JARVIS
JURNAL TEKNIK ELEKTRO

Evaluasi Kinerja Model Yolov8 dalam Deteksi Visual Kantuk dan Distraksi Pengemudi

Romi Rahmat Hidayatulloh [1], Andri Ulus Rahayu [2], Imam Taufiqurrahman [3], Okyza Maherdy Prabowo [4]

[1,2,3] *Program Studi Teknik Elektro, Universitas SIliwangi, Jl. Mugarsari, Kec. Tamansari, Kota Tasikmalaya, Indonesia*
[4] *Program Studi Sistem Informasi, STMIK Amik Bandung, Jl. Jakarta No.28, Kec. Batununggal, Kota Bandung, Indonesia*

INFORMASI ARTIKEL

Received: June, 01, 2025
Reviewed: June, 23, 2025
Available online: June, 30, 2025

KORESPONDEN

E-mail: 217002020@student.unsil.ac.id

A B S T R A C T

Traffic accidents caused by human factors such as drowsiness and distraction remain a leading cause of road fatalities globally. Computer vision-based approaches offer promising solutions to identify driver inattention through visual cues. This study aims to evaluate the performance of the YOLOv8 object detection model in recognizing visual indicators of driver drowsiness and distraction, including open eyes, closed eyes, yawning, and mobile phone usage. A custom dataset was manually annotated using the Roboflow platform, and model training was carried out using the Ultralytics YOLOv8 framework on Google Colab. Data augmentation techniques such as horizontal flipping, shear transformation, color jittering, and noise injection were applied to enhance the model's robustness. Evaluation was performed on the validation set using performance metrics including precision, recall, F1-score, and mean Average Precision (mAP). The results demonstrate that the model achieves high accuracy across all object classes. This study is limited to evaluating model performance and does not cover real-time implementation or system integration.

KEYWORD:
computer vision, drowsiness detection, driver distraction, model evaluation, object detection, YOLOv8.

A B S T R A K

Kecelakaan lalu lintas akibat faktor manusia seperti kantuk dan distraksi masih menjadi penyebab utama kematian di jalan raya secara global. Pendekatan berbasis visi komputer menawarkan solusi yang menjanjikan untuk mengidentifikasi ketidakfokusan pengemudi melalui indikator visual. Penelitian ini bertujuan untuk mengevaluasi performa model deteksi objek YOLOv8 dalam mengenali indikator visual kantuk dan distraksi, yaitu mata terbuka, mata tertutup, menguap, dan penggunaan ponsel. Dataset yang digunakan merupakan data kustom yang telah dianotasi secara manual menggunakan platform Roboflow, dan pelatihan model dilakukan dengan pustaka Ultralytics YOLOv8 di Google Colab. Teknik augmentasi seperti flipping horizontal, transformasi shear, color jittering, dan penambahan noise diterapkan untuk meningkatkan ketangguhan model. Evaluasi dilakukan pada data validasi menggunakan metrik precision, recall, F1-score, dan mean Average Precision (mAP). Hasil menunjukkan bahwa model mencapai akurasi tinggi pada semua kelas objek. Penelitian ini terbatas pada evaluasi performa model dan tidak mencakup implementasi sistem secara nyata.

KATA KUNCI:
computer vision, deteksi kantuk, deteksi objek, distraksi pengemudi, evaluasi model, YOLOv8.

PENDAHULUAN

Kecelakaan lalu lintas merupakan salah satu penyebab utama kematian di Indonesia. Pada tahun 2021, tercatat 103.645 kasus kecelakaan yang mengalami peningkatan sebesar 3,6% dibandingkan tahun sebelumnya. Menurut laporan KNKT, sekitar 80% kecelakaan tersebut disebabkan oleh kelelahan pengemudi yang memicu kondisi microsleep, yaitu tertidur singkat tanpa disadari saat berkendara[1]. Fakta ini menunjukkan urgensi pentingnya penelitian yang berfokus pada deteksi dini terhadap kondisi kantuk dan distraksi pengemudi.

