

Deteksi Kelelahan Pengemudi Menggunakan YOLOv11 dan CNN Berbasis Facial Landmark

Denny Pribadi¹, Moh. Yusup Nurhadi², Apip Supiandi³

^{1,2,3} Universitas Bina Sarana Informatika

e-mail: denny.dpi@bsi.ac.id¹, m.yusupnh06@gmail.com², apip.aup@bsi.ac.id³

Abstrak

Kecelakaan lalu lintas di Indonesia meningkat signifikan. Data tahun 2022 menunjukkan 61 persen kasus disebabkan oleh faktor manusia, khususnya kelelahan pengemudi. Penelitian ini bertujuan mengembangkan sistem deteksi kelelahan pengemudi berbasis penanda wajah secara waktu nyata. Sistem mengintegrasikan algoritma *You Only Look Once* versi 11 untuk deteksi objek wajah dan mata. Sistem juga menggunakan *Convolutional Neural Network* untuk mengklasifikasi kondisi mengantuk berdasarkan ekstraksi 68 titik penanda wajah. Peneliti menggunakan *dataset* awal sebanyak 2.200 citra wajah pengemudi yang setelah augmentasi menjadi 4.477 citra. Proses pra-pemrosesan meliputi normalisasi resolusi, augmentasi data, dan ekstraksi titik penanda wajah menggunakan pustaka *dlib*. Hasil evaluasi menunjukkan model deteksi objek mencapai presisi 92,9 persen, recall 93,1 persen, dan *mean Average Precision* pada ambang 50 sebesar 97,4 persen. Model klasifikasi mencapai akurasi 75,65 persen dengan skor F1 sebesar 74,55 persen pada 100 siklus pelatihan. Integrasi kedua model ini memberikan peringatan dini yang andal. Sistem ini terbukti efektif mendeteksi kelelahan pengemudi berbasis penanda wajah secara akurat.

Kata Kunci: Kelelahan pengemudi, *Facial landmark*, *You Only Look Once*, *Convolutional Neural Network*, Deteksi waktu nyata.

Abstract

Traffic accidents in Indonesia increased significantly. Data from 2022 shows that 61 percent of cases were caused by human factors, specifically driver fatigue. This research aimed to develop a driver fatigue detection system based on facial landmarks in real-time. The system integrated the You Only Look Once version 11 algorithm for face and eye object detection. The system also used a Convolutional Neural Network to classify drowsy conditions based on the extraction of 68 facial landmarks. The initial dataset consisted of 2,200 driver face images and increased to 4,477 images after augmentation. Preprocessing included resolution normalization, data augmentation, and facial landmark extraction using the dlib library. Evaluation results showed the object detection model achieved 92.9 percent precision, 93.1 percent recall, and a mean Average Precision at 50 threshold of 97.4 percent. The classification model achieved 75.65 percent accuracy with an F1-score of 74.55 percent at 100 training epochs. The integration of these two models provided reliable early warnings. This system proved effective in detecting driver fatigue based on facial landmarks accurately.

Keywords: *Driver fatigue, Facial landmark, You Only Look Once, Convolutional Neural Network, Real-time detection.*

1. Pendahuluan

Kecelakaan lalu lintas masih menjadi salah satu permasalahan utama keselamatan transportasi di Indonesia. Berdasarkan data Korps Lalu Lintas Kepolisian Republik Indonesia, sepanjang tahun 2022 tercatat sebanyak 94.617 kasus kecelakaan lalu lintas, dengan sekitar 61% di antaranya disebabkan oleh faktor manusia (*human error*), termasuk kelelahan dan rasa kantuk saat mengemudi

(Kementerian Perhubungan Republik Indonesia, 2023). (Badan Pusat Statistik, 2025) juga melaporkan bahwa jumlah kecelakaan lalu lintas meningkat menjadi 150.906 kasus pada tahun 2024. Kondisi kelelahan pengemudi (*driver fatigue*) dapat menurunkan tingkat kewaspadaan, memperlambat waktu reaksi, serta mengurangi kemampuan pengambilan keputusan sehingga diperlukan sistem deteksi dini sebagai upaya meningkatkan keselamatan berkendara.

