

Performa Model YOLOv8 untuk Deteksi Kondisi Mengantuk pada pengendara mobil

Edmund Ucok Armin¹, Anggun Purnama Edra², Fakhri Ikhwanul Alifin³,
Ikhwanussafa Sadidan⁴, Indri Purwita Sary⁵, Ulinnuha Latifa⁶

^{1,5,6}Program Studi Teknik Elektro, Fakultas Teknik, Universitas Singaperbangsa Karawang, Indonesia

²Program studi sistem informasi, Fakultas sains dan teknologi, Universitas Terbuka, Indonesia

³Program Studi Teknik Industri, Fakultas Teknik, Universitas Singaperbangsa Karawang, Indonesia

⁴Program Studi Teknik Lingkungan, Fakultas Teknik, Universitas Singaperbangsa Karawang, Indonesia

E-mail : ¹edmund.ucok@ft.unsika.ac.id, ²044742615@ecampus.ut.ac.id,
³fakhri.ikhwanul@ft.unsika.ac.id, ⁴Ikhwanussafa.sadidan@ft.unsika.ac.id,
⁵indri@ft.unsika.ac.id, ⁶ulinnuha.latifa@ft.unsika.ac.id

Abstract

Driving while drowsy is identified as a significant risk factor in traffic accidents, yet awareness of this risk is often lower compared to other hazards. Phenomena such as microsleep have been shown to increase the risk of inattention and accidents on the road. This study proposes a novel approach utilizing Deep Learning, specifically YOLOv8, to detect and address the risk of driver drowsiness. To train the model, the researchers employed a secondary dataset consisting of 3708 images, partitioned into 80% for model training and 20% for validation. Multiple models were compared during the training process, and the results indicated that the YOLOv8 model outperformed previous models, achieving a recall value of 0.95261, precision of 0.94655, F1-SCORE of 0.9496, and mAP of 0.98055. This research contributes to the development of more effective drowsiness detection systems using Deep Learning approaches, with promising evaluation results.

Keywords: microsleep, Deep Learning, YOLOv8

Abstrak

Kondisi mengemudi dalam keadaan mengantuk diidentifikasi sebagai faktor risiko utama dalam kecelakaan lalu lintas, namun kesadaran terhadap risiko ini sering kali minim dibandingkan dengan risiko lainnya. Fenomena seperti microsleep telah terbukti meningkatkan risiko ketidakwasan dan kecelakaan di jalan. Penelitian ini mengusulkan suatu pendekatan baru dengan memanfaatkan metode Deep Learning, khususnya YOLOv8, untuk mendeteksi dan mengatasi risiko mengantuk pengemudi. Untuk melatih model, peneliti menggunakan dataset sekunder berisi 3708 citra, membaginya menjadi 80% untuk pelatihan dan 20% untuk validasi. Beberapa model dibandingkan selama proses pelatihan, dan hasilnya menunjukkan bahwa model YOLOv8 mengungguli model-model sebelumnya dengan nilai recall sebesar 0.95261, precision sebesar 0.94655, F1-SCORE sebesar 0.9496, dan mAP sebesar 0.98055. Penelitian ini memberikan kontribusi pada pengembangan sistem deteksi kantuk yang lebih efektif menggunakan pendekatan Deep Learning, dengan hasil evaluasi yang menjanjikan.

Kata Kunci: microsleep, Deep Learning, YOLOv8

1. Pendahuluan

Mengemudi dalam kondisi mengantuk merupakan faktor risiko serius dalam terjadinya kecelakaan lalu lintas. Sayangnya, kesadaran terhadap risiko ini sering kali kurang dibandingkan dengan risiko mengemudi sambil mengirim pesan teks atau dalam keadaan mabuk. Keadaan mengantuk ini melibatkan fenomena seperti microsleep yang dapat meningkatkan risiko ketidakwasan dan kecelakaan di jalan. Untuk mengatasi risiko ini, penelitian ini memanfaatkan teknologi pemantauan dan deteksi kantuk pengemudi [1]. Sejumlah penelitian telah dilakukan terkait sistem deteksi kantuk pada pengemudi mobil. Penggunaan sensor EEG [2], [3], ataupun accelerometer [4] yang ditempatkan di bagian kepala pengemudi membuat sistem ini tidak efisien karena perangkat harus dipasang pada tubuh untuk dapat berfungsi. Penggunaan sensor kamera menjadi solusi yang efektif karena tidak memerlukan kontak fisik dengan pengemudi. Metode-metode pengolahan citra terus berkembang untuk mendeteksi kantuk pada pengendara mobil dengan memperhatikan fitur citra dari mata dan mulut pengemudi untuk mendeteksi kantuk [5]–[7]. Seiring berjalannya waktu, sistem deteksi kantuk menggunakan deep learning telah banyak dilakukan. Penggunaan metode Convolutional Neural Network (CNN) menjadi populer karena memiliki akurasi yang tinggi dibandingkan dengan metode konvensional [8]–[12].

