

Deteksi Kantuk Mahasiswa pada Pembelajaran *Online* Menggunakan YOLOv8 dengan Pendekatan CRISP-DM pada *Dataset Roboflow*

Dennis Ma'rifatul Kurnia¹, Arie Nugroho², Anita Sari Wardani³

^{1,2,3}Program Studi Sistem Informasi, Fakultas Teknik dan Ilmu Komputer, Universitas Nusantara PGRI Kediri

¹denniskurnia33@gmail.com*, ²arienugroho@unpkediri.ac.id, ³anita@unpkediri.ac.id

Abstract

The quality of online learning in higher education is highly influenced by the level of student alertness and concentration during the lecture process. However, the phenomenon of drowsiness resulting from physical fatigue and prolonged screen time often reduces learning effectiveness. The limitations of lecturers in visually monitoring students' conditions on online learning platforms drive the need for an automated monitoring system capable of operating in real-time. This study aims to develop a student drowsiness detection system using the YOLOv8 algorithm with a CRISP-DM approach. The dataset used originates from two public datasets on the Roboflow platform, namely SleepyDetect and Yawn, which then underwent curation, labeling, and augmentation processes to produce 8,573 images across three classes: open-eyes, pre-drowsy (yawning), and drowsy (eyes closed). The evaluation results show that the model achieves an mAP50 value of 0.992, Precision of 0.973, Recall of 0.979, and an F1-Score of 0.976, with an inference speed of 4.1 ms per image, demonstrating high detection performance on the utilized dataset. The model's implementation was carried out through the AwakeLens software integrated with a monitoring dashboard to support the online learning process. Nevertheless, because this research relies on a public dataset, the obtained results still require further validation using a primary dataset collected from Indonesian students to enhance the external validity and generalization capability of the model.

Keywords: drowsiness detection, student drowsiness, online learning, computer vision, YOLOv8

Abstrak

Kualitas pembelajaran daring di perguruan tinggi sangat dipengaruhi oleh tingkat kewaspadaan dan konsentrasi mahasiswa selama proses perkuliahan. Namun, fenomena kantuk akibat kelelahan fisik dan durasi penggunaan layar yang panjang sering menurunkan efektivitas pembelajaran. Keterbatasan dosen dalam memantau kondisi mahasiswa secara visual pada platform pembelajaran daring mendorong perlunya sistem pemantauan otomatis yang mampu bekerja secara *real-time*. Penelitian ini bertujuan mengembangkan sistem deteksi kantuk mahasiswa menggunakan algoritma YOLOv8 dengan pendekatan CRISP-DM. *Dataset* yang digunakan berasal dari dua dataset publik pada platform Roboflow, yaitu SleepyDetect dan Yawn, yang kemudian melalui proses kurasi, pelabelan, dan augmentasi sehingga menghasilkan 8.573 citra dengan tiga kelas, yaitu *open-eyes*, *pre-drowsy* (menguap), dan *drowsy* (mata tertutup). Hasil evaluasi menunjukkan bahwa model memperoleh nilai mAP50 sebesar 0,992, *Precision* 0,973, *Recall* 0,979, dan *F1-Score* 0,976, dengan kecepatan inferensi 4,1 ms per citra, sehingga menunjukkan performa deteksi yang tinggi pada dataset yang digunakan. Implementasi model dilakukan melalui perangkat lunak AwakeLens yang terintegrasi dengan dashboard pemantauan untuk mendukung proses pembelajaran daring. Meskipun demikian, karena penelitian ini menggunakan dataset publik, hasil yang diperoleh masih memerlukan validasi lebih lanjut menggunakan dataset primer yang dikumpulkan dari mahasiswa Indonesia agar validitas eksternal dan kemampuan generalisasi model dapat ditingkatkan.

