

Sistem Gaze Tracking Berbasis YOLOv8 dan Random Forest Melalui Ekstraksi Fitur Geometris Landmark Mata

Dadya Vio Hendraksa¹, Denny Budiansyach¹, Aditya Dwi Putro Wicaksono^{1*}

¹ Program Studi Teknik Informatika, Universitas Telkom, Indonesia

* Corresponding author: adityaw@telkomuniversity.ac.id

Sitasi: D. V. Hendraksa, D. Budiansyach and A. D. P. Wicaksono, "Sistem Gaze Tracking Berbasis YOLOv8 dan Random Forest Melalui Ekstraksi Fitur Geometris Landmark Mata," *Jurnal Teknologi Informasi dan Multimedia*, vol. 8, no. 4, pp. 790–804, 2026.

Diterima : 16-06-2026
Direvisi : 30-07-2026
Disetujui : 06-08-2026
Diterbitkan : 02-09-2026

Copyright: © 2026 oleh para penulis. Karya ini dilisensikan di bawah Creative Commons Attribution-Share Alike 4.0 International License.

Abstrak. *Gaze tracking* merupakan instrumen penting dalam realisasi interaksi adaptif *Human-Centric AI* di era Industri 5.0. Namun, sebagian besar metode estimasi tatapan berbasis penampilan (*appearance-based*) saat ini bergantung pada arsitektur regresi *deep learning end-to-end* berskala besar yang memiliki kompleksitas jaringan dan beban komputasi tinggi. Penelitian ini mengusulkan sebuah kerangka kerja hibrida yang mengintegrasikan model *computer vision* YOLOv8-Pose dengan algoritma *machine learning Random Forest Classifier* untuk klasifikasi spasial *gaze tracking*. Matriks piksel citra mentah secara efisien direduksi dengan melokalisasi 55 titik koordinat kritis (*keypoints*) area okular untuk diekstraksi menjadi 113 atribut fitur geometris tabular, yang mencakup parameter *Eye Aspect Ratio* (EAR) dan pergeseran posisi relatif iris. Model ansambel ini kemudian dilatih menggunakan dataset yang dibangun secara sintesis melalui simulator UnityEyes. Hasil eksperimen menunjukkan bahwa sistem hibrida yang diusulkan berhasil mencapai akurasi global sebesar 82,75%. Kemampuan diskriminasi ruang fitur tertinggi tercatat pada kelas spasial *TopRight* (F1-score 0,91) dan *Middle-Right* (F1-score 0,85). Pendekatan *feature engineering* terstruktur ini terbukti secara empiris mampu memitigasi galat klasifikasi yang bersifat anisotropik tanpa memerlukan sumber daya komputasi tingkat tinggi. Untuk pengembangan lebih lanjut, penelitian masa depan diarahkan pada eksplorasi hyperparameter tuning agresif, seleksi fitur berbasis meta-heuristik *Particle Swarm Optimization* (PSO), serta pengujian ketahanan model (*robustness limit*) menggunakan dataset dunia nyata di ranah publik yang tidak terkendali (*unconstrained open domains*).

Kata kunci: gaze tracking; YOLOv8; random forest; landmark mata; fitur geometris.

Abstract. Eye gaze tracking serves as a pivotal instrument in realizing adaptive Human-Centric AI interactions within the Industry 5.0 era. However, current appearance-based gaze estimation methods primarily rely on large-scale end-to-end deep learning regression architectures characterized by high network complexity and computational overhead. This research proposes a hybrid framework integrating the YOLOv8-Pose computer vision model with the Random Forest Classifier machine learning algorithm for the spatial classification of gaze directions. Raw image pixel matrices are efficiently reduced by localizing 55 ocular keypoints, which are subsequently extracted into 113 tabular geometric feature attributes, encompassing Eye Aspect Ratio (EAR) parameters and the relative spatial shifts of the iris. The ensemble model was trained utilizing a synthetic dataset generated via the UnityEyes simulator. Experimental results demonstrate that the proposed hybrid system achieved a global accuracy of 82.75%. The highest feature space discrimination capabilities were recorded in the TopRight (F1-score 0.91) and MiddleRight (F1-score 0.85) spatial classes. This structured feature engineering approach is empirically proven to effectively mitigate anisotropic classification errors without necessitating high-end computational resources. For further development, future research is directed towards exploring aggressive hyperparameter tuning, meta-heuristic feature selection using Particle Swarm Optimization (PSO), and evaluating the model's robustness limit using real-world datasets in unconstrained open domains.

Keywords: gaze tracking; YOLOv8; random forest; eye landmarks; geometric features.

1. Pendahuluan

Perkembangan teknologi interaksi manusia dan komputer (*Human-Computer Interaction*) telah mengalami transformasi menuju antarmuka pengguna cerdas (*Intelligent User Interfaces*) yang lebih alami dan responsif. Di era Industri 5.0, integrasi ilmu data (*data science*) dan otomatisasi melahirkan konsep *Human-Centric AI*, yang menempatkan pemahaman terhadap perilaku dan atensi manusia sebagai pusat sistem adaptif[1]. Melalui konsep tersebut, evaluasi antarmuka cerdas dirancang untuk meningkatkan keandalan interaksi antara manusia dan mesin cerdas[2]. Salah satu teknologi yang berperan penting dalam mendeteksi perhatian visual manusia adalah *gaze tracking*, yang telah diterapkan pada sistem otentikasi biometrik[3], pemantauan aktivitas kognitif dalam interaksi robot cerdas[4], serta sistem keselamatan transportasi untuk mendeteksi kelalaian maupun kelelahan pengemudi[5].

Secara umum, sistem estimasi *gaze tracking* diklasifikasikan menjadi dua pendekatan utama, yaitu *model-based* dan *appearance-based*. Pendekatan *model-based* memanfaatkan rekonstruksi geometri bola mata tiga dimensi menggunakan sensor khusus, seperti kamera *Time-of-Flight* (ToF) dan cahaya inframerah, sehingga mampu memberikan akurasi tinggi, tetapi kurang fleksibel untuk diterapkan pada lingkungan publik[6]. Sebaliknya, pendekatan *appearance-based* memanfaatkan citra wajah yang diproses menggunakan jaringan saraf tiruan (*Deep Neural Networks*)[7]. Berbagai penelitian menunjukkan bahwa pendekatan ini lebih fleksibel untuk diimplementasikan pada perangkat komersial tanpa memerlukan sensor tambahan[8]. Jaringan saraf tersebut mempelajari representasi visual mata secara dinamis dari berbagai kondisi citra[9]. Meskipun demikian, implementasinya masih dituntut mampu beradaptasi dengan keterbatasan sumber daya komputasi pada lingkungan *edge computing*[10].

Meskipun menawarkan fleksibilitas yang tinggi, metode *appearance-based gaze estimation* berbasis *deep learning end-to-end* masih menghadapi tantangan berupa kompleksitas model yang berdampak pada tingginya kebutuhan memori dan komputasi [8], [10]. Analisis sebelumnya juga menunjukkan adanya *trade-off* antara peningkatan akurasi dan kebutuhan sumber daya komputasi pada model berskala besar [11]. Selain itu, performa model *end-to-end* cenderung menurun ketika menghadapi perubahan pencahayaan, oklusi wajah, maupun variasi orientasi kepala[12]. Li dkk. [12] melaporkan bahwa model *Appearance-Based Gaze Estimation* (AGE) mengalami *anisotropic error*, yaitu galat prediksi pada sumbu vertikal (*pitch*) yang lebih besar dibandingkan sumbu horizontal (*yaw*) akibat *feature collapse* dan bias penyalarsan label geometris. Di sisi lain, Park dkk. [13] menunjukkan bahwa *MediaPipe Facemesh* mengalami penurunan akurasi pada rotasi wajah hingga 45° , sedangkan Dlib rentan mengalami kegagalan regresi pada kondisi pencahayaan yang tidak ideal [13]. Keterbatasan tersebut menunjukkan perlunya pendekatan yang lebih ringan, stabil, dan adaptif terhadap kondisi dunia nyata.