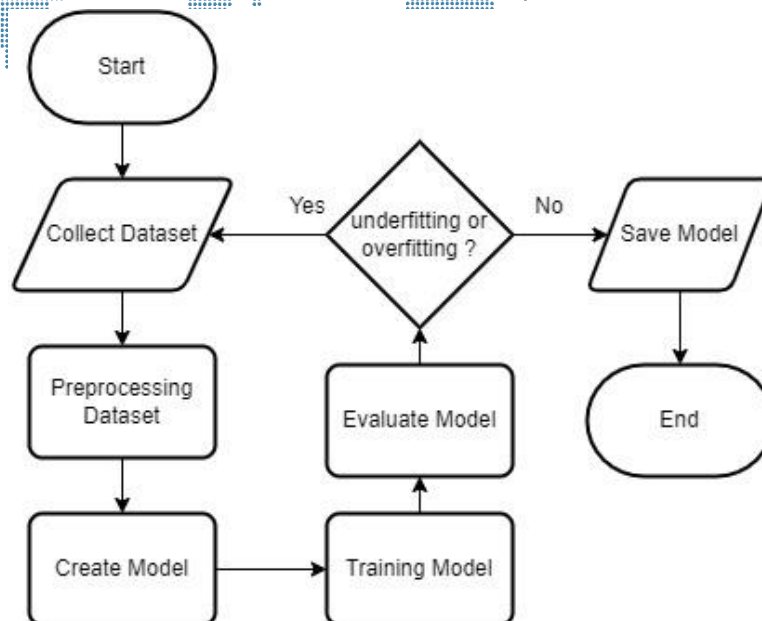
YOLO (You Only Look Once) adalah salah satu metode CNN untuk deteksi objek dan segmentasi gambar yang diperkenalkan oleh Joseph Redmon dan Ali Farhadi dari Universitas Washington pada tahun 2015. YOLOv8 [13], sebagai versi terkini yang dikembangkan oleh Ultralytics, menunjukkan kemajuan terbaru dalam model deteksi objek, memperkenalkan fitur-fitur baru dan peningkatan untuk meningkatkan kinerja dan fleksibilitas. YOLOv8 mendukung berbagai tugas dalam bidang kecerdasan buatan, seperti deteksi, segmentasi, estimasi pose, pelacakan, dan klasifikasi, memberikan fleksibilitas bagi pengguna dalam berbagai aplikasi dan domain. Ultralytics YOLOv8, sebagai model deteksi objek dan segmentasi gambar real-time terkini, mengintegrasikan kemajuan terbaru dalam pembelajaran mendalam dan visi komputer. Dalam percobaan ini, model YOLOv8 dan dataset dari ROBOFLOW digunakan, dengan melakukan pelatihan pada dataset menggunakan VSCode (Visual Studio Code). Eksekusi program dilakukan secara real-time menggunakan kamera. Hasil eksperimen menunjukkan kemampuan model untuk mendeteksi kondisi pengemudi secara real-time melalui pengenalan wajah.

Proses eksperimental ini menggunakan VSCode (Python), Visual Studio Code (VSCode), yang merupakan perangkat lunak pengedit kode sumber yang bersifat open source dan dikembangkan oleh Microsoft. Mampu beroperasi di berbagai sistem operasi seperti Linux, macOS, dan Windows, VSCode menawarkan fitur-fitur penting seperti penyorotan sintaksis, penyelesaian kode otomatis, manajemen kutipan kode, kemampuan refaktor kode, integrasi dengan pengawakutuan, dan dukungan untuk sistem kontrol versi seperti Git. VSCode sangat populer dalam pengembangan perangkat lunak dengan berbagai bahasa pemrograman, termasuk Python.