Kata kunci: deteksi kantuk, pembelajaran daring, YOLOv8, CRISP-DM, Roboflow

©This work is licensed under a Creative Commons Attribution -ShareAlike 4.0 International License

1. Pendahuluan

Pendidikan tinggi merupakan pilar utama dalam pengembangan sumber daya manusia yang berkualitas di Indonesia. Dalam prosesnya, konsentrasi dan kesiagaan mahasiswa menjadi faktor krusial yang menentukan efektivitas transfer pengetahuan. Namun, fenomena kantuk di ruang kelas, khususnya dalam lingkungan pembelajaran daring, sering kali menjadi hambatan utama yang mengganggu optimalisasi pembelajaran. Dalam bidang *computer vision*, sistem deteksi berbasis *deep learning* telah dikembangkan secara luas untuk mengenali gejala penurunan atensi dan kantuk melalui analisis fitur visual wajah secara otomatis dan terukur [1]. Implementasi rekayasa informasi visual secara *real-time* ini terbukti efektif

menghasilkan performa inferensi yang responsif dalam mengenali kondisi kantuk pada ekspresi wajah secara instan [2]. Lebih lanjut, dalam konteks pendidikan digital, otomatisasi pemantauan ekspresi wajah mahasiswa menggunakan jaringan saraf tiruan mendalam menjadi solusi krusial untuk mengevaluasi tingkat keterikatan (*engagement*) serta kesiagaan belajar secara objektif pada ekosistem tatap maya [3]. Situasi tersebut menunjukkan perlunya sistem pemantauan otomatis berbasis *computer vision* yang mampu membantu dosen dalam mengidentifikasi kondisi kantuk mahasiswa secara *real-time*.

Seiring berkembangnya teknologi *Artificial Intelligence* (AI), berbagai penelitian mulai berfokus pada pengembangan sistem deteksi objek yang cepat,

akurat, dan efisien dalam penggunaan komputasi. Dalam konteks pembelajaran daring, sistem pemantauan kantuk dituntut mampu bekerja secara *real-time* tanpa memberikan beban komputasi yang tinggi pada perangkat yang digunakan oleh dosen maupun mahasiswa. Salah satu algoritma *deep learning* yang banyak digunakan untuk kebutuhan deteksi objek *real-time* adalah YOLOv8 (*You Only Look Once versi 8*), yaitu arsitektur deteksi objek berbasis *single-stage detector* yang mampu melakukan proses pelokalan (*localization*) dan klasifikasi objek secara simultan dalam satu tahap pemrosesan sehingga menghasilkan kecepatan deteksi yang tinggi [4]. Pendekatan tersebut membuat YOLOv8 memiliki kecepatan inferensi yang tinggi dengan tetap mempertahankan akurasi deteksi yang baik. Selain itu, variasi model ringan seperti YOLOv8n dan YOLOv8s dirancang untuk kebutuhan komputasi rendah sehingga mampu menghasilkan *Frame Per Second* (FPS) yang tinggi dan cocok diterapkan pada sistem berbasis video *streaming* maupun *platform* konferensi video dengan spesifikasi perangkat yang terbatas [5].

Beberapa penelitian terdahulu telah mengembangkan sistem deteksi kantuk berbasis *computer vision* menggunakan berbagai pendekatan. Penelitian [6] menerapkan algoritma YOLOv8 untuk mendeteksi kondisi *microsleep* pada pengemudi kendaraan roda empat berdasarkan citra mata. Model yang dikembangkan memperoleh nilai *Precision* sebesar 97,2%, *Recall* 96,7%, dan *mAP@0.5* sebesar 97,9%, serta mampu mempertahankan performa yang stabil pada berbagai kondisi pencahayaan. Hasil tersebut menunjukkan bahwa YOLOv8 efektif dalam mendeteksi perubahan kondisi mata sebagai indikator kantuk. Namun, penelitian tersebut hanya memanfaatkan area mata sehingga informasi ekspresi wajah lainnya, seperti aktivitas mulut, belum dimanfaatkan sebagai indikator tambahan.

Penelitian oleh [7] memperluas pendekatan deteksi dengan menggunakan citra wajah pengemudi secara utuh. Hasil penelitian menunjukkan bahwa YOLOv8 memperoleh nilai *mAP* sebesar 0,98055, *Precision* 94,65%, dan *Recall* 95,26%, serta memberikan performa yang lebih baik dibandingkan SSD-ResNet, YOLOv3, dan YOLOv5. Meskipun demikian, penelitian ini masih berfokus pada domain transportasi dengan karakteristik dataset yang berasal dari lingkungan berkendara, sehingga penerapannya pada konteks pembelajaran daring memerlukan penyesuaian terhadap kondisi lingkungan, sudut pengambilan gambar, dan perilaku pengguna yang berbeda.