Perkembangan teknologi *Advanced Driver Assistance System* (ADAS) mendorong berkembangnya berbagai metode *Driver Monitoring System* berbasis *computer vision* yang mampu memantau kondisi pengemudi secara non-invasif (Arnita et al., 2022). Perkembangan *deep learning* semakin meningkatkan performa sistem tersebut. (Zhao et al., 2024) menjelaskan bahwa *Convolutional Neural Network* (CNN) memiliki kemampuan mengekstraksi fitur citra secara otomatis dengan performa klasifikasi yang tinggi. Selain itu, (Lin & Zuo, 2024) menunjukkan bahwa metode pengenalan wajah berbasis *deep learning* mampu mempertahankan akurasi deteksi pada berbagai kondisi pencahayaan, sedangkan (Kao & Chan, 2022) membuktikan bahwa kombinasi informasi wajah dan mata menghasilkan performa deteksi kantuk yang lebih baik dibandingkan penggunaan satu jenis fitur saja.

Berbagai penelitian telah menerapkan CNN maupun *You Only Look Once* (YOLO) untuk mendeteksi kantuk pengemudi. (Akbar et al., 2022) menunjukkan bahwa CNN mampu mengidentifikasi kondisi mengantuk melalui pembelajaran karakteristik wajah secara otomatis, sedangkan (Armin et al., 2023) melaporkan bahwa YOLOv8 memberikan performa deteksi objek yang baik untuk aplikasi *real-time*. Pengembangan lebih lanjut dilakukan oleh (Arava & Sundaram, 2024) melalui integrasi *facial landmark* dengan arsitektur YOLO yang dimodifikasi sehingga meningkatkan representasi fitur wajah dan ketahanan sistem terhadap variasi pose maupun ekspresi.

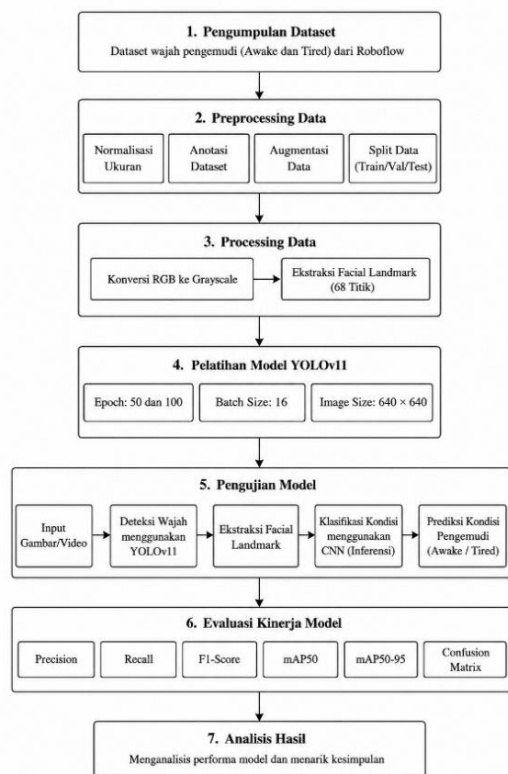
Representasi fitur wajah juga berperan penting dalam meningkatkan akurasi deteksi kelelahan. (Bao & Xu, 2024) menjelaskan bahwa ekstraksi 68 *facial landmark* menggunakan pustaka *dlib* mampu menghasilkan representasi wajah yang stabil. (Bachtiar & Wafi, 2021) menunjukkan bahwa *facial landmark* meningkatkan performa klasifikasi ekspresi wajah, sedangkan (Xiao et al., 2022) membuktikan efektivitasnya dalam mengenali indikator kelelahan melalui perubahan posisi mata dan mulut. Temuan tersebut diperkuat oleh (Kao & Chan, 2022), yang menunjukkan bahwa kombinasi fitur wajah dan mata menghasilkan representasi yang lebih informatif dibandingkan penggunaan salah satu fitur secara terpisah.

Meskipun demikian, sebagian besar penelitian masih berfokus pada deteksi objek atau klasifikasi secara terpisah. Penelitian berbasis CNN umumnya menggunakan citra wajah secara langsung sebagai masukan model (Akbar et al., 2022), sedangkan penelitian berbasis YOLO lebih menitikberatkan pada deteksi objek tanpa memanfaatkan representasi fitur wajah secara optimal (Armin et al., 2023). Oleh karena itu, penelitian ini mengusulkan integrasi YOLOv11 sebagai model deteksi wajah, 68 *facial landmark* sebagai representasi fitur, dan CNN sebagai model klasifikasi kondisi pengemudi ke dalam kategori *awake* dan *tired*. Integrasi tersebut diharapkan mampu menghasilkan sistem deteksi kelelahan pengemudi yang bekerja secara *real-time* dengan performa yang lebih baik serta mendukung pengembangan teknologi keselamatan berkendara berbasis *computer vision*. Sebagai pembanding, penelitian Mughni Rahmat Alaziz (2024) menggunakan YOLOv8 dan *facial landmark* untuk mendeteksi kelelahan pengemudi dengan *precision*, *recall*, dan *F1-score* sekitar 96%. Perbandingan tersebut dilakukan berdasarkan metode, arsitektur model, dan performa yang dilaporkan, bukan melalui pengujian langsung menggunakan *dataset* yang sama. Penelitian ini menggunakan YOLOv11 yang dikombinasikan dengan CNN serta dataset Roboflow sebanyak 2.200 citra yang setelah augmentasi menjadi 4.477 citra. Perbedaan *dataset*, teknik *preprocessing*, pembagian data, dan parameter pelatihan menyebabkan hasil tidak dapat dibandingkan secara langsung. Oleh karena itu, penelitian ini tidak bertujuan menyatakan bahwa YOLOv11 lebih unggul secara mutlak dibandingkan YOLOv8, melainkan menguji performa YOLOv11 dan CNN pada *dataset* dan skenario yang digunakan.