Kemajuan teknologi di bidang pengolahan citra dan visi komputer telah membuka peluang besar dalam pengembangan sistem pemantauan pengemudi secara otomatis. Salah satu pendekatan yang menjanjikan adalah penggunaan model deteksi objek berbasis deep learning, seperti You Only Look Once (YOLO), yang dikenal memiliki kecepatan dan akurasi tinggi dalam mengenali objek. Dalam konteks deteksi kondisi pengemudi, model ini dapat digunakan untuk mengidentifikasi tanda-tanda visual seperti mata tertutup, ekspresi menguap, penggunaan ponsel, dan kondisi mata terbuka.

Penelitian ini secara khusus bertujuan untuk mengevaluasi performa model YOLOv8 dalam mendeteksi empat jenis objek visual yang berkaitan dengan kondisi kantuk dan distraksi, yaitu mata terbuka, mata tertutup, ekspresi menguap, dan penggunaan ponsel. Model dilatih menggunakan dataset kustom yang telah diberi anotasi secara manual, serta diuji menggunakan metrik evaluasi standar seperti precision, recall, F1-score, dan mean Average Precision (mAP).

Penelitian ini dibatasi pada tahap evaluasi performa model, tanpa mencakup implementasi sistem nyata atau integrasi ke dalam perangkat kendaraan. Oleh karena itu, hasil penelitian ini memberikan gambaran awal mengenai kemampuan YOLOv8 dalam mendeteksi indikator visual kondisi pengemudi berdasarkan data pengujian.

TINJAUAN PUSTAKA

Deteksi Objek Dengan Deep Learning

Deteksi objek merupakan salah satu cabang penting dalam bidang visi komputer yang bertujuan untuk mengidentifikasi dan menentukan lokasi objek pada suatu citra atau video. Berbeda dengan klasifikasi citra yang hanya memberikan label terhadap keseluruhan gambar, deteksi objek mencakup dua tugas utama secara simultan, yaitu klasifikasi objek dan regresi koordinat dari kotak pembatas (*bounding box*). Teknik ini telah diterapkan secara luas pada berbagai bidang, termasuk pengawasan keamanan, kendaraan otonom, serta analisis perilaku manusia.

Perkembangan deteksi objek mengalami kemajuan signifikan dengan diperkenalkannya *Convolutional Neural Network* (*CNN*), yang menggantikan pendekatan berbasis fitur manual dengan proses pembelajaran representasi secara *end-to-end*. Arsitektur *CNN* menjadi fondasi utama dalam pengembangan berbagai model modern untuk deteksi objek[2]

Beberapa pendekatan berbasis *CNN* yang memiliki pengaruh besar antara lain *Faster R-CNN*, *Single Shot MultiBox Detector* (*SSD*), dan *You Only Look Once* (*YOLO*). *Faster R-CNN* merupakan pendekatan dua tahap yang memisahkan proses pencarian wilayah objek dengan proses klasifikasi akhir. Arsitektur ini menggunakan *Region Proposal Network* (*RPN*) untuk menghasilkan kandidat wilayah objek, yang kemudian diproses oleh jaringan klasifikasi dan regresi koordinat. Model ini dikenal memiliki akurasi tinggi, namun dengan waktu inferensi yang relatif lambat, yaitu sekitar 5 *frame per second* (fps) saat menggunakan arsitektur *VGG-16 [3]*.

SSD merupakan pendekatan satu tahap yang memanfaatkan fitur multi-skala serta *anchor boxes* untuk mendeteksi objek dalam berbagai ukuran. Model ini mampu melakukan deteksi dengan lebih cepat, mencapai kecepatan sekitar 58 fps pada *dataset* VOC2007, dengan rata-rata presisi (*mean Average Precision / mAP*) sebesar 72,1 persen pada resolusi 300 × 300 piksel[3].