Sementara itu, penelitian [8] mengembangkan sistem deteksi kantuk menggunakan kombinasi aktivitas mata, aktivitas mulut, dan kemiringan kepala melalui metode *Eye Aspect Ratio* (EAR) dan *pose estimation*. Pendekatan tersebut menunjukkan bahwa pemanfaatan beberapa indikator wajah mampu meningkatkan akurasi deteksi dibandingkan hanya menggunakan satu indikator. Namun, metode berbasis *facial landmark*

seperti EAR masih bergantung pada proses ekstraksi titik-titik wajah secara presisi sehingga kinerjanya dapat dipengaruhi oleh perubahan posisi wajah, kualitas citra, maupun kondisi pencahayaan.

Berdasarkan ketiga penelitian tersebut, dapat disimpulkan bahwa pemilihan dataset, fitur wajah, dan metode deteksi berpengaruh terhadap performa sistem deteksi kantuk. Sebagian besar penelitian masih menggunakan dataset pada sektor transportasi dan berfokus pada pengemudi kendaraan, sehingga belum merepresentasikan karakteristik pembelajaran daring. Selain itu, mayoritas penelitian hanya memanfaatkan satu indikator wajah atau hanya membedakan kondisi terjaga dan mengantuk tanpa mempertimbangkan fase awal kantuk (*pre-drowsy*) yang ditandai dengan aktivitas menguap. Oleh karena itu, penelitian ini mengembangkan sistem deteksi kantuk mahasiswa menggunakan YOLOv8 pada lingkungan pembelajaran daring dengan mengintegrasikan indikator mata dan aktivitas menguap ke dalam tiga kelas, yaitu *open-eyes*, *pre-drowsy*, dan *drowsy*, sehingga diharapkan mampu mendukung deteksi dini kondisi kantuk secara lebih adaptif.

Berdasarkan evaluasi studi terdahulu, terdapat kesenjangan penelitian (*research gap*) yang signifikan, di mana implementasi deteksi kantuk berbasis *deep learning* masih tersentralisasi pada domain transportasi dan minim adaptasi pada karakteristik ekosistem pendidikan digital. Selain itu, mayoritas sistem mengabaikan tahapan transisi penurunan fokus dengan hanya mengandalkan klasifikasi biner (terjaga atau tertidur absolut). Untuk mengatasi celah tersebut, penelitian ini mengintegrasikan varian model YOLOv8 Nano untuk mendeteksi tingkat kewaspadaan mahasiswa secara *real-time* melalui integrasi fitur orbital mata dan perioral mulut. Keterbaruan (*novelty*) yang ditawarkan adalah penyusunan kelas *pre-drowsy* (gejala menguap) sebagai instrumen intervensi peringatan dini. Landasan ilmiah penambahan kelas transisi ini didukung oleh penelitian komputasi perilaku yang membuktikan bahwa pemantauan dinamis pada area mulut memberikan sinyal preventif yang valid sebelum subjek memasuki fase *microsleep* [9]. Melalui pendekatan terintegrasi ini, sistem diharapkan mampu memberikan pemantauan yang lebih adaptif, cepat, dan ringan untuk meningkatkan efektivitas pengawasan instruksional pada pembelajaran daring.

2. Metode Penelitian

Penelitian ini mengembangkan sistem deteksi kantuk mahasiswa berbasis *computer vision* untuk menganalisis citra wajah secara *real-time* dalam pembelajaran daring. Menggunakan algoritma YOLOv8, sistem dirancang untuk mengklasifikasikan kondisi kewaspadaan menjadi tiga kelas: *awake*, *pre-drowsy*, dan *drowsy* melalui deteksi aktivitas mata dan mulut. Pengembangan sistem ini menerapkan kerangka kerja CRISP-DM (*Cross Industry Standard Process for*

Data Mining) yang terstruktur, meliputi enam tahapan utama: *Business Understanding*, *Data Understanding*, *Data Preparation*, *Modeling*, *Evaluation*, hingga *Deployment* berupa implementasi sistem berbasis antarmuka pengguna. Alur penelitian yang digunakan pada sistem ini ditunjukkan pada Gambar 1.



