti *MobileNet*, *ShuffleNet*, dan *EfficientNet* telah dikembangkan untuk menjawab tantangan ini. Long *et al.* (Long et al., 2022) mendemonstrasikan bahwa pendekatan berbasis fitur tingkat *patch* dapat secara signifikan mengurangi kompleksitas komputasi tanpa mengorbankan akurasi deteksi, dengan waktu inferensi yang 3-4 kali lebih cepat dibandingkan metode berbasis VGG atau ResNet standar.

Pendekatan multimodal yang menggabungkan pengenalan wajah dengan modalitas biometrik lainnya menunjukkan peningkatan signifikan dalam akurasi dan keamanan sistem (Hangaragi et al., 2023; Singh et al., 2023). Penelitian

tentang pengenalan multimodal yang mengintegrasikan fitur wajah dan tubuh menggunakan VGG Face-16 dan ResNet-50 mendemonstrasikan peningkatan akurasi hingga 7-12% dibandingkan sistem unimodal.

Berdasarkan ringkasan pada **Tabel 1**, terlihat bahwa peningkatan performa pengenalan wajah berbasis *deep learning* berjalan berdampingan dengan munculnya persoalan keamanan dan keterbatasan implementasi yang belum terselesaikan secara menyeluruh. Studi yang berfokus pada performa menunjukkan bahwa model modern cenderung unggul pada kondisi uji tertentu, tetapi temuan mengenai gangguan kualitas citra menegaskan bahwa performa dapat menurun drastis pada kondisi lapangan. Indikasi ini memperlihatkan *gap* yang nyata antara evaluasi berbasis benchmark atau kondisi terkontrol dengan kebutuhan real-world yang menghadirkan variasi pencahayaan, pose, oklusi, dan kualitas citra yang tidak stabil. Dengan kata lain, akurasi tinggi pada kondisi ideal belum cukup untuk menjamin reliabilitas sistem ketika dioperasikan pada skenario pengawasan, kontrol akses, atau verifikasi identitas yang dinamis.

Selain tantangan lingkungan, tabel juga menegaskan bahwa dimensi keamanan menjadi isu yang semakin dominan. Temuan terkait serangan adversarial menunjukkan bahwa keberhasilan serangan dapat meningkat ketika skala identitas dalam basis data bertambah, yang mengindikasikan bahwa sistem yang "aman" pada skenario kecil dapat menjadi jauh lebih rentan pada skenario operasional yang lebih besar. Hal ini menjadi masalah penting karena banyak penerapan pengenalan wajah di dunia nyata justru melibatkan populasi besar, data yang heterogen, dan kebutuhan respons real-time. Di sisi lain, studi morphing memperlihatkan bahwa manipulasi citra dapat mengancam integritas sistem verifikasi, khususnya pada konteks yang berdampak tinggi seperti paspor elektronik dan kontrol perbatasan. Sementara itu, ringkasan tentang deepfake menyoroti persoalan generalisasi: metode deteksi dapat sangat akurat pada dataset tertentu, tetapi performanya menurun ketika diuji lintas-domain atau pada generator yang berbeda. Pola ini menandakan adanya *gap* metodologis pada literatur, yaitu ketergantungan kuat pada dataset spesifik dan kurangnya pengujian yang merepresentasikan variasi ancaman di luar data latih.

Tabel 1 juga memperlihatkan bahwa solusi pertahanan memang tersedia, tetapi masih menyisakan masalah trade-off. Adversarial training dan mekanisme deteksi dapat menekan keberhasilan serangan, namun seringkali menuntut biaya komputasi lebih tinggi, kebutuhan data tambahan, atau kompleksitas implementasi yang meningkat. Kondisi ini menjadi semakin problematik pada skenario edge/mobile, karena keterbatasan komputasi dan memori menuntut penggunaan arsitektur ringan. Walaupun model lightweight memperbaiki kelayakan deployment, ringkasan studi menunjukkan bahwa isu robustness pada model-model efisien ini belum selalu dievaluasi secara memadai dalam konteks serangan dan gangguan real-world. Dengan demikian, terdapat *gap* praktis: kebutuhan efisiensi operasional belum sepenuhnya dipertemukan dengan kebutuhan ketahanan terhadap ancaman.