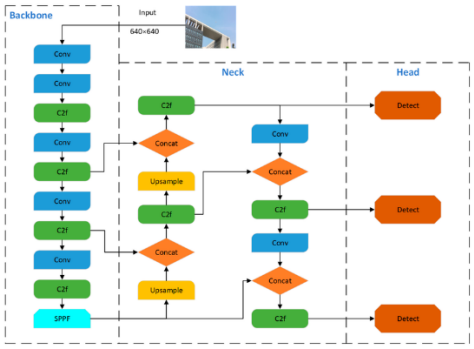
YOLO memperkenalkan pendekatan deteksi satu tahap dengan membagi citra ke dalam beberapa grid dan langsung memprediksi kotak pembatas beserta kelas objek pada setiap grid tersebut. Model ini dirancang untuk efisiensi tinggi dan mampu mencapai kecepatan sekitar 45 fps untuk versi dasar, dan hingga 155 fps untuk versi ringan seperti *Fast YOLO*, dengan akurasi yang kompetitif pada *dataset* VOC2007 [4].

Seiring waktu, berbagai varian dari pendekatan-pendekatan tersebut terus dikembangkan guna meningkatkan akurasi, efisiensi komputasi, serta kemampuan generalisasi. Inovasi-inovasi tersebut menjadi dasar dalam pengembangan YOLOv8, yaitu model deteksi objek modern yang menjadi fokus utama pada penelitian ini.

Arsitektur YOLOv8

YOLOv8 merupakan versi terbaru dari seri *You Only Look Once*, yang dikembangkan oleh Ultralytics dengan basis *PyTorch*. Model ini diklasifikasikan sebagai *single-stage detector*, sehingga selama inferensi hanya memerlukan satu kali proses untuk memprediksi kotak pembatas dan kelas objek secara bersamaan [5]. Salah satu aspek arsitektur yang membedakan *YOLOv8* dari pendahulunya adalah penggunaan kepala deteksi *anchor-free*, yang menghilangkan kebutuhan akan kotak jangkar (*anchor boxes*). Penerapan teknik ini terbukti menyederhanakan pelatihan dan mengurangi

kompleksitas regresi koordinat, sehingga meningkatkan efisiensi keseluruhan model [5]. Struktur arsitektur YOLOv8 dapat dilihat pada Gambar 1.



Gambar 1. Struktur Jaringan Yolov8

Sumber: Zhai et al., Electronics, 2023

Selain itu, *YOLOv8* juga telah mengintegrasikan modul C2f (*Cross Stage Partial with fused bottlenecks*) dalam bagian *backbone*, menggantikan modul C3 yang sebelumnya digunakan pada *YOLOv5*. Modul C2f ini dikabarkan mampu meningkatkan ekstraksi fitur spasial dari berbagai skala objek [5]. Bagian *neck* model mengadopsi kombinasi *Feature Pyramid Network* (FPN) dan *Path Aggregation Network* (PANet), yang memungkinkan agregasi fitur berlapis secara efisien untuk mendukung deteksi objek dengan ukuran yang bervariasi [6].

Pengorganisasian model secara modular juga diusung, yaitu dengan memisahkan antara *backbone* dan *head* agar memudahkan adaptasi konfigurasi dan pengujian model. Dukungan ekspor ke berbagai format inferensi seperti ONNX dan TorchScript turut disediakan guna melengkapi fleksibilitas integrasi dalam berbagai platform komputasi, misalnya server atau perangkat lokal yang serupa [5].

Dalam penelitian ini, *YOLOv8* dipilih karena arsitekturnya yang seimbang antara akurasi tinggi, efisiensi komputasi, dan ukuran model yang relatif sederhana. Kriteria tersebut ditegaskan melalui berbagai kajian awal yang menunjukkan bahwa *YOLOv8* mampu memberikan hasil deteksi objek yang reliabel dalam konteks citra pengemudi, dengan kebutuhan sumber daya yang tidak berlebihan [7].