Gambar 1. Alur Metode CRISP-DM

2.1 Business Understanding

Pada tahap pertama CRISP-DM yaitu *Business Understanding*, dilakukan pemahaman terhadap permasalahan dan tujuan penelitian yang akan diselesaikan menggunakan teknik *data mining* [10]. Pada tahap ini, penelitian difokuskan pada kebutuhan pengembangan sistem deteksi kantuk mahasiswa pada pembelajaran daring. Dalam proses kelas *online*, dosen mengalami kesulitan memantau kondisi mahasiswa secara langsung sehingga gejala kantuk seperti mata terpejam dan aktivitas menguap sering tidak terdeteksi. Kondisi tersebut dapat memengaruhi konsentrasi dan efektivitas proses pembelajaran [11]. Oleh karena itu, diperlukan sistem otomatis yang mampu mendeteksi kondisi kewaspadaan mahasiswa secara cepat dan *real-time* menggunakan teknologi *computer vision*. Penelitian ini menggunakan algoritma YOLOv8 untuk mengklasifikasikan kondisi mahasiswa ke dalam tiga kelas, yaitu *awake*, *pre-drowsy*, dan *drowsy* melalui analisis aktivitas mata dan mulut. Keberhasilan model diukur berdasarkan nilai *accuracy*, *precision*, *recall*, dan *F1-score* pada berbagai kondisi pengujian.

2.2 Data Understanding

Pada tahap *Data Understanding*, dilakukan proses pengumpulan dan analisis data untuk memahami karakteristik dataset yang digunakan dalam penelitian [12]. Dataset diperoleh dari platform *Roboflow* dengan judul “*SleepyDetect*” dan “*Yawn*” yang digunakan untuk mendukung proses deteksi kondisi kantuk berbasis *computer vision*. Total dataset yang digunakan sebanyak 4.486 citra yang terdiri dari 1.962 citra mata terbuka (*open-eyes*), 889 citra menguap (*pre-drowsy*), dan 1.635 citra mata tertutup (*drowsy*).

Dataset memiliki variasi kondisi pencahayaan, sudut wajah, dan jarak pengambilan gambar yang berbeda untuk meningkatkan kemampuan generalisasi model

pada kondisi pembelajaran daring yang beragam. Selain itu, dilakukan analisis kualitas citra seperti ketajaman dan kejelasan area mata serta mulut agar fitur visual dapat dikenali dengan baik oleh algoritma YOLOv8. Penggunaan dataset yang beragam diharapkan mampu membantu model menghasilkan deteksi yang lebih stabil dan akurat pada proses pengujian *real-time*.

2.3 Data Preparation

Tahap *Data Preparation* dilakukan untuk menyiapkan dataset sebelum proses pelatihan model melalui beberapa langkah sistematis [13]. Langkah pertama adalah *data cleaning* untuk menghapus citra berkualitas buruk, seperti *blur*, terlalu gelap, atau tidak menampilkan area mata dan mulut dengan jelas. Selanjutnya, dilakukan pelabelan citra ke dalam tiga kelas (*0_open-eyes*, *1_pre-drowsy*, dan *2_drowsy*). Untuk menyeragamkan format data, diterapkan *preprocessing* yang meliputi *cropping* area wajah, penyesuaian ukuran citra, dan normalisasi piksel.

Selain itu, dilakukan ekspansi data melalui penerapan lima teknik augmentasi sistematis, yang meliputi *horizontal flip*, *crop*, rotasi (*rotation*), manipulasi kecerahan (*brightness*), serta pemburaman (*blur*) guna memperkaya variasi spasial data sekaligus mencegah risiko *overfitting*. Terakhir, dataset dibagi secara acak menjadi tiga bagian dengan proporsi 70:20:10. Alokasi 70% sebagai data latih (*training data*) digunakan untuk memperbaiki bobot arsitektur YOLOv8 dalam mengenali pola fitur wajah. Alokasi 20% sebagai data validasi (*validation data*) berfungsi untuk menyetel *hyperparameter* serta memantau risiko *overfitting* selama proses pelatihan. Sementara itu, 10% sisanya dialokasikan sebagai data uji (*testing data*) yang sepenuhnya belum pernah dilihat oleh model, guna mengevaluasi performa akhir dan kemampuan generalisasi sistem secara objektif dan optimal.

2.4 Modeling

Tahap Modeling dilakukan untuk membangun model deteksi kantuk menggunakan algoritma YOLOv8 berbasis *deep learning* [9]. YOLOv8 dipilih karena menggunakan pendekatan *single-stage detector* yang mampu mendeteksi objek secara cepat dan simultan. Karakteristik komputasi yang cepat ini membuatnya sangat ideal untuk kebutuhan sistem pemantauan secara *real-time*. Melalui algoritma ini, fitur visual mata dan mulut diekstraksi untuk diklasifikasikan ke dalam kelas *0_open-eyes*, *1_pre-drowsy*, dan *2_drowsy*.