2. Metodologi Penelitian

Penelitian ini dimulai dengan pengumpulan dataset yang mencakup pengendara mobil dalam kondisi mengantuk. Setelah berhasil mengakuisisi dataset, analisis dataset dilakukan, dengan penerapan metode-metode augmentasi untuk meningkatkan jumlah data. Peneliti selanjutnya mengimplementasikan model deteksi objek ke dalam bahasa pemrograman yang kemudian menjalani proses pelatihan menggunakan dataset yang telah dikumpulkan. Setelah melalui tahap pelatihan model, evaluasi dilakukan untuk memeriksa apakah model menunjukkan indikasi overfitting atau underfitting. Jika model menunjukkan karakteristik tersebut, peneliti kembali menganalisis dataset dan melakukan penyesuaian terhadap model. **Overfitting terjadi ketika model memberikan hasil baik selama pelatihan, namun memiliki performa buruk pada tahap validasi, sementara**

underfitting terjadi jika model tidak mengalami peningkatan akurasi selama pelatihan. Jika model tidak menunjukkan indikasi overfitting atau underfitting, model disimpan untuk penggunaan lebih lanjut. Alur penelitian ini dapat dilihat pada flowchart yang tergambar pada Gambar 1.



Gambar 1. Alur Penelitian

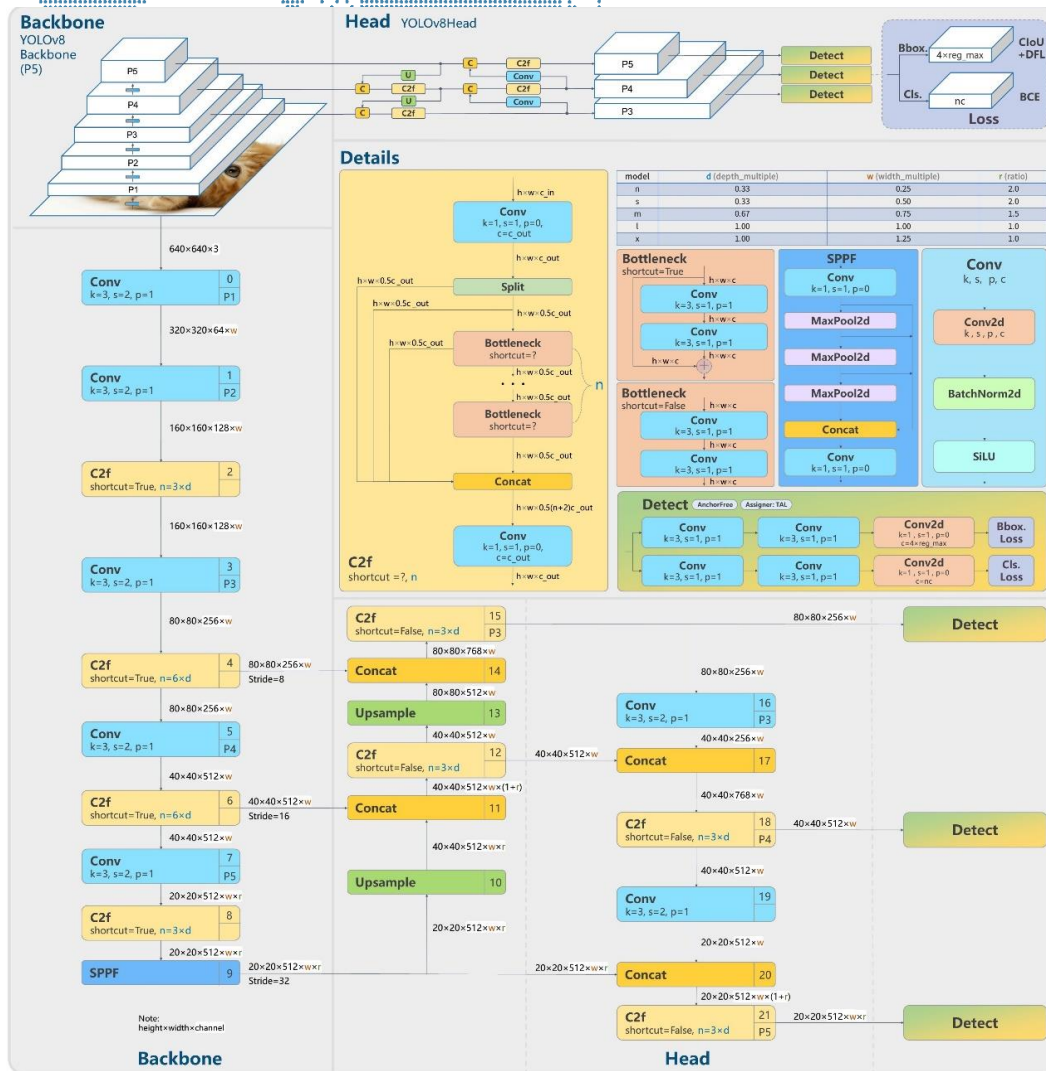
2.1. YOLOv8

Arsitektur YOLOv8 dari Ultralytics didasarkan pada konsep deteksi objek dalam satu tahap (one-stage), yang memungkinkan penggunaan yang efisien dan akurat dalam menangkap objek pada gambar. Arsitektur ini mengintegrasikan sejumlah inovasi dan peningkatan dari versi-versi YOLO sebelumnya.