Sebagai upaya mengatasi keterbatasan tersebut, pendekatan *hybrid machine learning* mulai dikembangkan dengan memisahkan proses ekstraksi fitur visual dan klasifikasi. Dalam tugas estimasi pose dan pelokalan landmark, keluarga model You Only Look Once (YOLO) dikenal sebagai salah satu metode *state-of-the-art* yang mampu melakukan deteksi objek, estimasi pose, dan orientasi kepala secara simultan[14]. Dibandingkan generasi sebelumnya, YOLOv8 menghadirkan peningkatan melalui mekanisme *anchor-free detection* dan optimasi *pose loss*, sehingga mampu meningkatkan presisi regresi koordinat spasial dua dimensi pada objek yang kompleks maupun teroklusi[15]. Koordinat landmark yang diperoleh kemudian dapat ditransformasikan menjadi fitur geometris, seperti jarak *Euclidean* dan *Eye Aspect Ratio* (EAR), menggunakan transformasi matematis[16]. Fitur-fitur tersebut selanjutnya dapat diklasifikasikan secara efisien menggunakan Random Forest[17], yang memiliki interpretabilitas tinggi[18], mampu menangani fitur geometris berdimensi tinggi dengan risiko *overfitting* yang rendah melalui

mekanisme *majority voting* [19], serta menghasilkan representasi numerik yang lebih ringkas dan efisien untuk diproses [20].

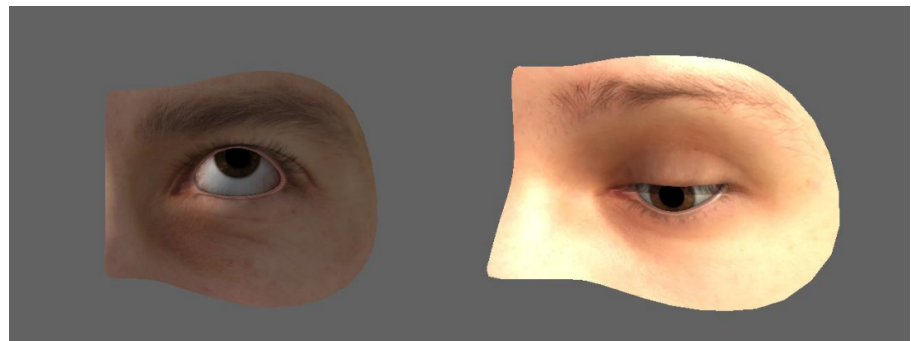
Meskipun berbagai penelitian telah menunjukkan keunggulan YOLOv8 dalam estimasi pose dan Random Forest dalam klasifikasi data tabular, integrasi kedua metode tersebut untuk klasifikasi arah tatapan berbasis representasi geometris lokal masih belum banyak dieksplorasi. Berdasarkan kesenjangan tersebut, penelitian ini mengembangkan sistem *gaze tracking* hibrida yang mengintegrasikan model regresi landmark berbasis YOLOv8 dengan algoritma Random Forest. Integrasi kedua algoritma dipilih untuk mengatasi keterbatasan model *deep learning end-to-end* yang membutuhkan sumber daya komputasi tinggi [8], [10] serta keterbatasan Dlib dan MediaPipe yang mengalami penurunan kinerja pada rotasi wajah ekstrem dan kondisi pencahayaan yang berubah-ubah [13]. YOLOv8 digunakan untuk mengekstraksi 55 titik landmark kontur kelopak mata dan pupil dari *Eye Gaze Detection Dataset* [21]. Koordinat landmark hasil regresi kemudian dikonversi menjadi parameter geometris berupa *Eye Aspect Ratio* (EAR) dan pergeseran posisi relatif iris sebagai representasi fitur numerik. Selanjutnya, fitur tersebut dimasukkan ke dalam algoritma *Random Forest Classifier* untuk mengklasifikasikan arah tatapan ke dalam delapan kelas spasial (*gaze direction*). Kinerja sistem dievaluasi menggunakan *accuracy matrix*, *classification report*, dan *confusion matrix*. Tujuan utama penelitian ini adalah membuktikan bahwa representasi geometris lokal hasil ekstraksi *landmark* mampu menghasilkan klasifikasi *gaze tracking* yang stabil dan berakurasi tinggi, sekaligus mengurangi ketergantungan terhadap arsitektur *deep learning end-to-end* yang kompleks dan membutuhkan sumber daya komputasi besar.

2. Bahan dan Metode

Sistem *gaze tracking hybrid* yang dikembangkan dalam penelitian ini mengintegrasikan pemrosesan visi komputer tingkat lanjut untuk ekstraksi *landmark* lokal dengan pemodelan pembelajaran mesin berbasis statistik. Kerangka kerja dirancang menggunakan skema hibrida untuk mereduksi kompleksitas masukan citra visual menjadi representasi data tabular numerik yang ringkas dan andal.

2.1. Dataset dan Sumber Bahan Penelitian

Penelitian ini menggunakan data sekunder yang bersumber dari public repository, yaitu *Eye Gaze Detection Dataset* yang diunggah oleh Hazem Fahmy [21]. Karakteristik instansiasi data sintetis di dalam basis data tersebut dibangkitkan melalui perangkat simulator UnityEyes, sehingga menyajikan variasi geometri mata, pantulan spekular kornea, dan pergeseran posisi pupil beresolusi tinggi yang menyerupai kondisi anatomi mata manusia nyata. Karena dataset ini murni merupakan hasil simulasi berbasis grafik komputer tiga dimensi (3D) tanpa melibatkan intervensi langsung terhadap partisipan manusia atau hewan, penelitian ini secara hukum tidak memerlukan kode persetujuan etik khusus.



Gambar 1. Contoh Citra pada Eye Gaze Detection Dataset

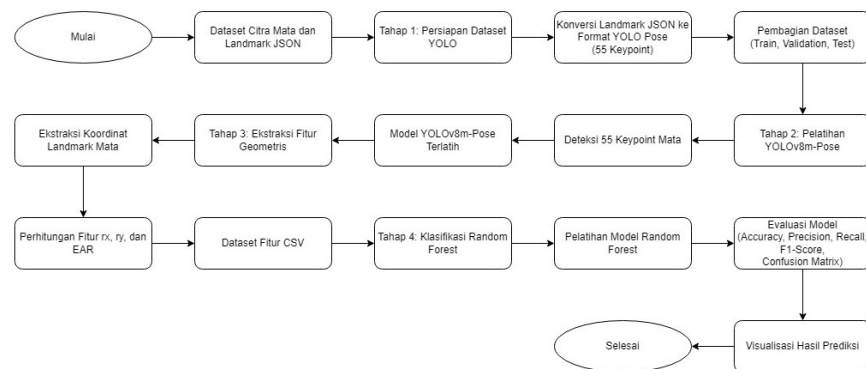
Dataset visual ini memuat citra area mata terisolasi yang telah dianotasi

secara terstruktur. Dataset yang digunakan terdiri atas 4.000 sampel citra dengan resolusi 640×480 piksel. Dataset dibagi menjadi data pelatihan sebanyak 2.800 sampel (70%), data validasi sebanyak 800 sampel (20%), dan data pengujian sebanyak 400 sampel (10%). Untuk kebutuhan pemodelan, dataset dibagi menjadi dua komponen utama:

1. Dataset Visi Komputer: Digunakan untuk melatih model regresi landmark mata berbasis YOLOv8. Citra masukan memiliki sebaran koordinat titik kritis (*keypoints*) yang menandai kontur kelopak mata dan batas iris.
2. Dataset Tabular Numerik: Hasil konversi geometris dari koordinat *landmark* untuk melatih dan menguji model klasifikasi arah tatapan. Basis data tabular ini mendistribusikan sampel ke dalam 8 kelas spasial arah tatapan (*gaze direction*), yaitu: TopCenter, TopLeft, TopRight, MiddleLeft, MiddleRight, BottomLeft, BottomRight, BottomCenter.