Proses pelatihan model dieksekusi menggunakan pembagian dataset dengan proporsi 70% data latih, 20% data validasi, dan 10% data uji. Selama proses ini, beberapa *hyperparameter* kunci disetel secara presisi untuk meminimalkan nilai *loss function*. Parameter tersebut meliputi *learning rate* untuk stabilitas, *batch size* untuk efisiensi memori, serta jumlah *epoch*. Pengaturan parameter tersebut dilakukan secara

optimal guna menghasilkan performa model terbaik sekaligus menghindari risiko *overfitting*.

2.5 Evaluation

Tahap *Evaluation* dilakukan untuk mengukur performa model YOLOv8 dalam mendeteksi dan mengklasifikasikan kondisi kantuk mahasiswa berdasarkan ekspresi wajah. Pengujian dilakukan menggunakan data uji guna mengetahui kemampuan generalisasi model dalam mengenali pola baru yang belum pernah dipelajari sebelumnya. Selain itu, performa deteksi objek dievaluasi menggunakan nilai *confidence score* untuk menilai tingkat keyakinan prediksi model. Metrik *Mean Average Precision* (mAP) juga digunakan sebagai parameter utama untuk mengukur ketepatan pelokalan area wajah secara keseluruhan.

Evaluasi klasifikasi model diukur menggunakan metrik *accuracy*, *precision*, *recall*, dan *F1-score*. Seluruh metrik tersebut berfungsi untuk menilai tingkat ketepatan, sensitivitas, dan keseimbangan prediksi pada kelas 0 *open-eyes*, 1 *pre-drowsy*, dan 2 *drowsy*. Analisis performa ini dilakukan secara mendalam untuk mendeteksi potensi terjadinya kesalahan diagnosis klasifikasi selama sistem dijalankan. Model akhir dinyatakan layak diimplementasikan apabila mampu menghasilkan nilai evaluasi yang tinggi dan stabil pada berbagai skenario pengujian.

2.6 Deployment

Tahap *Deployment* merupakan proses implementasi model YOLOv8 yang telah dilatih ke dalam sistem deteksi kantuk berbasis antarmuka pengguna. Pada tahap ini, model diintegrasikan dengan kamera untuk melakukan deteksi kondisi mahasiswa secara *real-time* selama pembelajaran daring. Sistem dirancang agar mampu menampilkan hasil deteksi kondisi *awake*, *pre-drowsy*, dan *drowsy* secara otomatis berdasarkan analisis area mata dan mulut.

Implementasi sistem dilakukan menggunakan aplikasi berbasis web sehingga dapat diakses dengan lebih mudah pada berbagai perangkat. Tahap ini bertujuan untuk memastikan model yang telah dikembangkan dapat digunakan pada kondisi nyata serta mampu membantu proses pemantauan tingkat kewaspadaan mahasiswa selama kegiatan pembelajaran *online*.

3. Hasil dan Pembahasan

Implementasi sistem deteksi kantuk mahasiswa dalam penelitian ini dilakukan melalui pendekatan terstruktur berbasis kerangka kerja CRISP-DM. Seluruh tahapan eksperimen, mulai dari analisis kebutuhan awal hingga penyebaran model, diuraikan secara sistematis untuk memberikan gambaran komprehensif mengenai performa sistem yang dibangun. Fokus pembahasan diarahkan pada proses optimasi algoritma YOLOv8 dan efektivitas penerapannya dalam mengenali kondisi kewaspadaan mahasiswa secara *real-time*. Hasil

mendalam dari masing-masing fase tersebut dipaparkan secara bertahap pada subbab berikut.

3.1 Business Understanding

Tahap awal penelitian difokuskan pada identifikasi hambatan instruksional dalam pelaksanaan pembelajaran daring di perguruan tinggi serta perumusan kebutuhan teknologi. Berdasarkan analisis situasi, kendala utama yang ditemukan adalah penurunan efektivitas transfer pengetahuan yang dipicu oleh fenomena kantuk mahasiswa akibat kelelahan visual di depan layar. Di sisi lain, dosen mengalami keterbatasan akomodasi dalam memantau kondisi fisiologis mahasiswa secara manual karena keterbatasan ruang pandang pada platform konferensi video. Oleh karena itu, dirumuskan kebutuhan mendesak akan sebuah sistem berbasis *computer vision* yang mampu melakukan pemantauan kondisi kewaspadaan secara otomatis, objektif, dan *real-time*.