Pertama-tama, YOLOv8 menggunakan sebuah jaringan saraf utama (backbone) yang kuat untuk mengekstraksi fitur dari gambar. Dalam konteks ini, Ultralytics telah memilih arsitektur CSPDarknet53 sebagai jaringan saraf utama, yang dikenal karena kemampuannya dalam menangkap representasi gambar yang dalam.

Setelah proses ekstraksi fitur, YOLOv8 menggunakan lapisan deteksi untuk memprediksi bounding box dan kelas objek. Arsitektur ini mempartisi gambar ke dalam kisi dan untuk setiap sel kisi, model memprediksi beberapa bounding box bersamaan dengan probabilitas kelas. Penggunaan deteksi multi-skala memungkinkan model untuk menangkap objek dengan berbagai skala, sehingga tidak terbatas pada objek dengan ukuran tertentu.

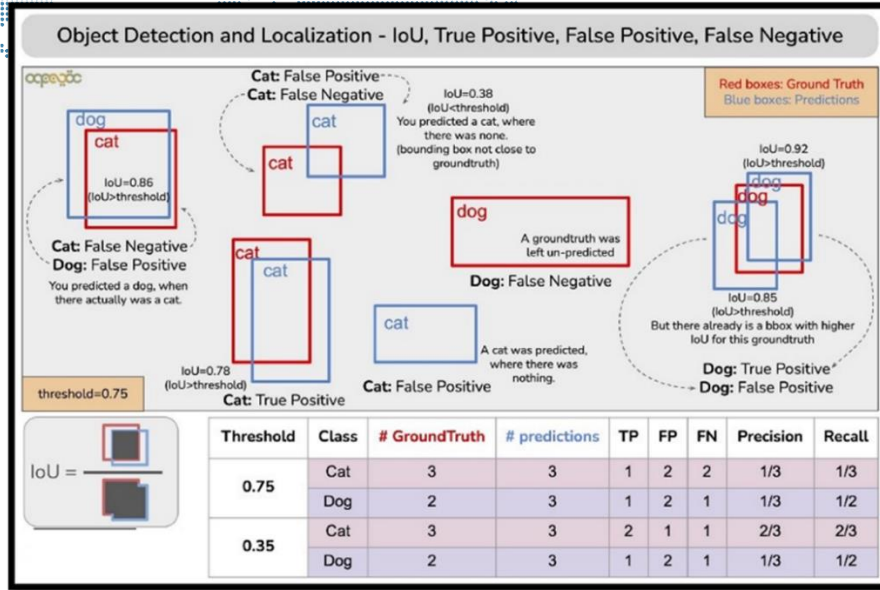
Selanjutnya, YOLOv8 mengimplementasikan teknik koneksi loncat (skip connections), yang memfasilitasi integrasi informasi yang lebih rinci dari lapisan-lapisan awal ke dalam lapisan-lapisan yang lebih dalam. Ini berkontribusi pada peningkatan kemampuan model dalam mengenali objek, terutama dalam kasus objek yang kecil atau kompleks. Dengan demikian, arsitektur YOLOv8 mencapai keseimbangan optimal antara kecepatan dan akurasi, menjadikannya solusi yang sangat efektif untuk tugas deteksi objek yang kompleks. Arsitektur YOLOv8 dapat dilihat pada Gambar 2.