2. Metode Penelitian

2.1 Tahapan Penelitian

Penelitian ini menggunakan metode eksperimen yang terdiri atas tujuh tahapan. Alur penelitian ditunjukkan pada Gambar 1.



Gambar 1. Tahapan Penelitian
Sumber: Penulis (2025)

2.2 Pengumpulan Dataset

Dataset yang digunakan merupakan data sekunder berupa citra wajah pengemudi yang diperoleh melalui *platform Roboflow*. *Dataset* terdiri atas dua kelas, yaitu *awake* (tidak mengantuk) dan *tired* (mengantuk). Seluruh citra telah diberi anotasi sesuai kelas masing-masing sehingga dapat digunakan sebagai data pelatihan maupun pengujian model. *Dataset* awal terdiri atas 2.200 citra. Setelah melalui proses augmentasi, jumlah citra menjadi 4.477 citra. Selanjutnya, *dataset* dibagi menjadi tiga *subset*, yaitu data *training* sebanyak 3.940 citra (88%), data *validation* sebanyak 269 citra (6%), dan data *testing* sebanyak 268 citra (6%). Data *training* digunakan untuk proses pembelajaran model, data *validation* digunakan untuk memantau dan mengevaluasi performa model selama proses pelatihan, sedangkan data *testing* digunakan untuk mengevaluasi performa akhir model terhadap data yang tidak digunakan dalam proses pembelajaran. *Dataset* dipilih karena memiliki variasi kondisi pencahayaan, posisi wajah, serta ekspresi pengemudi yang cukup beragam sehingga mampu

merepresentasikan kondisi nyata saat berkendara.

2.3 Preprocessing Data

Tahap *preprocessing* bertujuan meningkatkan kualitas data sebelum digunakan dalam proses pelatihan model. Proses ini meliputi normalisasi ukuran citra, anotasi objek, pembagian dataset menjadi data latih, validasi, dan pengujian (*split data*), serta augmentasi data.

Augmentasi dilakukan menggunakan beberapa teknik, yaitu rotasi acak, *horizontal flipping*, rotasi 90°, dan *cropping*. Pada data *training* diterapkan *preprocess_input* dan augmentasi berupa *rotation* 20°, *width_shift_range* = 0,2, *height_shift_range* = 0,2, *zoom_range* = 0,2, *brightness_range* = [0,8, 1,2], *horizontal_flip* = *True*, serta *fill_mode* = '*nearest*'. Data *validation* dan *testing* hanya menggunakan *preprocess_input* tanpa augmentasi. Data *training* menggunakan *shuffle* = *True*, sedangkan *validation* dan *testing* menggunakan *shuffle* = *False*. Tahapan ini bertujuan meningkatkan keragaman data sehingga model menjadi lebih *robust* terhadap perubahan orientasi objek, variasi pencahayaan, maupun perubahan posisi wajah pengemudi.

2.4 Processing Data

Setelah proses *preprocessing*, citra diproses melalui dua tahapan utama, yaitu konversi warna dan ekstraksi fitur wajah.

a. Konversi RGB ke Grayscale

Citra RGB terlebih dahulu dikonversi menjadi citra *grayscale* untuk menyederhanakan informasi warna sehingga proses komputasi menjadi lebih efisien tanpa mengurangi informasi penting yang diperlukan pada tahap ekstraksi fitur wajah.

b. Ekstraksi Facial Landmark

Ekstraksi *facial landmark* dilakukan menggunakan pustaka *dlib* untuk memperoleh titik-titik penting pada wajah, seperti posisi mata, alis, hidung, dan mulut. *Landmark* tersebut digunakan sebagai representasi fitur wajah yang berkaitan dengan kondisi kelelahan pengemudi. Proses ini bertujuan meningkatkan kemampuan sistem dalam mengenali perubahan ekspresi wajah yang menjadi indikator utama kondisi mengantuk.