Evaluasi Kinerja Model

Dalam sistem deteksi objek berbasis *deep learning*, performa model umumnya dievaluasi menggunakan sejumlah metrik standar. Metrik-metrik ini digunakan untuk mengukur kualitas prediksi model dalam mengenali dan mengklasifikasikan objek pada citra uji. Beberapa metrik evaluasi yang paling umum digunakan antara lain *precision*, *recall*, *F1-score*, dan *mean Average Precision* (*mAP*).

Precision didefinisikan sebagai rasio antara jumlah prediksi positif yang benar (*true positive*) dengan total jumlah prediksi positif yang dihasilkan model. Nilai *precision* yang tinggi menunjukkan bahwa sebagian besar prediksi positif yang dihasilkan adalah benar. Sementara itu, *recall* merupakan rasio antara jumlah prediksi positif yang benar dengan total jumlah objek positif yang sebenarnya ada dalam citra. Metrik ini mengukur sejauh mana model mampu mendeteksi seluruh objek yang relevan.

F1-score merupakan metrik harmonik yang menggabungkan *precision* dan *recall* ke dalam satu nilai komposit. Nilai ini digunakan untuk memberikan penilaian yang lebih seimbang, terutama dalam kondisi di mana terdapat ketimpangan antara jumlah objek positif dan negatif. Nilai *F1-score* yang tinggi mengindikasikan keseimbangan yang baik antara akurasi dan kelengkapan deteksi.

Salah satu metrik yang paling banyak digunakan dalam deteksi objek adalah *mean Average Precision* (*mAP*). Metrik ini mengukur rata-rata akurasi deteksi untuk semua kelas objek pada berbagai nilai ambang batas *Intersection over Union* (*IoU*). *IoU* sendiri adalah ukuran seberapa besar kesesuaian antara kotak prediksi dan kotak ground-truth. Umumnya, evaluasi mAP dilakukan pada ambang batas IoU ≥ 0.5 (*mAP@0.5*) atau pada rentang 0.5 hingga 0.95 (*mAP@[.5:.95]*) sesuai standar COCO dataset [8].

Penggunaan metrik-metrik ini memungkinkan evaluasi performa model secara kuantitatif dan objektif, serta menjadi dasar utama dalam membandingkan efektivitas berbagai pendekatan deteksi objek dalam literatur.

Tanda Visual Kantuk dan Distraksi

Tanda visual kantuk dan distraksi merupakan indikator perilaku yang dapat dikenali melalui ekspresi wajah dan aktivitas pengemudi. Berbagai penelitian menyatakan bahwa kondisi kantuk dapat diidentifikasi melalui ciri-ciri seperti peningkatan frekuensi kedipan lambat (*blink frequency*), rasio penutupan kelopak mata (*Percentage of Eyelid Closure over the Pupil* atau *PERCLOS*), serta frekuensi menguap yang tinggi. Indikator-indikator tersebut telah terbukti relevan dalam deteksi kantuk secara visual pada pengemudi kendaraan [9].

Selain itu, distraksi visual akibat penggunaan ponsel selama berkendara telah dilaporkan sebagai salah satu penyebab utama terganggunya perhatian pengemudi. Aktivitas tersebut menyebabkan pengalihan pandangan dari jalan dan penurunan waktu reaksi terhadap kondisi lalu lintas, sehingga meningkatkan risiko kecelakaan secara signifikan[10].

Dalam penelitian ini, tiga objek visual dipilih sebagai representasi kondisi tidak normal pada pengemudi, yaitu mata tertutup, menguap, dan ponsel.