Lebih jauh, temuan multimodal yang meningkatkan akurasi membuka peluang, namun juga menimbulkan masalah baru terkait kompleksitas integrasi, konsistensi kualitas antar modalitas, serta kebutuhan evaluasi keamanan yang lebih luas. Secara keseluruhan, Tabel 1 menegaskan bahwa riset pengenalan wajah berbasis *deep learning* masih memerlukan sintesis integratif yang menghubungkan performa, ancaman, pertahanan, dan kendala implementasi dalam satu kerangka evaluasi yang konsisten. Gap utama yang muncul adalah kurangnya standardisasi protokol uji (terutama lintas-domain dan lintas-skenario), serta belum kuatnya pemetaan trade-off antara akurasi, robustness, dan biaya komputasi untuk kebutuhan operasional yang berbeda.

4 Kesimpulan

Penelitian ini meninjau literatur 2021–2025 untuk menilai dampak *deep learning* terhadap akurasi dan keamanan pengenalan wajah. Hasil menunjukkan bahwa arsitektur CNN (mis. VGG, ResNet, FaceNet) mampu mencapai akurasi hingga ±99% pada kondisi ideal, namun performa dan keandalan menurun pada skenario *real-world* akibat variasi pencahayaan, pose, oklusi, dan kualitas citra rendah. Dari sisi keamanan, strategi pertahanan seperti *adversarial training*, deteksi berbasis *ensemble*, dan pendekatan multimodal terbukti meningkatkan robustness serta menurunkan keberhasilan serangan sekitar 50–70% pada beberapa studi, meski tetap berpotensi menimbulkan trade-off komputasi dan kompleksitas. Ke depan, pengembangan diarahkan pada arsitektur yang lebih robust secara inheren, deteksi anomali real-time terhadap manipulasi/adversarial, integrasi multimodal, penyusunan dataset yang lebih representatif

kondisi lapangan, serta pendekatan *privacy-preserving face recognition* untuk menyeimbangkan akurasi, keamanan, dan privasi.

5 Referensi

Al-dmour, H., Tareef, A., Alkalbani, A. M., Hammouri, A., & Alrahmani, B. (2023). Masked Face Detection and Recognition System Based on Deep Learning Algorithms. *Journal of Advances in Information Technology*, *14*(2). https://doi.org/10.12720/jait.14.2.224-232

Chi, J., Kim On, C., Zhang, H., & See Chai, S. (2023). Interactive Mobile Technologies. *International Journal of Interactive Mobile*, *17*(23), 4–19. https://doi.org/https://doi.org/10.3991/ijim.v17i23.40867

Deng, N., Xu, Z., Li, X., Gao, C., & Wang, X. (2024). applied sciences Deep Learning and Face Recognition : Face Recognition Approach Based on the DS-CDCN Algorithm. *Applied Sciences*. https://doi.org/https://doi.org/10.3390/app14135739

Fakrogha, B. E., Olapegba, P. O., & Uye, E. E. (2025). Anxiety and Depression as Predictors of Quality of Life among University Undergraduates. *TRILOGI: Jurnal Ilmu Teknologi, Kesehatan, Dan Humaniora*, *6*(1), 83–90. https://doi.org/https://doi.org/10.33650/trilogi.v6i1.10799

Hangaragi, S., Singh, T., & Neelima, N. (2023). ScienceDirect ScienceDirect ScienceDirect Face Detection and Recognition Using Face and Mesh and Deep Neural Machine Learning and Data Engineering Face Detection and Recognition Using Face Mesh and * Deep Neural Network Network. *Procedia Computer Science*, *218*, 741–749. https://doi.org/10.1016/j.procs.2023.01.054