2.5 Pelatihan Model YOLOv11

Model YOLOv11 digunakan sebagai model deteksi objek untuk mengenali kondisi wajah pengemudi berdasarkan dua kelas, yaitu *awake* dan *tired*. Pelatihan dilakukan menggunakan ukuran citra (*image size*) 640 × 640 piksel, *batch size* 16, serta dua skenario jumlah *epoch*, yaitu 50 *epoch* dan 100 *epoch*.

Perbandingan kedua skenario dilakukan untuk mengetahui pengaruh jumlah epoch terhadap performa model. Selama proses pelatihan dilakukan pemantauan nilai *loss*, *precision*, *recall*, *mAP50*, dan *mAP50-95* hingga model mencapai kondisi konvergen.

2.6 Pengujian Model Menggunakan CNN

Setelah model YOLOv11 selesai dilatih, proses pengujian dilakukan menggunakan data uji. Pada tahap ini, YOLOv11 digunakan untuk mendeteksi area wajah pengemudi pada setiap citra atau frame video. Selanjutnya, hasil deteksi tersebut diproses melalui ekstraksi *facial landmark* dan digunakan sebagai masukan bagi *Convolutional Neural Network* (CNN) untuk mengklasifikasikan kondisi pengemudi ke dalam dua kelas, yaitu *awake* dan *tired*.

CNN pada penelitian ini tidak digunakan pada proses pelatihan model deteksi, melainkan pada tahap inferensi untuk melakukan klasifikasi kondisi pengemudi berdasarkan fitur wajah yang telah diperoleh. Dengan pendekatan ini, YOLOv11 berperan sebagai model deteksi objek, sedangkan CNN berfungsi sebagai model klasifikasi sehingga kedua model memiliki fungsi yang saling melengkapi dalam proses deteksi kelelahan pengemudi. Proses konversi *facial landmark* menjadi input CNN dimulai dari citra wajah pengemudi. Citra tersebut kemudian diproses untuk memperoleh *facial landmark*, yaitu titik-titik tertentu pada wajah yang merepresentasikan bagian penting seperti area mata dan wajah. Setiap *landmark* direpresentasikan dalam bentuk koordinat numerik *x* dan *y*. Koordinat tersebut kemudian disusun secara terstruktur berdasarkan urutan titik *landmark* sehingga menghasilkan representasi fitur yang konsisten untuk setiap citra. Selanjutnya, representasi *landmark* tersebut digunakan sebagai informasi fitur untuk proses klasifikasi menggunakan CNN. CNN

mempelajari pola dari representasi tersebut untuk membedakan dua kondisi pengemudi, yaitu *awake* dan *tired*. Alur prosesnya adalah: Citra Pengemudi → YOLOv11 → Deteksi Wajah → *Facial Landmark* → Ekstraksi Fitur → CNN → Klasifikasi *Awake/Tired*.

2.7 Evaluasi Kinerja Model

Evaluasi dilakukan untuk mengukur kemampuan sistem dalam mendeteksi kondisi kelelahan pengemudi. Kinerja YOLOv11 dievaluasi menggunakan metrik *Precision*, *Recall*, *F1-score*, *mAP50*, dan *mAP50-95*, sedangkan hasil klasifikasi CNN dianalisis menggunakan *accuracy*, *precision*, *recall*, *F1-score*, serta *confusion matrix*. Selain itu, dilakukan perbandingan performa model pada pelatihan 50 *epoch* dan 100 *epoch* untuk menentukan konfigurasi yang memberikan hasil terbaik.

3. Hasil dan Pembahasan

3.1 Evaluasi Model Deteksi Objek YOLOv11

Pelatihan model YOLOv11 dilakukan menggunakan dua skenario, yaitu 50 *epoch* dan 100 *epoch*, untuk mengevaluasi pengaruh jumlah iterasi terhadap performa deteksi wajah pengemudi. Evaluasi dilakukan menggunakan metrik *precision*, *recall*, *F1-score*, *mAP@50*, dan *mAP@50-95*, sehingga dapat memberikan gambaran mengenai kemampuan model dalam mendeteksi objek secara akurat pada berbagai tingkat *Intersection over Union* (IoU).