Untuk menjawab kebutuhan tersebut, ditetapkan spesifikasi teknis dan parameter keberhasilan dari sistem deteksi kantuk yang dikembangkan. Sistem dirancang untuk mampu mengekstraksi fitur wajah mahasiswa secara langsung menggunakan algoritma YOLOv8 dan mengklasifikasikannya ke dalam tiga kondisi: *awake*, *pre-drowsy*, dan *drowsy*. Kriteria keberhasilan pemodelan ditetapkan berdasarkan pencapaian metrik evaluasi yang tinggi, dengan nilai mAP minimal 0,90 dan kecepatan inferensi di bawah 10 milidetik per citra. Luaran akhir dari fase ini menjadi acuan utama dalam merancang arsitektur perangkat lunak *AwakeLens* serta penyusunan *dashboard* pemantauan bagi dosen.

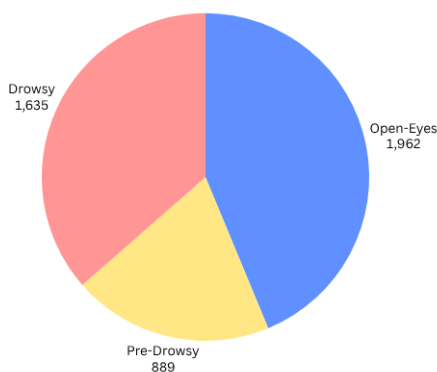
3.2 Data Understanding

Pada tahap ini dilakukan proses pemahaman terhadap data yang digunakan dalam penelitian, meliputi pengumpulan data, struktur serta format anotasi, hingga konfigurasi *dataset* yang digunakan untuk pelatihan model deteksi kantuk mahasiswa. *Dataset* dalam penelitian ini diperoleh dari platform Roboflow melalui penggabungan dua sumber dataset publik, yaitu "*SleepyDetect*" dan "*Yawn*". Dataset gabungan tersebut kemudian melalui proses seleksi, kurasi visual, dan penyesuaian kembali untuk memastikan relevansinya dalam mengenali karakteristik citra ekspresi wajah mahasiswa pada berbagai kondisi pencahayaan dan sudut pandang. *Dataset* tersebut dapat diakses melalui tautan berikut:

<https://universe.roboflow.com/drowsy-detection-bglc2/awakelens-2-5>

Berdasarkan hasil kurasi awal, *dataset* dikategorikan ke dalam tiga kelas utama untuk mendukung analisis perilaku mahasiswa. Kelas *open-eyes* menggunakan 1.962 citra mata terbuka sebagai representasi kondisi terjaga, kelas *pre-drowsy* menggunakan 889 citra mulut terbuka (menguap) sebagai indikasi awal gejala kantuk, dan kelas *drowsy* terdiri dari 1.635 citra mata tertutup untuk merepresentasikan fase kantuk tingkat lanjut atau

microsleep. Rincian distribusi jumlah data untuk setiap kategori kondisi wajah tersebut disajikan dalam bentuk grafik lingkaran pada Gambar 2.



Gambar 2. Chart Penyebaran Dataset per Kelas

3.3 Data Preparation

Tahap berikutnya adalah data preparation yang bertujuan untuk mempersiapkan dataset agar siap digunakan dalam proses pelatihan model deteksi kantuk mahasiswa menggunakan YOLOv8. Tahapan yang dilakukan mencakup verifikasi kesesuaian koordinat *bounding box*, pelabelan ekspresi wajah ke dalam format standar YOLO, pembagian *dataset* awal, serta penerapan *preprocessing* dan *augmentasi* data. Langkah awal difokuskan pada pengondisian data melalui metode penyeimbangan data (*manual balancing*). Setelah *dataset* awal bersih dan seimbang, data langsung dibagi menjadi tiga *subset* utama dengan proporsi 70:20:10 guna menjamin objektivitas evaluasi model, sebagaimana disajikan secara lengkap pada Tabel 1.