Hung Hwang, R., You Lin, J., Ying Hsieh, S., Yu Lin, H., & Liang Lin, C. (2023). Adversarial Patch Attacks on Deep-Learning-Based Face Recognition Systems Using Generative Adversarial Networks. *Sensors*. https://doi.org/https://doi.org/10.3390/s23020853

Iqbal, M. J., Iqbal, M. M., Ahmad, I., Alassafi, M. O., Alfakeeh, A. S., & Alhomoud, A. (2021). Real-Time Surveillance Using Deep Learning. *Security and Communication Networks*, *2021*(1). https://doi.org/10.1155/2021/6184756

Kristanto, V. N., Riadi, I., & Prayudi, Y. (2023). Analisa Deteksi dan Pengenalan Wajah pada Citra dengan Permasalahan Visual. *JISKA (Jurnal Informatika Sunan Kalijaga)*, *8*(1), 78–89. https://doi.org/https://doi.org/10.14421/jiska.2023.8.1.78-89

Long, M., Zhao, X., Zhang, L., & Peng, F. (2022). Detection of Face Morphing Attacks Based on Patch-Level Features and Lightweight Networks. *Security and Communication Networks*, *2022*(1), 12. https://doi.org/10.1155/2022/7460330

Maulidiansyah, M., & Yaqin, M. A. (2023). Deteksi Tumpukan Sampah dengan Metode You Only Look Once (YOLO). *TRILOGI: Jurnal Ilmu Teknologi, Kesehatan, Dan Humaniora*, *4*(2), 76–79. https://doi.org/https://doi.org/10.33650/trilogi.v4i2.6185

Prabha, B., Poonkodi, M., & Joseph, L. (2025). An optimal hybrid deep learning-aided facial emotion detection and classification scheme to identify criminal activities. *Automatika*, *66*(4). https://doi.org/10.1080/00051144.2025.2560147

Ramachandran, S., Nadimpalli, A. V., & Rattani, A. (2021). An Experimental Evaluation on Deepfake Detection using Deep Face Recognition. *International Carnahan Conference on Security Technology (ICCST)*, 1–6. https://doi.org/10.1109/ICCST49569.2021.9717407

Singh, A., Bhatt, S., Nayak, V., & Shah, M. (2023). Automation of surveillance systems using deep learning and facial recognition. *International Journal of System Assurance Engineering and Management*, *14*(1), 236–245. https://doi.org/10.1007/s13198-022-01844-6

Sya'roni, W., Auliya'Abdillah, Y., Arifin, S., & Hibatullah, R. (2025). Klasifikasi Penyakit Daun Sawi Menggunakan VGG19 Berbasis Citra Digital. *TRILOGI: Jurnal Ilmu Teknologi, Kesehatan, Dan Humaniora*, *6*(3), 20–30. https://doi.org/https://doi.org/10.33650/trilogi.v6i3.12106

Wang, M., & Deng, W. (2021). Deep Face Recognition : A Survey. *Neurocomputing*, *429*, 215–244. https://doi.org/https://doi.org/10.48550/arXiv.1804.06655

Xing, H., Tan, S. Y., Qamar, F., & Jiao, Y. (2025). Face Anti-Spoofing Based on Deep Learning : A Comprehensive Survey. *Applied Sciences*. https://doi.org/https://doi.org/10.3390/app

15126891

Yan, L., Zhang, Y., & Zhang, Y. (2023). A fast face recognition system based on annealing algorithm to optimize operator parameters. *The Imaging Science Journal ISSN:*, *71*(4), 323–330. https://doi.org/10.1080/13682199.2023.2182261

Zhang, Y., Wang, Z., Zhang, X., Cui, Z., Zhang, B., Cui, J., & Janneh, L. L. (2022). in face recognition. *CAAI Transactions on Intelligence Technology*, *8*(4), 1391–1402. https://doi.org/10.1049/cit2.12115