Tabel 1. Perbandingan Kinerja Model YOLOv11 pada 50 dan 100 Epoch

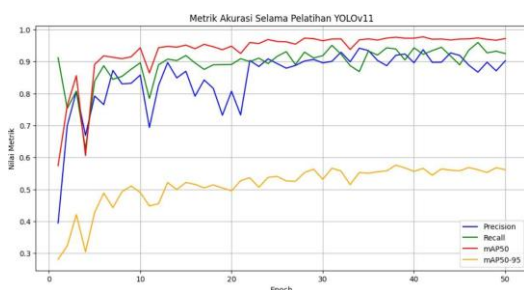
Metrik	50 Epoch	100 Epoch
<i>Precision</i>	91,8%	92,9%
<i>Recall</i>	93,9%	93,1%
<i>F1-Score</i>	92,8%	93,0%
<i>mAP@50</i>	97,6%	97,4%
<i>mAP@50-95</i>	57,6%	58,0%

Sumber: Penulis (2025)

Berdasarkan hasil evaluasi pada Tabel 1, model yang dilatih selama 100 *epoch* memberikan performa yang lebih baik dibandingkan model 50 *epoch*. Nilai *precision* meningkat dari 91,8% menjadi 92,9%, sedangkan *F1-score* meningkat dari 92,8% menjadi 93,0%. Nilai *recall* relatif

stabil pada kisaran 93%, yang menunjukkan bahwa penambahan jumlah *epoch* mampu meningkatkan kemampuan model dalam mengurangi kesalahan prediksi positif (*false positive*) tanpa mengurangi sensitivitas deteksi objek. Selain itu, peningkatan $mAP@50-95$ dari 57,6% menjadi 58,0% mengindikasikan bahwa model memiliki kemampuan lokalisasi objek yang lebih baik pada berbagai ambang *IoU*.

Nilai $mAP@50$ yang mencapai 97,4% menunjukkan bahwa model mampu mendeteksi wajah pengemudi dengan tingkat akurasi yang sangat baik pada ambang *IoU* sebesar 0,50. Sementara itu, nilai $mAP@50-95$ yang lebih rendah menunjukkan bahwa ketelitian lokalisasi objek masih dapat ditingkatkan ketika menggunakan ambang *IoU* yang lebih ketat. Hasil tersebut menunjukkan bahwa konfigurasi pelatihan 100 *epoch* memberikan keseimbangan yang lebih baik antara ketepatan deteksi dan kemampuan generalisasi model. Temuan ini sejalan dengan penelitian (Armin et al., 2023) yang menunjukkan bahwa model YOLO mampu menghasilkan performa deteksi yang tinggi pada sistem pemantauan pengemudi berbasis *computer vision*. Meskipun penelitian tersebut menggunakan YOLOv8, hasil penelitian ini menunjukkan bahwa penggunaan YOLOv11 tetap mampu mempertahankan tingkat *precision* dan *recall* yang tinggi sehingga layak diterapkan pada sistem deteksi kelelahan pengemudi secara *real-time*.



Gambar 2. Grafik Pelatihan Model YOLOv11 pada 100 Epoch
Sumber: Penulis (2025)

Gambar 2 menunjukkan perkembangan metrik evaluasi selama proses pelatihan. Nilai *precision*, *recall*, $mAP@50$, dan $mAP@50-95$ mengalami peningkatan secara bertahap hingga mencapai kondisi yang relatif stabil pada akhir pelatihan. Pola tersebut menunjukkan bahwa model telah mencapai kondisi

konvergen dan mampu mempelajari karakteristik objek secara konsisten. Stabilitas kurva pelatihan menunjukkan bahwa proses optimasi telah mencapai kondisi konvergen sehingga risiko *underfitting* maupun *overfitting* dapat diminimalkan. Kondisi tersebut sesuai dengan karakteristik pelatihan model *deep learning* yang dijelaskan oleh (Zhao et al., 2024), yaitu performa model akan meningkat hingga mencapai titik konvergen ketika proses ekstraksi fitur berlangsung secara optimal.

3.2 Evaluasi Model Klasifikasi CNN

Setelah wajah pengemudi berhasil dideteksi oleh model YOLOv11, citra hasil deteksi diproses melalui ekstraksi *facial landmark* untuk memperoleh representasi fitur wajah. Selanjutnya, fitur tersebut digunakan sebagai masukan bagi *Convolutional Neural Network* (CNN) untuk mengklasifikasikan kondisi pengemudi ke dalam dua kelas, yaitu *awake* dan *tired*. Berbeda dengan evaluasi YOLOv11 yang menggunakan metrik *precision*, *recall*, dan *mAP*, performa CNN dievaluasi menggunakan *accuracy*, *precision*, *recall*, *F1-score*, serta *confusion matrix* karena berfungsi sebagai model klasifikasi.