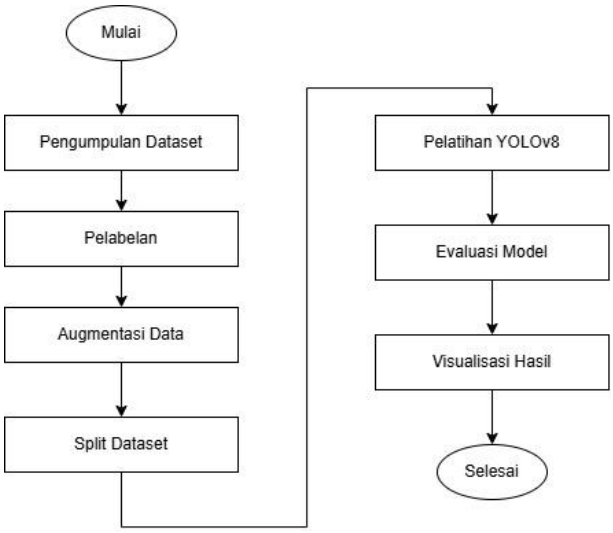
Pemilihan objek-objek tersebut didasarkan pada kajian literatur yang menyatakan bahwa ekspresi mata tertutup dan menguap berkaitan erat dengan kantuk, sedangkan penggunaan ponsel merupakan bentuk distraksi visual yang nyata.

Label untuk masing-masing objek diterapkan secara manual pada dataset citra wajah pengemudi, mengikuti pendekatan *event-based annotation*. Mata tertutup didefinisikan sebagai kondisi di mana kelopak menutupi pupil selama lebih dari 500 milidetik, sedangkan menguap diidentifikasi dari bukaan mulut yang signifikan selama lebih dari 1 detik. Sementara itu, penggunaan ponsel dikenali dari keberadaan objek ponsel di tangan atau posisi yang mengganggu pandangan ke depan [10].

Dengan mendasarkan klasifikasi pada ketiga tanda visual tersebut, sistem ini diharapkan mampu mendeteksi potensi gangguan perhatian yang dapat berdampak pada keselamatan berkendara.

METODE

Pengemudi berbasis pengolahan citra menggunakan algoritma YOLOv8. Proses dimulai dari pengumpulan data berupa citra wajah pengemudi yang direkam secara langsung melalui kamera dengan mempertimbangkan berbagai variasi pose wajah, ekspresi kantuk, dan gangguan visual seperti penggunaan ponsel. Data citra yang terkumpul kemudian dilakukan pelabelan secara manual menggunakan platform Roboflow. Setiap citra dianotasi dengan bounding box untuk menandai objek-objek penting seperti mata terbuka, mata tertutup, mulut menguap, dan ponsel, serta dikonversi ke dalam format yang kompatibel dengan arsitektur YOLOv8.



Gambar 2. Diagram Alir Penelitian

Untuk memperkaya keragaman data dan meningkatkan generalisasi model, dilakukan augmentasi terhadap dataset menggunakan teknik transformasi seperti flipping horizontal, rotasi, penyesuaian brightness, dan penambahan noise. Dataset yang telah diaugmentasi kemudian dibagi ke dalam tiga subset, yaitu data pelatihan, data validasi, dan data pengujian, masing-masing dengan proporsi 70%, 20%, dan 10%.

Pelatihan model dilakukan menggunakan pustaka Ultralytics YOLOv8 yang dijalankan pada platform Google Colaboratory dengan dukungan GPU Tesla T4. Konfigurasi pelatihan menggunakan parameter seperti ukuran input 640×640 piksel, batch size 16, learning rate 0.001, dan jumlah epoch sebanyak 60. Model dilatih ulang dari bobot awal (pretrained weights) yang telah tersedia agar mampu mengenali objek visual spesifik pada wajah pengemudi secara efisien.

Evaluasi performa model dilakukan menggunakan metrik umum dalam tugas deteksi objek, yaitu precision, recall, F1-score, dan mean Average Precision (mAP) pada threshold IoU 0.5 dan 0.5:0.95. Selain evaluasi kuantitatif, digunakan juga visualisasi seperti confusion matrix dan grafik hasil pelatihan untuk menganalisis kinerja model secara kualitatif terhadap data pengujian.