Gambar 2. YOLOv8 Architecture

2.2. Evaluasi Model

Performa model YOLOv8 dievaluasi melalui *confusion matrix* yang diperoleh selama proses pelatihan. *Confusion matrix* ini memungkinkan perhitungan metrik seperti akurasi, presisi, recall, dan F1-skor. Akurasi mengukur persentase data yang berhasil dideteksi dengan benar dari total data pengujian. Presisi mencerminkan seberapa tepat model mengidentifikasi anggota suatu kelas dibandingkan dengan total data yang diklasifikasikan sebagai kelas tersebut. Recall mengukur sejauh mana model dapat mengidentifikasi anggota kelas tersebut dibandingkan dengan total data pengujian yang sebenarnya. F1-skor adalah metrik yang menggabungkan presisi dan recall menjadi rata-rata harmonik, memberikan gambaran holistik tentang performa model. Semakin tinggi nilai akurasi, presisi, recall, dan F1-skor dari model, semakin baik kinerja model tersebut dalam tugas deteksi objek menggunakan YOLOv8.



Gambar 3. Confusion Matrix for Object Detection

Terdapat 3 istilah sebagai representasi hasil proses deteksi objek pada *confusion matrix*, diantaranya :

1. *True Positive* (TP): Model memprediksi dengan benar bahwa objek adalah positif.
2. *True Negative* (FP): Model memprediksi dengan benar bahwa sesuatu adalah negatif.
3. *False Positive* (FN): Model salah memprediksi objek sebagai positif ketika seharusnya negatif.

Penentuan jumlah TP, FP dan FN dapat dilihat pada gambar 3. Nilai dari confusion matrix digunakan untuk mengukur performa dari model objek deteksi. Beberapa metode dapat digunakan untuk melihat performa dari model deteksi objek yaitu:

- 1) Precision merupakan sejauh mana prediksi positif model benar.

$$\text{Precision} = \frac{TP}{TP + FP} \quad (1)$$

- 2) Recall merupakan kemampuan model untuk mengidentifikasi semua data positif yang sebenarnya.

$$\text{Recall} = \frac{TP}{TP + FN} \quad (2)$$

- 3) F1 Score merupakan keseimbangan antara precision dan recall dalam satu nilai, yang membantu mengevaluasi kinerja model.

$$F1_{\text{SCORE}} = \frac{2 \times \text{recall} \times \text{precision}}{\text{recall} + \text{precision}} \quad (3)$$

- 4) Mean Average Precicion (mAP) merupakan presisi rata-rata yang sangat sering digunakan dalam mengukur keakuratan model deteksi.

$$mAP = \frac{1}{N} \sum_{i=1}^N \text{Average Precision}_i \quad (4)$$

3. Hasil dan Pembahasan

Studi ini menggunakan sebuah komputer pribadi dengan spesifikasi prosesor Core i5 Generasi 9, RAM 8 GB, dan Nvidia RTX2060 sebagai unit pemrosesan data. Peneliti memanfaatkan bahasa pemrograman Python dengan library PyTorch untuk menerapkan

model YOLOv8. Proses penelitian terdiri dari tiga tahap, yaitu persiapan dataset, pelatihan model YOLOv8, dan evaluasi kinerja model YOLOv8.

3.1. Dataset

Penelitian ini memanfaatkan dataset sekunder yang diperoleh dari sumber Roboflow dengan nama *dowsy_detection Dataset* [1]. Dataset ini terdiri dari 3708 citra dan telah terbagi menjadi dua bagian, yaitu 80% digunakan untuk melatih model dan 20% untuk mengevaluasi kinerja model. Pembagian dataset ini menjadi dua bagian dilakukan agar peneliti dapat menguji kemampuan model dalam mendeteksi objek pada citra yang tidak pernah dilihat oleh model sebelumnya. Rincian distribusi dataset dapat ditemukan dalam Tabel 1 dan Gambar 4.