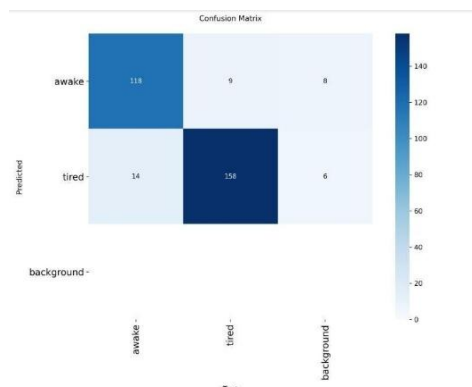
Tabel 2. Perbandingan Kinerja Model CNN pada 50 dan 100 *Epoch*

Metrik	Kelas	50 Epoch	100 Epoch
<i>Precision</i>	Mengantuk	68,0%	77,3%
	Tidak Mengantuk	80,0%	74,0%
<i>Recall</i>	Mengantuk	86,0%	72,0%
	Tidak Mengantuk	58,0%	78,0%
Akurasi Keseluruhan		72,0%	75,7%

Sumber: Penulis (2025)

Hasil pengujian menunjukkan bahwa model CNN memperoleh akurasi sebesar 75,65% pada konfigurasi pelatihan 100 *epoch*. Nilai *precision* kelas *tired* mencapai 77,29%, sedangkan *recall* sebesar 72,00% dengan *F1-score* sebesar 74,55%. Pada kelas *awake*, model menghasilkan *precision* sebesar 74,12%, *recall* sebesar 79,25%, dan *F1-score* sebesar 76,60%. Nilai akurasi tersebut menunjukkan bahwa CNN mampu mempelajari pola perubahan ekspresi wajah

yang berkaitan dengan kondisi kelelahan pengemudi. Temuan ini sejalan dengan penelitian (Akbar et al., 2022) yang menunjukkan bahwa CNN efektif digunakan untuk mengklasifikasikan kondisi kantuk berdasarkan karakteristik wajah pengemudi.



Gambar 3. Confusion Matrix Hasil Pengujian CNN pada 100 Epoch
Sumber: Penulis (2025)

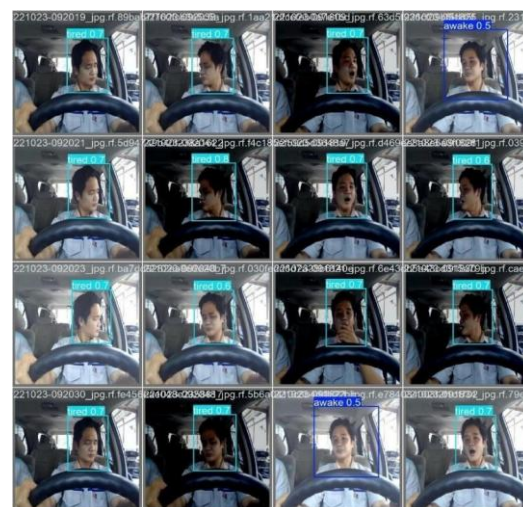
Gambar 3 menunjukkan bahwa sebagian besar data berhasil diklasifikasikan dengan benar pada kedua kelas. Namun demikian, masih terdapat sejumlah *false positive* dan *false negative*, yang menunjukkan adanya kesalahan klasifikasi antara kondisi *awake* dan *tired*. Kesalahan tersebut kemungkinan dipengaruhi oleh beberapa faktor, seperti variasi pencahayaan, orientasi wajah, kualitas citra, maupun kemiripan ekspresi wajah antara kondisi mengantuk dan tidak mengantuk. Kondisi serupa juga dilaporkan oleh (Nggiku et al., 2023), yang menyatakan bahwa perubahan posisi kepala dan intensitas cahaya dapat memengaruhi akurasi sistem deteksi kantuk berbasis *computer vision*.

Penggunaan 68 *facial landmark* memberikan representasi geometris wajah yang lebih informatif melalui posisi mata, alis, hidung, dan mulut sehingga membantu CNN mempelajari perubahan ekspresi yang berkaitan dengan kondisi kelelahan pengemudi. Hasil ini sejalan dengan penelitian (Bachtiar & Wafi, 2021) yang menunjukkan bahwa *facial landmark* mampu meningkatkan kualitas representasi fitur dibandingkan penggunaan citra wajah secara langsung. Selain itu, (Xiao et al., 2022) menjelaskan bahwa informasi geometris yang diperoleh dari posisi mata dan mulut merupakan indikator penting dalam mendeteksi kondisi kelelahan

pengemudi. Meskipun demikian, peningkatan akurasi masih dapat dilakukan melalui penambahan variasi dataset, optimasi arsitektur CNN, maupun pengembangan teknik ekstraksi fitur yang lebih adaptif terhadap perubahan kondisi lingkungan.