2.2. Deskripsi Diagram Alir Pemrosesan Sistem

Sesuai dengan metodologi perancangan arsitektur cerdas, seluruh rangkaian proses pemrosesan data di dalam sistem harus berjalan secara sekuensial. Alur kerja operasional dari sistem gaze tracking hybrid yang diusulkan dijabarkan secara rinci melalui tahapan-tahapan di bawah ini:



Gambar 2. Flowchart metode penelitian sistem gaze tracking berbasis YOLOv8 dan Random Forest melalui ekstraksi fitur geometris landmark mata

1. Mulai (Citra Area Mata Masukan): Sistem menerima masukan berupa berkas citra digital area mata (baik mata kiri maupun kanan) dalam format RGB (Red, Green, Blue) yang diekstrak dari repositori dataset [21].
2. Tahap 1 (Prapemrosesan Citra Masukan): Citra mentah dikonversi ke dalam ruang matriks numerik melalui pustaka visi komputer. Proses ini melibatkan operasi *resizing* untuk menyelaraskan dimensi resolusi gambar agar sesuai dengan konfigurasi lapisan input model YOLOv8, diikuti dengan normalisasi nilai intensitas piksel menuju rentang matematis $[0, 1]$ untuk mempercepat konvergensi gradien.
3. Tahap 2 (Inferensi & Regresi Spasial YOLOv8): Matriks citra diumpankan ke dalam arsitektur model YOLOv8 yang telah dilatih secara khusus untuk ekstraksi *keypoint*. Lapisan output head model melakukan regresi spasial dua dimensi (2D) untuk memprediksi lokasi presisi dari 55 titik *landmark* kustom yang mengitari kontur kelopak mata dan lingkaran iris[14].
4. Tahap 3 (Ekstraksi Fitur Geometris): Nilai koordinat spasial (x, y) hasil proyeksi YOLOv8 diisolasi. Logika matematika diterapkan untuk menghitung jarak Euclidean antar-titik kelopak mata guna menghasilkan nilai *Eye Aspect Ratio* (EAR) serta melokalisasi selisih jarak absolut antara pusat massa iris dengan sudut mata interior dan eksterior[16].

5. Tahap 4 (Rekonstruksi Matriks Tabular): Seluruh parameter fitur skalar hasil kalkulasi geometris disusun menjadi satu baris vektor fitur tunggal. Kumpulan vektor dari ribuan citra diakumulasikan menjadi sebuah matriks data tabular berdimensi tinggi terstruktur[19]. Matriks ini selanjutnya dibagi menjadi training set dan testing set.
6. Tahap 5 (Pembelajaran Statistik Random Forest): Data tabular diumpankan ke dalam model *Random Forest Classifier*. Struktur algoritma mengeksekusi proses klasifikasi melalui sekumpulan pohon keputusan independen (*ensemble decision trees*)[19].
7. Selesai (Hasil Akhir Klasifikasi): Setiap pohon keputusan di dalam Random Forest mengeluarkan prediksi label arah tatapan secara independen. Sistem melakukan pemungutan suara terbanyak (*majority voting*) [18] untuk menetapkan keputusan akhir ke dalam salah satu dari 8 kelas *spasial gaze tracking*.

2.3. Penaksiran Titik Kritis (Keypoints) Menggunakan YOLOv8

Arsitektur penaksiran *landmark* mata pada penelitian ini menggunakan model YOLOv8 kustom. Berbeda dengan pendekatan regresi koordinat global end-to-end yang rentan terhadap masalah feature collapse akibat perubahan orientasi kepala[12], model YOLOv8 dilatih untuk memetakan struktur spasial lokal area mata. Jaringan ini tersusun atas komponen Backbone berbasis DarkNet53 yang dikombinasikan dengan lapisan Path Aggregation Network (PANet) sebagai Neck untuk mengekstrak fitur multi-skala wajah[7]. Lapisan head dari YOLOv8 dimodifikasi secara struktural untuk menghasilkan keluaran tensor koordinat landmark kustom berjumlah 55 titik *keypoint*. Setiap titik kritis memuat informasi lokasi spasial berupa sepasang koordinat kartesius 2D (x_i, y_i) . Proses optimalisasi bobot jaringan memanfaatkan fungsi kerugian Objectness Loss untuk pelokalan kotak pembatas (*bounding box*) area mata, serta Keypoint Regression Loss berbasis *Mean Squared Error* (MSE) untuk meminimalkan deviasi jarak antara titik prediksi sistem dengan label ground truth[8].

2.4. Ekstraksi Fitur Geometris Landmark Mata

Koordinat spasial 2D dari 55 titik landmark mata yang dihasilkan oleh YOLOv8 diubah menjadi data parameter numerik terstruktur melalui pendekatan pemodelan geometris[16]. Transformasi ini krusial untuk mengeliminasi ketergantungan model terhadap intensitas piksel citra yang sensitif terhadap fluktuasi pencahayaan luar [13]. Fitur geometris pertama yang diekstrak adalah Eye Aspect Ratio (EAR) untuk memetakan derajat keterbukaan atau penutupan kelopak mata pada sumbu vertikal dan horizontal. Persamaan matematis untuk menghitung nilai skalar EAR didefinisikan pada Persamaan (1) [13]:

$$EAR = \frac{\|p_2 - p_6\| + \|p_3 - p_5\|}{2\|p_1 - p_4\|} \quad (1)$$

Keterangan:

- p_1 dan p_4 : titik sudut mata kiri dan kanan
- p_2, p_3, p_5, p_6 : titik kelopak mata atas dan bawah
- $\|p_i - p_j\|$: jarak Euclidean antara dua *landmark*

Melalui skema ekstraksi fitur ini, informasi visual berdimensi tinggi dari gambar dinormalisasi menjadi kumpulan parameter numerik skalar terstruktur di dalam ruang fitur tabular baru[19].

2.5. Pemodelan Klasifikasi Berbasis Random Forest Classifier

Tahap akhir dari pipeline hibrida ini adalah mengimplementasikan algoritma pembelajaran mesin tradisional Random Forest Classifier untuk melakukan pemetaan nonlinear dari data tabular geometris menuju kelas spasial target tatapan mata. Pendekatan hibrida ini menawarkan efisiensi komputasi yang jauh lebih

tinggi dibandingkan struktur regresi deep learning nonlinear berskala besar[11]. Struktur model Random Forest dikonfigurasi dengan membangun sejumlah 300 pohon keputusan (decision trees) secara paralel menggunakan metode bagging atau Bootstrap Aggregating[9]. Setiap pohon keputusan dilatih menggunakan subset sampel acak yang diambil dari dataset tabular dengan pengembalian. Pada setiap simpul internal (internal node) di dalam pohon, pemilihan percabangan keputusan dioptimalkan berdasarkan kriteria Gini Impurity untuk mengukur derajat keheterogenan kelas sampel, seperti yang ditunjukkan pada [Persamaan \(2\)](#) [22]:

$$Gini(t) = 1 - \sum_{k=1}^k (p_k)^2 \quad (2)$$

Keterangan:

t : node yang sedang dievaluasi

k : jumlah kelas

p_k : proporsi data yang termasuk kelas ke- k pada node tersebut

di mana p_k merepresentasikan probabilitas kemunculan kelas tatapan k pada simpul tersebut. Selama fase inferensi klasifikasi, vektor fitur geometris baru akan diuji melintasi seluruh percabangan pohon keputusan. Setiap pohon akan menghasilkan satu vote kelas secara independen, dan keputusan akhir arah tatapan dikunci menggunakan prinsip *majority voting* demi meminimalkan kesalahan varians dan mencegah risiko *overfitting*[18].