Table 1. Distribusi Dataset

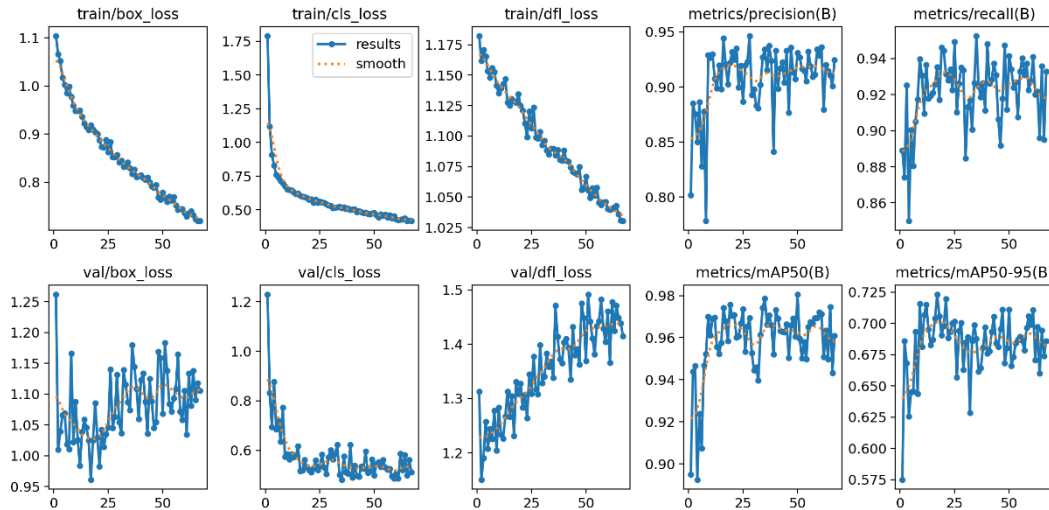
<i>Classes ID</i>	<i>Classes Name</i>	<i>Training</i>	<i>Validation</i>	<i>All</i>
0	<i>Awake</i>	1090	322	1412
1	<i>Drowsy</i>	1906	390	2296
<i>Total</i>		2996	712	3708



Gambar 4. Dataset

3.2. Training Model

Dalam tahap pelatihan model, peneliti memanfaatkan dataset pelatihan untuk melatih model YOLOv8 guna mendeteksi kelelahan pada pengemudi mobil. Proses pelatihan dilakukan dengan mengonfigurasi jumlah epoch sebanyak 100 dan ukuran batch sebesar 8. Saturasi pelatihan tercapai pada epoch ke-67 dengan mencapai nilai recall sebesar 0.93287, presisi sebesar 0.92442, mAP sebesar 0.96092, box loss sebesar 0.71979, dan classes loss sebesar 0.4189. Kurva hasil pelatihan dapat ditemukan pada Gambar 5.