HASIL DAN PEMBAHASAN

Penelitian ini menghasilkan model deteksi visual berbasis YOLOv8 yang dievaluasi melalui tiga aspek utama, yaitu performa pelatihan, akurasi per kelas objek, serta analisis kesalahan deteksi. Pembahasan pada bab ini disusun untuk memberikan gambaran menyeluruh mengenai kinerja model dalam mendeteksi tanda visual kantuk dan distraksi pada pengemudi.

Hasil Pelatihan Model

Pelatihan model YOLOv8 dilakukan pada Google Colaboratory dengan memanfaatkan GPU Tesla T4. Model dilatih menggunakan konfigurasi berupa ukuran gambar 640 × 640 piksel, batch size sebesar 16, dan jumlah epoch sebanyak 60. Nilai learning rate diatur secara otomatis (default) oleh pustaka Ultralytics.

Distribusi data pelatihan, validasi, dan pengujian untuk masing-masing kelas objek dapat dilihat pada Tabel 1.

Tabel 1. Distribusi Dataset

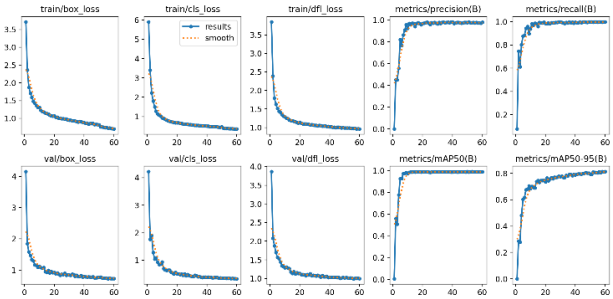
Kelas	Training	Validasi	Testing
Mata Kanan Terbuka	498	43	17
Mata Kanan Tertutup	423	38	22
Mata Kiri Terbuka	480	41	18
Mata Kiri Tertutup	423	38	21
Menguap	435	37	18
Ponsel	378	48	26
Total	2637	245	122

Dari tabel tersebut terlihat bahwa jumlah citra pada setiap kelas relatif seimbang, meskipun kelas ponsel memiliki data pelatihan yang sedikit lebih rendah. Distribusi ini diharapkan dapat membantu model dalam belajar mengenali variasi visual dari masing-masing objek secara efektif. Parameter pelatihan utama model YOLOv8 selama proses training dapat diringkas pada Tabel 2.

Tabel 2. Parameter Pelatihan YOLO

Parameter	Nilai
Model	YOLOv8s (small)
Ukuran gambar	640 x 640 px
Batch size	16
Epoch	60
Learning rate	Default (auto)

Dari evaluasi terhadap data uji, diperoleh performa umum model seperti tersaji pada Tabel 3. Nilai precision sebesar 87,40 persen dan recall sebesar 84,10 persen menunjukkan keseimbangan yang baik antara ketepatan prediksi dan kemampuan mendeteksi seluruh objek. Sementara itu, nilai mAP pada ambang batas IoU 0,5 sebesar 89,20 persen, serta mAP pada rentang IoU 0,5 hingga 0,95 sebesar 67,30 persen, memperlihatkan performa deteksi yang memadai pada berbagai tingkat ketelitian overlap.

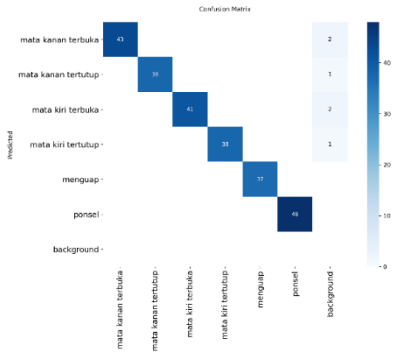


Gambar 3. Grafik Loss dan mAP Selama Proses Pelatihan Model Yolov8

Tabel 3. Hasil Evaluasi Model

Metrik	Nilai
Precision	87.40%
Recall	84.10%
F1-score	85.70%
mAP@0.5	89.20%
mAP@0.5:0.95	67.30%

Selain itu, untuk mengetahui pola klasifikasi silang antar kelas, dilakukan analisis confusion matrix yang ditampilkan pada Gambar 3 . Dari visualisasi tersebut dapat terlihat beberapa kesalahan prediksi antar kelas, meskipun secara umum distribusi prediksi benar mendominasi.