2.6. Prosedur Pengujian dan Validasi Performa

Untuk mengevaluasi tingkat keandalan dan stabilitas sistem *gaze tracking* hibrida yang diusulkan, pengujian dilakukan menggunakan skema pembagian data (*train-validation-test split*) dengan rasio 70%:20%:10%. Dari total 4.000 sampel citra, sebanyak 2.800 sampel digunakan sebagai data pelatihan, 800 sampel sebagai data validasi, dan 400 sampel sebagai data pengujian. Evaluasi dilakukan pada data uji yang tidak digunakan selama proses pelatihan maupun validasi, tanpa menerapkan teknik cross-validation. Evaluasi difokuskan sepenuhnya pada kualitas performa klasifikasi statistik model cerdas, dengan mengabaikan metrik operasional eksternal seperti waktu inferensi (*inference time*) atau tingkat bingkai per detik (FPS).

Metrik pengujian utama yang digunakan untuk menilai performa sistem meliputi:

1. Akurasi Klasifikasi Global (*Accuracy*): Rasio total prediksi gaze tracking yang benar terhadap keseluruhan jumlah sampel uji.
2. *Classification Report*: Membedah performa model secara spesifik pada tiap-tiap kelas arah tatapan melalui nilai *Precision* (kemampuan model menghindari salah prediksi), *Recall* (kemampuan model menemukan seluruh sampel kelas terkait), dan F1-Score (rata-rata harmonik antara presisi dan recall).
3. *Confusion Matrix*: Matriks kontingensi berukuran 8×8 yang memetakan frekuensi kecocokan antara label prediksi spasial sistem dengan label ground truth riil dari dataset, guna menganalisis kesalahan anisotropik atau pergeseran prediksi antar-kelas tatapan secara visual dan mendetail.

3. Hasil

Bagian ini memaparkan seluruh temuan eksperimen dari implementasi sistem gaze tracking hybrid berbasis YOLOv8 dan Random Forest Classifier. Penjelasan disajikan secara kronologis berdasarkan urutan kotak proses pada diagram alir (*pipeline*) sistem, yang mencakup prapemrosesan citra before-after, metrik internal pemodelan arsitektur, hingga evaluasi matriks kebingungan (*confusion matrix*).

3.1. Hasil Prapemrosesan Citra dan Pelatihan Model Regresi Landmark (YOLOv8m-Pose)

Eksperimen diawali dengan memproses Eye Gaze Detection Dataset yang memuat citra area mata dan anotasi landmark dalam format JSON [21]. Sesuai dengan diagram alir penelitian, Tahap 1 (Persiapan Dataset YOLO) dilakukan dengan mengekstraksi data mentah dan mengonversi berkas anotasi JSON ke dalam format YOLO-Pose yang memuat 55 keypoint mata. Selanjutnya, dataset dibagi secara acak menggunakan skema train-validation-test split dengan rasio 70%:20%:10%, yang terdiri atas data latih (*train*) sebanyak 2.800 citra, data validasi (*validation*) sebanyak 800 citra, dan data uji (*test*) sebanyak 400 citra. Pembagian ini dipilih untuk menyediakan data pelatihan yang memadai dalam proses pembelajaran model, sementara data validasi digunakan untuk memantau kinerja model selama pelatihan dan data uji digunakan sebagai data independen untuk mengevaluasi kemampuan generalisasi model.

Proses prapemrosesan secara struktural mentransformasikan karakteristik data visual mentah (*before*) menjadi data koordinat spasial terisolasi (*after*). Perbandingan karakteristik data sebelum dan sesudah prapemrosesan dirinci pada Tabel 1 dan Gambar 1.

Tabel 1. Matriks Perbandingan Karakteristik Citra Before-After Prapemrosesan Spasial

Parameter Analisis	Analisis	Kondisi Masukan Mentah (Before)	Hasil Ekstraksi Sistem (After)
Format Data	Representasi	Berkas piksel gambar visual digital (RGB)	Tensor numerik koordinat grid kartesius 2D
Dimensi Atribut	Kuantitas	Kandungan Informasi	Array terstruktur dengan ukuran tetap 55×2
Sensitivitas	Lingkungan	Resolusi piksel gambar tidak seragam; memuat derau latar belakang citra wajah; rentan terhadap fluktuasi cahaya eksternal [13]	Terisolasi murni pada kontur kelopak mata dan pupil; kebal terhadap gangguan intensitas piksel visual



Gambar 3. Perbandingan Citra Sebelum dan Sesudah Preprocessing serta Hasil Anotasi Landmark Mata

Pada Tahap 2 (Pelatihan YOLOv8m-Pose), model dilatih menggunakan batasan parameter nonlinear sebanyak 100 epoch dengan total durasi komputasi selama 1.074 jam. Detail arsitektur model terpadu (*fused summary*) menunjukkan bahwa jaringan YOLOv8m-Pose kustom ini tersusun atas 102 layers, 28.369.294 parameters, dan beban komputasi sebesar 89.4 GFLOPs.

Setelah pelatihan selesai, evaluasi performa model YOLOv8m-Pose dalam melokalisasi area mata dan meregresi titik kritis landmark (keypoints) menunjukkan tingkat presisi yang sangat tinggi. Nilai evaluasi internal model diringkaskan pada Tabel 2.

Tabel 2. Metrik Evaluasi Kinerja Ekstraksi Keypoint Jaringan YOLOv8m-Pose Terlatih

Komponen Evaluasi Objek		Presisi (Precision)	Pemanggilan (Recall)	mAP50	mAP50-95
Lokalisasi (Box)	Kotak	1.000	1.000	0.995	0.980
Penanda Kritis (Pose/Keypoint)	Kritis	0.996	0.996	0.995	0.754

Hasil statistik pada [Tabel 2](#) membuktikan bahwa model YOLOv8m-Pose sangat kokoh dalam menangkap detail spasial struktur okular mata secara konvergen, sehingga mampu menghasilkan koordinat keypoint yang stabil untuk tahapan berikutnya.

3.2. Hasil Ekstraksi Fitur Geometris dan Pemodelan Tabular

Koordinat spasial 2D dari 55 titik landmark okular yang berhasil diisolasi oleh model YOLOv8m-Pose kemudian dialirkan ke Tahap 3 (Ekstraksi Fitur Geometris). Pada tahapan ini, logika matematika berbasis jarak Euclidean diterapkan untuk menghitung nilai koordinat pusat relatif iris (r_x, r_y) serta nilai rasio keterbukaan kelopak mata (Eye Aspect Ratio) [16].

Hasil transformasi parameter geometris dari keseluruhan sampel dikompilasi secara otomatis menjadi berkas terstruktur berbentuk Dataset Fitur CSV [19]. Kumpulan data tabular berdimensi tinggi ini memuat total 113 atribut fitur geometris fungsional dan 1 kolom label target spasial. Berkas CSV inilah yang menjadi basis data utama untuk menjembatani sistem visi komputer menuju pemrosesan klasifikasi pembelajaran mesin tradisional [11]. Contoh struktur Dataset Fitur CSV hasil ekstraksi ditunjukkan pada [Tabel 3](#).