Gambar 5. Training Result

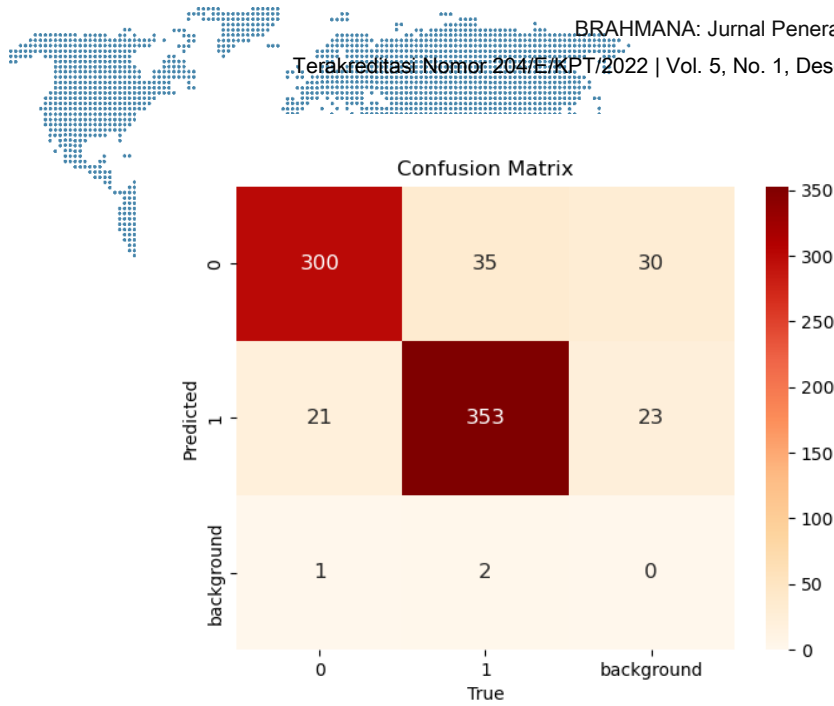
3.3. Evaluasi Model



Gambar 6. Implementation

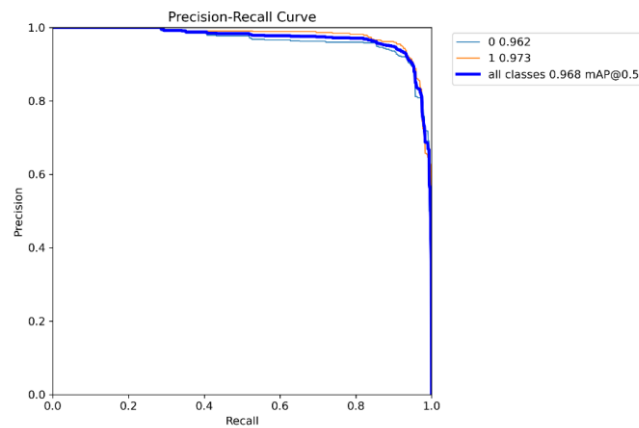
Setelah menyelesaikan pelatihan pada model, peneliti kemudian melanjutkan ke tahap evaluasi menggunakan dataset validasi, terdiri dari 322 citra untuk kelas 0 dan 390 citra untuk kelas 1. Implementasi YOLOv8 dalam mendeteksi citra terlihat pada Gambar 6. YOLOv8 berhasil mengidentifikasi objek secara tepat pada 300 citra untuk kelas 0 dan 353 citra untuk kelas 1, yang menghasilkan nilai True Positive (TP) sebanyak 653 pada dataset validasi.

Namun, model YOLOv8 juga mengalami kesalahan deteksi, dengan 35 citra kelas 0 yang keliru diprediksi sebagai kelas 1, dan 30 citra yang seharusnya merupakan latar belakang. Untuk kelas 1, terdapat 21 citra yang salah diprediksi sebagai kelas 0 dan 23 citra yang seharusnya merupakan latar belakang. Hal ini menyebabkan nilai False Negative (FN) sebanyak 109 dan nilai False Positive (FP) sebanyak 59. Rincian evaluasi model YOLOv8 dapat ditemukan pada *Confusion Matrix* yang tergambar pada Gambar 7.



Gambar 7. Confusion Matrix

Setelah memperoleh Confusion Matrix, peneliti melakukan perhitungan terhadap nilai precision, recall, F1-Score, dan mAP untuk mengevaluasi kinerja model yang telah dilatih. Gambar 8 menunjukkan kurva presisi terhadap recall sebagai representasi visual dari evaluasi tersebut. Selain itu, peneliti melakukan perbandingan dengan model deteksi objek lainnya, seperti SSD-Resnet, YOLOv3, dan YOLOv5, dengan menggunakan dataset yang sama untuk proses pelatihan dan evaluasi model. Dari Tabel 2, dapat diamati bahwa model YOLOv8 mencapai nilai evaluasi yang lebih unggul dalam mendeteksi kelelahan pada pengemudi menggunakan dataset yang digunakan oleh peneliti.



Gambar 8. Precicion Recall Curve

Table 2. Model Comparison

Model	Recall	Precicion	F1-SCORE	mAP	FPS (Frame per Seconds)
SSD-Resnet	0.72253	0.83162	0.72253	0.8201	132
YOLOv3	0.8112	0.8312	0.821	0.85	75
YOLOv5	0.91	0.9621	0.9353	0.953	80
YOLOv8	0.95261	0.94655	0.9496	0.98055	82

4. Kesimpulan

Dari hasil penelitian ini, model YOLOv8, berhasil mencapai nilai mAP sebesar 0.96092 selama proses pelatihan dengan memanfaatkan 2996 citra dari dataset sekunder. Peneliti selanjutnya melakukan evaluasi pada kinerja model YOLOv8 dan membandingkannya dengan model deteksi objek lainnya, termasuk SSD-Resnet, YOLOv3, dan YOLOv5, dengan menggunakan dataset yang sama yang telah dikumpulkan oleh peneliti. Hasil evaluasi menunjukkan bahwa meskipun YOLOv8 memiliki FPS yang lebih rendah dibandingkan dengan SSD-Resnet dengan selisih sebanyak 50 FPS, namun model ini menunjukkan tingkat ketepatan yang lebih baik dalam mendeteksi objek.