Gambar 4. Confusion Matrix Deteksi Objek Pada Data Uji

Evaluasi Deteksi per Kelas

Evaluasi lanjutan dilakukan pada masing-masing kelas objek untuk mengetahui sejauh mana model mampu mendeteksi tanda visual spesifik yang mengindikasikan kantuk maupun distraksi pengemudi. Tabel 4 merangkum hasil evaluasi performa per kelas.

Tabel 4. Evaluasi Performa per Kelas

Kelas	**Precision**	**Recall**	**mAP @0.5**	**mAP @0.5: 0.95**
Mata kanan terbuka	97.7%	99.1%	98.5%	76.3%
Mata kanan tertutup	97.1%	100%	98.8%	79.7%
Mata kiri terbuka	96.7%	100%	97.4 %	77.5%
Mata kiri tertutup	97.2%	100%	99.4 %	78.4%
Menguap	99.5%	100%	99.5 %	82.8%
Ponsel	99.6%	100%	99.5%	94.1%

Berdasarkan tabel tersebut, dapat diketahui bahwa performa deteksi terbaik dicapai pada kelas penggunaan ponsel dan ekspresi menguap, dengan nilai precision dan recall yang mendekati sempurna. Nilai mAP pada kelas ponsel juga tercatat paling tinggi, yakni sebesar 94,1 persen pada rentang IoU 0,5 hingga 0,95. Sementara itu, performa pada deteksi mata terbuka maupun tertutup juga tetap berada pada tingkat akurasi yang sangat baik, menunjukkan konsistensi model dalam mengenali berbagai kondisi visual wajah pengemudi.

Meskipun demikian, hasil evaluasi ini juga memperlihatkan adanya sejumlah kesalahan deteksi dalam bentuk false positive maupun false negative, yang umumnya terjadi pada kondisi ekstrem. Kesalahan semacam ini banyak dijumpai ketika pose wajah pengemudi tidak frontal, menoleh ke samping, sebagian wajah tertutup oleh tangan, rambut, atau aksesoris seperti kacamata tebal, serta pada citra dengan pencahayaan

yang rendah. Kondisi-kondisi tersebut menyulitkan model dalam mengenali batas objek secara tepat, sehingga dapat mengakibatkan objek target terlewat terdeteksi (missed detection) ataupun terdeteksi ganda atau keliru (misclassification). Temuan ini menjadi catatan penting untuk pengembangan lebih lanjut, misalnya melalui peningkatan variasi dataset dan penerapan teknik preprocessing guna memperkuat ketahanan model terhadap kondisi nyata.

KESIMPULAN

Penelitian ini telah berhasil mengevaluasi kinerja model deteksi objek berbasis YOLOv8 dalam mengidentifikasi empat jenis ekspresi visual yang berkaitan dengan kondisi kantuk dan distraksi pada pengemudi, yaitu mata terbuka, mata tertutup, ekspresi menguap, dan penggunaan ponsel. Berdasarkan hasil pelatihan dan pengujian, model menunjukkan performa yang cukup baik dengan nilai *mean Average Precision* (mAP) sebesar 89,2% pada skala IoU 0.5 dan 67,3% pada skala 0.5–0.95. Evaluasi per kelas juga menunjukkan tingkat presisi dan recall yang tinggi, khususnya pada deteksi mata terbuka dan penggunaan ponsel.