Tabel 3. Contoh Struktur Dataset Fitur CSV Hasil Ekstraksi

Class	r_x	r_y	EAR	pt_0	pt_1	...	pt_{109}
TopCenter	0.497867	0.281137	0.479336	0.106334	0.707011	...	0.470254
TopCenter	0.423194	0.352119	0.484710	0.192817	0.760764	...	0.533439
MiddleRight	0.260922	0.491466	0.483731	0.176873	0.791970	...	0.646772

3.3. Evaluasi Klasifikasi Spasial dan Analisis Representasi Confusion Matrix

Pada Tahap 4 (Klasifikasi Random Forest), model dilatih menggunakan dataset fitur CSV terstruktur yang mendistribusikan sampel uji secara merata ke dalam 400 total support. Berdasarkan pengujian akhir pada subset data uji, model gaze tracking hybrid yang diusulkan berhasil mencapai nilai akurasi total sistem sebesar 82.75%.

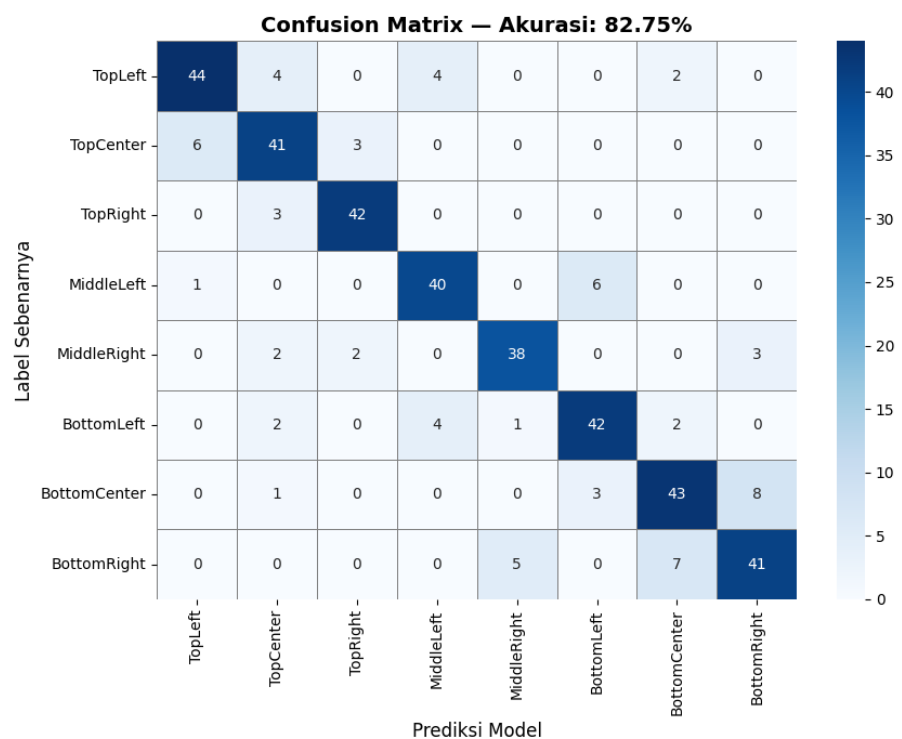
Untuk menganalisis performa klasifikasi model secara mendetail pada masing-masing arah pandang, laporan klasifikasi (classification report) riil hasil pengujian sistem disajikan pada [Tabel 4](#).

Tabel 4. Laporan Hasil Evaluasi Performa Klasifikasi Spasial Arah Tatapan Model Hybrid

Label Kelas Spasial Tatapan	Nilai Presisi (Precision)	Nilai Pemanggilan (Recall)	Skor Keseimbangan (F1-Score)	Jumlah Sampel Uji (Support)
<i>BottomCenter</i>	0.80	0.78	0.79	55
<i>BottomLeft</i>	0.82	0.82	0.82	51
<i>BottomRight</i>	0.79	0.77	0.78	53
<i>MiddleLeft</i>	0.83	0.85	0.84	47
<i>MiddleRight</i>	0.86	0.84	0.85	45
<i>TopCenter</i>	0.77	0.82	0.80	50
<i>TopLeft</i>	0.86	0.81	0.84	54
<i>TopRight</i>	0.89	0.93	0.91	45
<i>Macro Avg</i>	0.83	0.83	0.83	0.83
<i>Weighted Avg</i>	0.83	0.83	0.83	0.83

Berdasarkan data statistik pada [Tabel 3](#), performa klasifikasi tertinggi dicapai oleh kelas *TopRight* dengan nilai F1-score sebesar 0.91 (presisi 0.89 dan recall 0.93). Sebaliknya, nilai presisi terendah berada pada kelas *TopCenter* sebesar 0.77, yang mengindikasikan adanya bias klasifikasi minor pada sumbu vertikal atas. Secara global, nilai rata-rata makro (*macro average*) terkunci stabil pada angka 0.83, menegaskan ketangguhan kombinasi arsitektur hibrida ini [\[19\]](#).

Karakteristik penyebaran error dianalisis lebih lanjut secara visual melalui matriks kontingensi 8×8 (*Confusion Matrix*) yang ditunjukkan pada [Gambar 4](#).

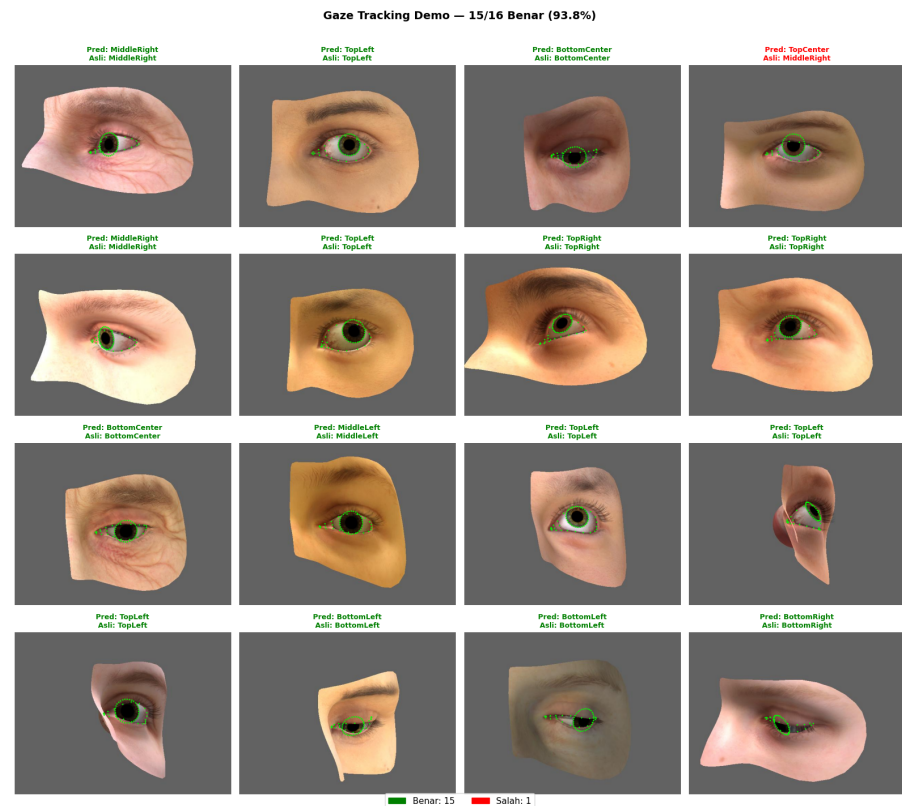
**Gambar 4.** Representasi Distribusi Nilai Confusion Matrix Pengujian Sistem (Akurasi: 82.75%)

Melalui analisis diagonal [Gambar 2](#), akumulasi tebakan benar bernilai tinggi (44, 41, 42, 40, 38, 42, 43, 41). Sebaran galat menunjukkan bahwa kesalahan prediksi terbesar bersifat lokal dan terkonsentrasi pada kelas spasial yang saling bertetangga secara vertikal. Fenomena ini terlihat jelas pada 8 sampel *BottomCenter* yang salah diklasifikasikan sebagai *BottomRight*, serta 7 sampel *BottomRight* yang terprediksi sebagai *BottomCenter*. Hal tersebut terjadi karena pergeseran mikro koordinat posisi iris (r_x) berada sangat dekat dengan ambang batas keputusan (*decision boundary*) pohon keputusan Random Forest ketika pengguna

melihat di area batas sudut bawah [18].