Daftar Pustaka

- [1] S. Dong, T. Zhang, H. Zhu, K. Li, and A. Wang, "Research on the effect of shift work system on rail transit drivers' sleep and fatigue," in *Proceedings - 2020 13th International Symposium on Computational Intelligence and Design, ISCID 2020*, Institute of Electrical and Electronics Engineers Inc., Dec. 2020, pp. 270–273. doi: 10.1109/ISCID51228.2020.00066.
- [2] M. S. Saleab, M. A. Abd, E. Ghany, R. M. Toma, and K. Hofmann, "Real-Time Sleep Detection and Warning System to Ensure Driver's Safety Based on EEG."
- [3] IEEE Staff, *2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. IEEE, 2017.
- [4] Y. Siregar and F. K. Sani, "Accelerometer-Based Head Movement Detection System Design as a Sleep Detection Tool," in *Proceedings - 2022 9th International Conference on Information Technology, Computer and Electrical Engineering, ICITACEE 2022*, Institute of Electrical and Electronics Engineers Inc., 2022, pp. 86–89. doi: 10.1109/ICITACEE55701.2022.9924097.
- [5] SCAD College of Engineering and Technology and Institute of Electrical and Electronics Engineers, *Proceedings of the International Conference on Trends in Electronics and Informatics (ICOEI 2019) : 23-25, April 2019*.
- [6] Universitatea de Medicină și Farmacie "Gr.T. Popa" Iași, IEEE Engineering in Medicine and Biology Society. Romania Chapter, Academia Română (1990-). Institute of Computer Science. Iași Branch, and Institute of Electrical and Electronics Engineers, *2020 E-Health and Bioengineering Conference (EHB) : EHB 2020 - 8-th edition : Web Conference, Iasi, Romania, October 29-30, 2020*.
- [7] P. Sudarshan, V. Bhardwaj, and Virender, "Real-time Driver Drowsiness Detection and Assistance System using Machine Learning and IoT," in *Proceedings of the 8th International Conference on Communication and Electronics Systems, ICCES 2023*, Institute of Electrical and Electronics Engineers Inc., 2023, pp. 1128–1132. doi: 10.1109/ICCES57224.2023.10192831.
- [8] N. Ali, I. Hasan, T. Özyer, and R. Alhajj, "Driver drowsiness detection by employing CNN and DLIB," in *2021 22nd International Arab Conference on Information Technology, ACIT 2021*, Institute of Electrical and Electronics Engineers Inc., 2021. doi: 10.1109/ACIT53391.2021.9677197.
- [9] "Neural Network Based Driver Warning System."
- [10] M. V. Sowmya Laxshmi, U. Prabu, L. Chandana, and N. Sunny, "An Enhanced Driver Drowsiness Detection System using Transfer Learning," in *Proceedings of the 5th International Conference on Electronics, Communication and Aerospace Technology, ICECA 2021*, Institute of Electrical and Electronics Engineers Inc., 2021, pp. 1671–1678. doi: 10.1109/ICECA52323.2021.9676050.
- [11] SCAD Institute of Technology and Institute of Electrical and Electronics Engineers, *Proceedings of the 3rd International Conference on IoT in Social, Mobile, Analytics and Cloud (ISMAC 2019) : 12-14 December, 2019*.

- [12] Fla.) IEEE Engineering in Medicine and Biology Society. Annual International Conference (38th : 2016 : Orlando, IEEE Engineering in Medicine and Biology Society, and Institute of Electrical and Electronics Engineers, *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) : 16-20 Aug. 2016*.
- [13] I. Purwita Sary, E. Ucok Armin, S. Andromeda, E. Engineering, and U. Singaperbangsa Karawang, "Performance Comparison of YOLOv5 and YOLOv8 Architectures in Human Detection Using Aerial Images," *Ultima Computing : Jurnal Sistem Komputer*, vol. 15, no. 1, 2023.