Temuan ini menunjukkan bahwa YOLOv8 mampu menjadi solusi efektif untuk mendeteksi indikator visual perilaku pengemudi secara real-time, dengan tingkat akurasi yang layak untuk diterapkan pada sistem monitoring berbasis visi komputer. Meskipun demikian, model masih memiliki keterbatasan dalam menghadapi kondisi ekstrem, seperti pose wajah tidak frontal atau pencahayaan yang kurang optimal, yang dapat memengaruhi akurasi deteksi.

Untuk pengembangan lebih lanjut, penelitian ini menyarankan integrasi model deteksi ini dengan sistem klasifikasi berbasis logika atau time series guna menganalisis durasi kantuk atau distraksi secara kontinu. Selain itu, implementasi sistem ke perangkat embedded seperti Raspberry Pi perlu dilakukan untuk menguji kinerja model dalam kondisi nyata di lingkungan kendaraan.

REFERENSI

[1] Firdaus, F. Utaminingrum, and E. R. Widasari, "Sistem Pendeteksi Kantuk Pengemudi berbasis Eye Aspect Ratio dan Mouth Opening Ratio menggunakan Algoritme C-LSTM," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 7, no. 2, pp. 927–933, Mar. 2023, Accessed: Jun. 28, 2025. [Online]. Available: https://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/12347

[2] Y. Zhou, L. Xia, J. Zhao, R. Yao, and B. Liu, "Efficient convolutional neural networks and network compression methods for object detection: a survey," *Multimed Tools Appl*, vol. 83, no. 4, pp. 10167–10209, Jan. 2023, doi: 10.1007/S11042-023-15608-2.

[3] Q. Zhang, "CNNA: A study of Convolutional Neural Networks with Attention," *Procedia Comput Sci*, vol. 188, pp. 26–32, Jan. 2021, doi: 10.1016/J.PROCS.2021.05.049.

[4] M. A. R. Alif and M. Hussain, "YOLOv1 to YOLOv10: A comprehensive review of YOLO variants and their application in the agricultural domain," Jun. 2024, [Online]. Available: http://arxiv.org/abs/2406.10139

[5] J. Terven, D. M. Córdova-Esparza, and J. A. Romero-González, "A Comprehensive Review of YOLO Architectures in Computer Vision: From YOLOv1 to YOLOv8 and YOLO-NAS," Dec. 01, 2023, *Multidisciplinary Digital Publishing Institute (MDPI)*. doi: 10.3390/make5040083.

[6] S. Liu, F. Shao, W. Chu, J. Dai, and H. Zhang, "An Improved YOLOv8-Based Lightweight Attention Mechanism for Cross-Scale Feature Fusion," *Remote Sensing 2025, Vol. 17, Page 1044*, vol. 17, no. 6, p. 1044, Mar. 2025, doi: 10.3390/RS17061044.

[7] M. Yaseen, "What is YOLOv8: An In-Depth Exploration of the Internal Features of the Next-Generation Object Detector," Aug. 2024, Accessed: Jun. 28, 2025. [Online]. Available: https://arxiv.org/pdf/2408.15857

[8] R. Padilla, W. L. Passos, T. L. B. Dias, S. L. Netto, and E. A. B. Da Silva, "A comparative analysis of object detection metrics with a companion open-source toolkit," *Electronics (Switzerland)*, vol. 10, no. 3, pp. 1–28, Feb. 2021, doi: 10.3390/electronics10030279.

[9] M. F. Yusri, P. Mangat, and O. Wasenmüller, "Detection of Driver Drowsiness by Calculating the Speed of Eye Blinking," Oct. 2021, Accessed: Jun. 29, 2025. [Online]. Available: https://arxiv.org/pdf/2110.11223

[10] R. Ezzati Amini *et al.*, "Driver distraction and in-vehicle interventions: A driving simulator study on visual attention and driving performance," *Accid Anal Prev*, vol. 191, p. 107195, Oct. 2023, doi: 10.1016/J.AAP.2023.107195.