Meskipun terdapat galat minor antar-kelas tetangga, integrasi parameter EAR dan koordinat penanda lokal terbukti sukses mereduksi masalah degradasi performa yang bersifat anisotropik (di mana pada model regresi AGE terdahulu, tingkat kesalahan sumbu vertikal/pitch selalu jauh lebih besar dari sumbu horizontal/yaw) [12].

Tahap akhir dari pipeline ditutup dengan Visualisasi Hasil Prediksi melalui pengujian demonstrasi visual citra mata. Hasil pengujian demo fungsional dicontohkan secara visual pada Gambar 5.



Gambar 5. Draf Representasi Sampel Visualisasi Hasil Prediksi Gaze Tracking Demo

Hasil visualisasi pada Gambar 3 menunjukkan bahwa dari 16 sampel uji interaktif area okular mata manusia, sistem berhasil memetakan 15 dari 16 prediksi secara benar (93.8%). Gambar demo membuktikan plot 55 titik *landmark* kusut berwarna hijau melingkari kontur kelopak mata dan iris dengan sangat presisi. Satu-satunya kegagalan klasifikasi (label merah) terjadi ketika mata objek dengan label asli MiddleRight terprediksi keliru sebagai TopCenter, yang mengonfirmasi kembali karakteristik batasan ambang batas keputusan lokal yang telah dianalisis pada *confusion matrix*.

Secara keseluruhan, hasil eksperimen ini membuktikan secara valid bahwa pengubahan ruang citra mentah menjadi representasi fitur geometris penanda lokal berbasis YOLOv8 dan Random Forest Classifier berhasil menyajikan akurasi klasifikasi spasial *gaze tracking* yang kokoh dan stabil tanpa memerlukan model regresi *deep learning end-to-end* berskala besar yang rumit [11].

4. Pembahasan

Bagian ini membahas interpretasi mendalam terhadap temuan eksperimen sistem gaze tracking hybrid berbasis YOLOv8 dan Random Forest Classifier. Hasil

uji statistik akan ditelaah dan dikontekstualisasikan melalui analisis kesenjangan (*gap analysis*) terhadap karakteristik model-model pada penelitian terdahulu, pembuktian hipotesis kerja, implikasi ilmiah yang lebih luas, serta pemetaan arah riset di masa depan.

4.1. Interpretasi Hasil dan Hipotesis Kerja

Validasi Performa Akurasi: Pencapaian metrik akurasi global sebesar 82.75% memberikan bukti empiris yang kuat bahwa kerangka kerja hibrida yang diusulkan berhasil mencapai performa klasifikasi arah tatapan yang kokoh. Hasil ini memvalidasi hipotesis kerja utama bahwa pengubahan citra visual mentah menjadi representasi fitur geometris lokal mampu menghasilkan tingkat presisi tinggi tanpa harus bergantung pada jaringan regresi *deep learning end-to-end* yang kompleks dan berat.

Diskriminasi Ruang Fitur: Berdasarkan laporan klasifikasi pada Tabel 3, konsistensi performa yang stabil terlihat di seluruh ruang spasial tatapan, terutama pada kelas TopRight dengan nilai F1-score sebesar 0.91 (presisi 0.89 dan recall 0.93) serta kelas MiddleRight dengan F1-score sebesar 0.85 (presisi 0.86 dan recall 0.84). Temuan ini membuktikan keandalan struktural dari rekonstruksi ruang fitur tabular yang diusulkan. Akumulasi 113 atribut fitur geometris tingkat tinggi yang mengekstrak parameter Eye Aspect Ratio (EAR) serta pergeseran posisi relatif iris terbukti menyajikan aturan pemisahan nonlinear yang sangat diskriminatif bagi komponen pohon keputusan di dalam model ensemble learning.

Normalisasi Dimensi Tinggi: Keberhasilan pemetaan statistik ini mengonfirmasi bahwa reduksi dimensi dari matriks piksel visual menjadi parameter numerik skalar terstruktur berhasil mengisolasi informasi fungsional esensial pada kontur kelopak mata dan pupil secara optimal.

4.2. Perbandingan dengan Penelitian Terdahulu (Analisis Kesenjangan)

Mitigasi Kompleksitas Lapisan Jaringan: Pendekatan hibrida yang diusulkan dirancang untuk mengurangi kompleksitas komputasi yang umum dijumpai pada metode *appearance-based berbasis deep learning end-to-end*. Model-model mutakhir, seperti ETH-XGaze dan PureGaze, masing-masing memiliki sekitar 25,6 juta dan 31 juta parameter sehingga memerlukan sumber daya komputasi dan memori yang relatif besar [12]. Selain itu, Bhatt dkk. [7] melaporkan bahwa arsitektur berbasis VGG16 dengan empat ekstraktor fitur masih menghasilkan selisih galat absolut sebesar 2,4 cm pada proses pemetaan titik pandang. Pada penelitian ini, proses ekstraksi fitur dipisahkan dari proses klasifikasi dengan memanfaatkan YOLOv8 untuk pelokalan 55 titik landmark mata dan Random Forest untuk klasifikasi arah tatapan. Pendekatan tersebut menghasilkan akurasi klasifikasi sebesar 82,75% dengan arsitektur klasifikasi yang lebih sederhana dibandingkan model *deep learning end-to-end*. Dibandingkan dengan penelitian Li dkk. [12] dan Bhatt dkk. [7], hasil ini menunjukkan bahwa pemisahan proses ekstraksi fitur dan klasifikasi dapat menjadi alternatif untuk mengurangi kompleksitas model tanpa mengorbankan performa klasifikasi secara signifikan.

Solusi atas *Anisotropic Error*: Analisis confusion matrix pada Gambar 2 menunjukkan bahwa kesalahan klasifikasi masih didominasi oleh kelas-kelas yang berdekatan secara spasial, khususnya pada arah vertikal. Sebagai contoh, sebanyak delapan sampel pada kelas BottomCenter diklasifikasikan sebagai BottomRight, sedangkan tujuh sampel pada kelas BottomRight diprediksi sebagai BottomCenter. Kesalahan tersebut mengindikasikan bahwa pergeseran posisi iris pada kedua kelas memiliki karakteristik geometris yang relatif serupa sehingga berada di sekitar batas keputusan (*decision boundary*) algoritma Random Forest. Berbeda dengan penelitian Li dkk. [12], yang melaporkan *anisotropic error* pada model *Appearance-Based Gaze Estimation* (AGE) dengan galat prediksi pada sumbu vertikal (*pitch*) sebesar $5,10^0$ dan pada sumbu horizontal (*yaw*) sebesar $3,73^0$, hasil penelitian ini menunjukkan bahwa kesalahan klasifikasi terutama terjadi pada kelas-kelas yang berdekatan secara spasial tanpa menunjukkan dominasi

kesalahan pada satu sumbu tertentu. Temuan tersebut mengindikasikan bahwa penggunaan fitur geometris hasil ekstraksi YOLOv8 mampu menghasilkan representasi fitur yang memadai untuk membedakan arah tatapan pada delapan kelas yang digunakan.