3.3 Implementasi Sistem dan Analisis Kinerja

Sistem deteksi kelelahan pengemudi diimplementasikan dalam bentuk aplikasi berbasis web yang mampu melakukan deteksi kondisi pengemudi secara *real-time* menggunakan kamera. Implementasi ini mengintegrasikan model YOLOv11 sebagai detektor wajah dengan proses ekstraksi 68 titik *facial landmark* menggunakan pustaka *dlib*, kemudian dilanjutkan dengan klasifikasi kondisi pengemudi menggunakan *Convolutional Neural Network* (CNN). Antarmuka sistem dirancang untuk menampilkan hasil deteksi secara langsung sehingga memudahkan pengguna dalam memantau kondisi pengemudi.



Gambar 4. Implementasi Sistem Deteksi Kelelahan Pengemudi Berbasis YOLOv11 dan CNN

Sumber: Penulis (2025)

Gambar 4 menunjukkan antarmuka aplikasi yang menampilkan hasil deteksi wajah beserta status kondisi pengemudi, yaitu *awake* atau *tired*. Setelah wajah berhasil dideteksi oleh YOLOv11, sistem melakukan ekstraksi *facial landmark* untuk memperoleh representasi fitur wajah, kemudian CNN mengklasifikasikan kondisi pengemudi berdasarkan fitur tersebut. Hasil

klasifikasi ditampilkan secara langsung pada layar sebagai bentuk peringatan dini terhadap potensi kelelahan pengemudi.

Secara keseluruhan, implementasi sistem menunjukkan bahwa integrasi YOLOv11, *facial landmark*, dan CNN mampu bekerja secara berurutan dalam mendukung deteksi kelelahan pengemudi. YOLOv11 memberikan kemampuan deteksi wajah yang cepat, sedangkan CNN memanfaatkan informasi dari *facial landmark* untuk melakukan klasifikasi kondisi pengemudi. Pendekatan ini menunjukkan potensi yang baik untuk diterapkan sebagai sistem pemantauan kelelahan pengemudi secara *real-time*, terutama sebagai teknologi pendukung keselamatan berkendara.

Hasil implementasi menunjukkan bahwa integrasi model deteksi objek dan model klasifikasi dapat mendukung proses pemantauan pengemudi secara *real-time*. Pendekatan tersebut sejalan dengan perkembangan sistem *Driver Monitoring System* (DMS) yang memanfaatkan kombinasi deteksi wajah dan analisis fitur wajah untuk meningkatkan keselamatan berkendara sebagaimana dijelaskan oleh (Kao & Chan, 2022).

4. Kesimpulan

Penelitian ini berhasil mengembangkan sistem deteksi kelelahan pengemudi berbasis YOLOv11, *facial landmark*, dan *Convolutional Neural Network* (CNN). YOLOv11 digunakan untuk mendeteksi wajah pengemudi secara cepat dan akurat, sedangkan CNN memanfaatkan fitur hasil ekstraksi 68 titik *facial landmark* untuk mengklasifikasikan kondisi pengemudi ke dalam kategori *awake* dan *tired*. Integrasi kedua model tersebut memungkinkan sistem bekerja secara berurutan dalam mendukung proses deteksi kelelahan pengemudi secara *real-time*.

Hasil evaluasi menunjukkan bahwa model YOLOv11 memberikan performa deteksi yang sangat baik dengan nilai *precision* sebesar 92,9%, *recall* sebesar 93,1%, *F1-score* sebesar 93,0%, dan *mAP@50* sebesar 97,4% pada konfigurasi pelatihan 100 *epoch*. Sementara itu, CNN menghasilkan tingkat akurasi sebesar 75,65% dalam mengklasifikasikan kondisi pengemudi berdasarkan fitur *facial landmark*. Hasil tersebut menunjukkan bahwa sistem yang diusulkan mampu mendeteksi wajah secara akurat sekaligus

melakukan klasifikasi kondisi pengemudi dengan performa yang cukup baik.

Penelitian ini memberikan kontribusi berupa penerapan arsitektur yang mengintegrasikan YOLOv11 sebagai model deteksi objek dengan CNN sebagai model klasifikasi berbasis *facial landmark* untuk mendukung sistem deteksi kelelahan pengemudi. Pendekatan tersebut menunjukkan potensi yang baik untuk diterapkan sebagai teknologi pendukung keselamatan berkendara melalui pemantauan kondisi pengemudi secara *real-time*.