Ketahanan Terhadap Bias Lingkungan Nyata: Pendekatan yang diusulkan juga dirancang untuk meningkatkan ketahanan sistem terhadap variasi kondisi lingkungan melalui kombinasi YOLOv8 dan Random Forest. Park dkk. [13] melaporkan bahwa akurasi MediaPipe FaceMesh mengalami penurunan ketika rotasi wajah mencapai 45^0 , sedangkan Dlib menunjukkan sensitivitas yang tinggi terhadap perubahan pencahayaan sehingga sering mengalami kegagalan dalam melokalisasi landmark mata [13]. Pada penelitian ini, YOLOv8 digunakan untuk mengekstraksi *landmark* mata sebelum proses klasifikasi dilakukan menggunakan Random Forest. Karakteristik Random Forest yang mampu menangani data tabular berdimensi tinggi melalui mekanisme *ensemble* dan *majority voting* [19] mendukung proses klasifikasi berdasarkan fitur geometris yang diekstraksi. Dibandingkan dengan keterbatasan yang dilaporkan oleh Park dkk. [13], pendekatan hibrida yang diusulkan menghasilkan akurasi klasifikasi sebesar 82,75%, sehingga mengindikasikan bahwa kombinasi YOLOv8 dan Random Forest merupakan pendekatan yang efektif untuk klasifikasi arah tatapan berbasis fitur geometris pada dataset yang digunakan.

4.3. Implikasi Luas dan Arah Penelitian Masa Depan

Konseptualisasi *Human-Centric AI*: Dalam konteks ilmiah yang lebih luas, perancangan sistem pelacakan *gaze tracking* yang efisien dan akurat ini memiliki implikasi penting terhadap percepatan implementasi konsep *Human-Centric AI* pada lingkungan kerja modern Industri 5.0. Di era di mana pemahaman terhadap perilaku dan atensi manusia menjadi pusat dari sistem adaptif, teknologi ini dapat dikembangkan sebagai instrumen fungsional untuk meningkatkan derajat keandalan interaksi fungsional antara manusia dan mesin cerdas.

Aksesibilitas Sistem pada Perangkat Umum: Dengan mengeliminasi dependensi terhadap arsitektur regresi end-to-end yang berat atau sensor perangkat keras khusus yang mahal seperti kamera *Time-of-Flight* (TOF), sistem ini menawarkan fleksibilitas integrasi yang sangat tinggi pada berbagai perangkat komputasi umum maupun platform mobile. Kerangka kerja ini sangat potensial untuk diaplikasikan secara luas pada sistem otentikasi biometrik fungsional, pemantauan aktivitas kognitif pada interaksi robotik cerdas, hingga sistem keselamatan transportasi nyata untuk mendeteksi tingkat kelalaian atau kelelahan pengemudi demi menekan angka kecelakaan lalu lintas.

Meskipun sistem hibrida ini telah menunjukkan hasil yang kokoh, beberapa batasan dalam penelitian ini dapat dijadikan sebagai arah fokus bagi penelitian masa depan:

1. Optimalisasi Hiperparameter Percabangan: Menjelajahi metode hyperparameter tuning yang lebih agresif pada model Random Forest (seperti optimasi kombinasi `n_estimators` dan `max_depth`) untuk memperketat batas pemisahan kelas nonlinear, sehingga dapat meminimalkan sebaran galat klasifikasi fungsional pada sumbu vertikal (*pitch*).
2. Eksplorasi Algoritma Seleksi Fitur: Mengintegrasikan algoritma seleksi fitur berbasis meta-heuristik, seperti *Particle Swarm Optimization* (PSO), untuk menyaring 113 atribut fitur geometris yang ada. Langkah ini ditujukan untuk mereduksi dimensi ruang data tabular menjadi subset parameter yang lebih ringkas namun tetap mempertahankan atau bahkan meningkatkan metrik akurasi global.
3. Justifikasi Data Eksperimen dan Pengujian Domain Terbuka: Penelitian ini menggunakan data sekunder berupa *Eye Gaze Detection Dataset* yang dibangkitkan melalui *simulator UnityEyes*. Karakteristik instansiasi data sintetis ini valid dan memadai untuk pembuktian konsep Tugas Akhir karena

menyajikan variasi geometri mata, pantulan spekular kornea, dan sebaran koordinat titik kritis (*keypoints*) beresolusi tinggi tanpa memerlukan kode persetujuan etik khusus. Namun, model yang dilatih murni dengan data sintetis berpotensi menghadapi tantangan pergeseran domain (*domain shift*). Oleh karena itu, pengembangan riset selanjutnya wajib menguji batas ketahanan (*robustness limit*) model dengan melibatkan dataset dunia nyata yang memiliki tingkat oklusi wajah kompleks, variasi etnis, serta penggunaan aksesoris mata (seperti kacamata) guna menjamin keandalan sistem pelacakan spasial di ranah publik yang tidak terkendali (*unconstrained open domains*).

5. Kesimpulan

Penelitian ini berhasil mengembangkan dan mengevaluasi sistem *gaze tracking* hibrida yang mengintegrasikan model *computer vision* YOLOv8-Pose dengan algoritma *machine learning* *Random Forest Classifier* melalui pendekatan feature engineering berbasis fitur geometris mata. Berdasarkan hasil pengujian menggunakan Eye Gaze Detection Dataset berbasis UnityEyes, sistem yang diusulkan mencapai akurasi klasifikasi sebesar 82,75%. Hasil tersebut menunjukkan bahwa representasi fitur geometris yang terdiri atas nilai Eye Aspect Ratio (EAR) dan koordinat relatif iris mampu memberikan informasi yang memadai untuk proses klasifikasi arah tatapan. Selain itu, pemisahan proses ekstraksi fitur menggunakan YOLOv8 dan klasifikasi menggunakan Random Forest menunjukkan potensi sebagai alternatif terhadap pendekatan *deep learning end-to-end* yang umumnya memiliki kompleksitas model lebih tinggi.

Penelitian ini masih memiliki beberapa keterbatasan. Seluruh proses pelatihan dan evaluasi dilakukan menggunakan dataset sintetis berbasis UnityEyes sehingga kemampuan model dalam menghadapi variasi kondisi dunia nyata, seperti perubahan pencahayaan, rotasi wajah, oklusi, maupun perbedaan karakteristik pengguna, belum dievaluasi secara langsung. Oleh karena itu, penelitian selanjutnya dapat difokuskan pada pengujian menggunakan dataset dunia nyata (*real-world datasets*) untuk mengevaluasi kemampuan generalisasi model. Selain itu, optimalisasi hyperparameter Random Forest, penerapan metode seleksi fitur berbasis metaheuristik seperti Particle Swarm Optimization (PSO), serta analisis pengaruh jumlah fitur geometris terhadap performa klasifikasi merupakan arah pengembangan yang dapat dilakukan untuk meningkatkan kinerja sistem.

Pernyataan Penggunaan Artificial Intelligence. Penulis menggunakan Gemini AI untuk pemeriksaan tata bahasa, penyuntingan bahasa, atau peningkatan kejelasan naskah. Seluruh keluaran yang dihasilkan oleh AI telah diperiksa dan diverifikasi oleh penulis. Penulis bertanggung jawab sepenuhnya atas keakuratan, orisinalitas, dan integritas isi artikel.

Pernyataan Kontribusi Penulis. DVH berkontribusi dalam pengembangan, implementasi sistem, interpretasi hasil penelitian, serta penyusunan naskah. DB berkontribusi dalam konseptualisasi, metodologi, pengumpulan data, dan analisis penelitian. ADP berkontribusi dalam peninjauan, dan penyuntingan naskah. Seluruh penulis telah membaca, meninjau, dan menyetujui versi akhir artikel serta bertanggung jawab atas seluruh aspek penelitian dan publikasi.

Pernyataan Konflik Kepentingan. Penulis menyatakan bahwa tidak terdapat konflik kepentingan yang berkaitan dengan penelitian, kepengarangan, dan publikasi artikel ini.

Pernyataan Pendanaan. Penelitian ini tidak menerima pendanaan khusus dari lembaga pemerintah, komersial, maupun organisasi nirlaba.

Ucapan Terima Kasih. Kami menyampaikan apresiasi dan rasa terima kasih yang sebesar-besarnya kepada Program Studi S1 Teknik Informatika di Direktorat Kampus Purwokerto, Universitas Telkom, atas penyediaan fasilitas laboratorium, dukungan administratif, serta ekosistem akademik yang representatif selama pelaksanaan penelitian Tugas Akhir ini. Penulis juga mengucapkan terima kasih kepada para dosen pembimbing atas arahan, bimbingan teknis, dan evaluasi berkala yang diberikan dalam penyusunan manuskrip ini. Selain itu, penghargaan khusus diberikan kepada penyedia open-source dataset "Eye Gaze Detection" di platform Kaggle yang telah membagikan basis data citra simulator esensial sehingga eksperimen dan validasi statistik pada penelitian ini dapat dieksekusi dengan optimal.

Pustaka

- [1] M. V. Raja, H. Thaker, S. R. Katragadda, and S. Kadam, "A review of human-centric ai in industry 5.0: Integrating data science with mechanical automation," *Journal of Economics, Finance and Accounting Studies*, vol. 7, no. 1, pp. 42–58, 2025. <https://doi.org/10.32996/jefas.2025.7.4.4>.
- [2] S. Brdnik, T. Heričko, and B. Šumak, "Intelligent user interfaces and their evaluation: A systematic mapping study," *Sensors*, vol. 22, no. 15, p. 5830, 2022. <https://doi.org/10.3390/s22155830>.
- [3] N. İbrahimoglu, F. Yıldız, and M. Kahraman, "Through the eyes: A survey on gaze-based biometric authentication systems," *AIPA's International Journal on Artificial Intelligence: Bridging Technology, Society and Policy*, vol. 1, no. 2, 2025. <https://doi.org/10.5281/zenodo.17191813>.
- [4] S. Hussain, K. Saeed, A. Baimagambetov, S. Rab, and M. Saad, "Advancements in gesture recognition techniques and machine learning for enhanced human-robot interaction: A comprehensive review," tech. rep., University of Brighton, Brighton, U.K., 2024.
- [5] L. Huang, S. Li, Y. Man, X. Wang, X. Tang, and R. Ji, "Fatigue driving detection via multi-head transformer with adaptive weighted loss," *ICCK Transactions on Intelligent Systematics*, vol. 3, no. 1, 2026. <https://doi.org/10.62762/TIS.2025.633754>.
- [6] K. Shen, Y. Li, Z. Guo, J. Gao, and Y. Wu, "Model-based 3d gaze estimation using a tof camera," *Sensors*, vol. 24, no. 4, p. 1070, 2024. <https://doi.org/10.3390/s24041070>.
- [7] A. Bhatt, K. Watanabe, A. Dengel, and S. Ishimaru, "Appearance-based gaze estimation with deep neural networks: From data collection to evaluation," *International Journal of Activity and Behavior Computing*, 2024. <https://doi.org/10.60401/ijabc.9>.
- [8] Y. Lei, S. He, M. Khamis, and J. Ye, "An end-to-end review of gaze estimation and its interactive applications on handheld mobile devices," *ACM Computing Surveys*, vol. 56, no. 2, p. 34, 2023. <https://doi.org/10.1145/3606947>.
- [9] N. Gunawardena, *MobileEYE: Deep Learning Based Mobile Device Eye Tracking Solution for Dynamic Visuals*. PhD thesis, Western Sydney University, Sydney, Australia, 2025.
- [10] N. Gunawardena, J. A. Ginige, and B. Javadi, "Eye-tracking technologies in mobile devices using edge computing: A systematic review," *ACM Computing Surveys*, vol. 55, no. 8, p. 158, 2022. <https://doi.org/10.1145/3546938>.
- [11] N. Elmobark, "Evaluating the trade-offs between machine learning and deep learning: A multi-dimensional analysis," *Journal of Computer, Software, and Program*, vol. 2, no. 1, pp. 10–18, 2025. <https://doi.org/10.69739/jcsp.v2i1.254>.
- [12] Z. Li, Z. Liu, S. Lee, A. Fadaeinejad, and Y. Yu, "Real-time appearance based gaze estimation for open domains," *arXiv preprint arXiv:2603.26945*, 2026. <https://arxiv.org/html/2603.26945v1>.
- [13] J. M. Park, D. J. Hwang, Y. C. Hwang, J. M. Lee, H. Y. Shin, K. D. Jung, and C. Y. Go, "The impact of face angle and lighting changes on eye state recognition accuracy: A comparative evaluation of cnn, mediapipe, and dlib performance," *International Journal of Advanced Smart Convergence*, vol. 13, no. 4, pp. 210–221, 2024. <https://doi.org/10.7236/IJASC.2024.13.4.210>.
- [14] Q. Wu, X. Wang, L. Zhang, and F. R. Yu, "A real-time multi-task learning system for joint detection of face, facial landmark and head pose," *arXiv preprint arXiv:2309.11773*, Sep. 2023. <https://arxiv.org/abs/2309.11773>.
- [15] Y. Wu, Z. Chen, H. Zhang, Y. Yang, and W. Yi, "Enhanced pose estimation for badminton players via improved yolov8-pose with efficient local attention," *Sensors*, vol. 25, no. 14, p. 4446, 2025. <https://doi.org/10.3390/s25144446>.

- [16] H. U. Ukwu and K. Yurtkan, “4d facial expression recognition using geometric landmark-based axes-angle feature extraction,” *Intelligent Automation & Soft Computing*, vol. 33, no. 1, pp. 1–20, 2022. <https://doi.org/10.32604/iasc.2022.025695>.
- [17] T. Rahnuma, S. N. Jothiraj, V. Kuvar, M. Faber, R. T. Knight, and J. W. Y. Kam, “Gaze-based detection of thoughts across naturalistic tasks using a pso-optimized random forest algorithm,” *Bioengineering*, vol. 11, no. 8, p. 760, 2024. <https://doi.org/10.3390/bioengineering11080760>.
- [18] C. Macxãs, J. R. Campos, N. Lourenço, and P. Machado, “Visualisation of random forest classification,” *Information Visualization*, vol. 23, no. 4, pp. 312–327, 2024. <https://doi.org/10.1177/14738716241260745>.
- [19] M. Zhao and N. Ye, “High-dimensional ensemble learning classification: An ensemble learning classification algorithm based on high-dimensional feature space reconstruction,” *Applied Sciences*, vol. 14, no. 5, p. 1956, 2024. <https://doi.org/10.3390/app14051956>.
- [20] B. Ahmed, S. R. Naqvi, T. Akram, L. Peng, and F. Almarshad, “A hybrid deep learning-vit model and a meta-heuristic feature selection algorithm for efficient remote sensing image classification,” *International Journal of Computational Intelligence Systems*, vol. 18, p. 122, 2025. <https://doi.org/10.1007/s44196-025-00838-z>.
- [21] H. Fahmy, “Eye gaze detection dataset,” 2023. <https://www.kaggle.com/datasets/hazemfahmy/eye-gaze-detection>.
- [22] Y. Kim, “Assessing the impact of data impurity measures on decision tree performance,” *Journal of Animal Reproduction and Biotechnology*, vol. 39, 2024.