Penelitian selanjutnya dapat difokuskan pada peningkatan akurasi klasifikasi melalui penambahan variasi dataset, optimasi arsitektur CNN, pengujian pada perangkat *edge*, serta implementasi langsung pada sistem pemantauan kendaraan agar mampu beroperasi secara lebih adaptif pada berbagai kondisi lingkungan.

Referensi

- Akbar, H., Aryani, D., & Junaedi, S. (2022). Deteksi Kantuk Pengendara Mobil Berbasis Citra Menggunakan Convolutional Neural Networks. *INFORMATIKA*, 10(1), 13–20. <https://doi.org/https://doi.org/10.36987/informatika.v10i1.2454>
- Arava, M., & Sundaram, D. M. (2024). Enhancing Driver Drowsiness Detection: A Fusion of Facial Landmarks and Modified YOLOv5 Architecture. *Intelligent Systems and Applications in Engineering*, 12(11s), 437–449.
- Armin, E. U., Edra, A. P., Alifin, F. I., Sadidan, I., Sary, I. P., & Latifa, U. (2023). Performa Model YOLOv8 untuk Deteksi Kondisi Mengantuk pada Pengendara Mobil. *BRAHMANA: Jurnal Penerapan Kecerdasan Buatan*, 5(1), 67–76. <https://doi.org/https://doi.org/10.30645/brahmana.v5i1.279>
- Arnita, Marpaung, F., Aulia, F., Suryani, N., & Nabila, R. C. (2022). *Computer Vision Dan Pengolahan Citra Digital*. Pustaka Aksara.
- Bachtiar, F. A., & Wafi, M. (2021). Komparasi Metode Klasifikasi untuk Deteksi Ekspresi Wajah dengan Fitur Facial Landmark. *Jurnal Teknologi Informasi Dan Ilmu Komputer (JTIIK)*, 8(5), 949–956.

- <https://doi.org/10.25126/jtiik.202184434>
- Badan Pusat Statistik. (2025). *Jumlah Kecelakaan, Korban Mati, Luka Berat, Luka Ringan, dan Kerugian Materi, 2024*.
<https://www.bps.go.id/id/statistics-table/2/NTEzIzI=/jumlah-kecelakaan-korban-mati-luka-berat-luka-ringan-dan-kerugian-materi.html>
- Bao, Y., & Xu, W. (2024). Design and Implementation of a Fatigue Detection System Based on Dlib for Driver Facial Features. *Electronics, Communications and Networks*, 792–799.
<https://doi.org/10.3233/FAIA231266>
- Kao, I., & Chan, C. (2022). Comparison of Eye and Face Features on Drowsiness Analysis. *Sensors*, 22(17), 1–20.
<https://doi.org/https://doi.org/10.3390/s22176529>
- Kementerian Perhubungan Republik Indonesia. (2023). *Tekan Angka Kecelakaan Lalu Lintas, Kemenhub Ajak Masyarakat Beralih ke Transportasi Umum dan Utamakan Keselamatan Berkendara*.
<https://dephub.go.id/post/read/tekan-angka-kecelakaan-lalu-lintas,-kemenhub-ajak-masyarakat-beralih-ke-transportasi-umum-dan-utamakan-keselamatan-berkendara>
- Lin, N., & Zuo, Y. (2024). Advancing Driver Fatigue Detection in Diverse Lighting Conditions for Assisted Driving Vehicles with Enhanced Facial Recognition Technologies. *PLOS ONE*, 19(7), 1–25.
<https://doi.org/10.1371/journal.pone.0304669>
- Nggiku, C. K. U., Rabi, A., & Subairi. (2023). Deteksi Kantuk Untuk Keamanan Berkendara Berbasis Pengolahan Citra. *JEETech: Journal of Electrical Engineering And Technology*, 4(1), 48–56.
<https://doi.org/https://doi.org/10.32492/jeetech.v4i1.4107>
- Xiao, W., Liu, H., Ma, Z., Chen, W., Sun, C., & Shi, B. (2022). Fatigue Driving Recognition Method Based on Multi-Scale Facial Landmark Detector. *Electronics*, 11(24), 1–18.
<https://doi.org/https://doi.org/10.3390/electronics11244103>
- Zhao, X., Wang, L., Zhang, Y., Han, X., Deveci, M., & Parmar, M. (2024). A Review of Convolutional Neural Networks in Computer Vision. In *Artificial Intelligence Review* (Vol. 57, Issue 4). Springer Netherlands.
<https://doi.org/10.1007/s10462-024-10721-6>