Tabel 1. Pembagian *Dataset* Penelitian

Subset	Proporsi	Jumlah
Data Training	70 %	2.497 sampel
Data Validation	20 %	713 sampel
Data Testing	10 %	357 sampel

Agar model tangguh menghadapi variasi kondisi riil perkuliahan daring, dilakukan ekspansi data melalui teknik augmentasi sistematis yang meliputi manipulasi geometris (*rotation* dan *horizontal flip*) serta modifikasi eksposur pencahayaan (*brightness* dan *contrast*). Melalui implementasi rangkaian teknik augmentasi tersebut, volume *dataset* berhasil ditingkatkan secara signifikan dari total awal sebanyak 3.567 sampel menjadi 8.573 citra pasca-augmentasi. Adapun tahapan detail *preprocessing* yang diterapkan meliputi: data (*data cleaning*) dan penyelarasan resolusi gambar (*resizing*) menjadi ukuran seragam 640×640 piksel guna mengoptimalkan proses ekstraksi fitur visual pada area orbital mata dan *perioral* mulut. Seluruh teknik augmentasi beserta deskripsi parameternya dirinci secara lengkap pada Tabel 2.

Tabel 2. Teknik Augmentasi yang diterapkan

No	Deskripsi	Gambar Asli	Hasil Asugmentasi
1.	<i>Horizontal Flip</i> , Membalik citra secara horizontal untuk merekayasa variasi sudut pandang kamera simetris.		
2.	<i>Crop</i> , Pemotongan area objek (skala 0%-30%) agar model adaptif terhadap fitur wajah yang terpotong bingkai.		
3.	<i>Rotation</i> , Memutar citra dalam rentang -15° hingga +15° untuk mensimulasikan kemiringan kepala mahasiswa.		
4.	<i>Brightness</i> , Memodifikasi kecerahan (rentang -30% hingga +30%) guna meniru fluktuasi intensitas cahaya ruangan.		
5.	<i>Blur</i> , Menerapkan efek kabur hingga 1.7 px untuk mengantisipasi penurunan fokus akibat <i>motion blur</i> kamera.		

3.4 Modeling

Penelitian ini mengimplementasikan algoritma YOLOv8 berbasis *deep learning* yang dikonfigurasi menggunakan varian arsitektur YOLOv8 Nano (YOLOv8n) melalui pendekatan *transfer learning* [12]. Varian ini dipilih karena strukturnya yang ringan dengan kapasitas sekitar 3,2 juta parameter, sehingga sangat ideal untuk menghasilkan komputasi berkecepatan tinggi yang dibutuhkan pada sistem pemantauan kelas daring secara *real-time*. Proses ekstraksi fitur dilakukan oleh jaringan untuk mengenali representasi piksel area kelopak mata dan pergerakan mulut guna mengklasifikasikan kondisi kewaspadaan

ke dalam tiga kelas target: *open-eyes*, *pre-drowsy*, dan *drowsy*.

Proses pelatihan model dieksekusi menggunakan lingkungan komputasi berbasis perangkat keras GPU (*Graphics Processing Unit*) untuk mempercepat waktu iterasi dan mengoptimalkan efisiensi pemrosesan data gambar. Guna menghasilkan performa model yang optimal dan mencegah terjadinya risiko *overfitting*, beberapa *hyperparameter* kunci disetel secara presisi selama proses *training* berlangsung. Konfigurasi seluruh parameter utama yang diterapkan dalam pengujian pemodelan ini dirinci secara sistematis pada Tabel 3.

Tabel 3. Rincian Parameter Pelatihan Model YOLOv8

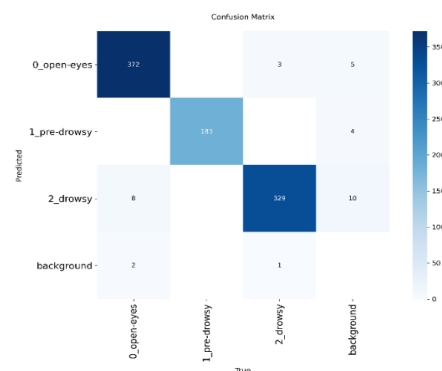
Parameter	Nilai	Keterangan
<i>epochs</i>	100	Jumlah siklus pelatihan maksimal. Model akan mempelajari seluruh <i>dataset</i> sebanyak 100 kali putaran.
<i>patience</i>	10	Mekanisme <i>Early Stopping</i> . <i>Training</i> dihentikan otomatis apabila akurasi validasi tidak meningkat selama 10 <i>epoch</i> berturut-turut, mencegah <i>overfitting</i> .
<i>imgsz</i>	640	Resolusi input citra (640 x 640 piksel). Ukuran ini merupakan standar YOLOv8 yang optimal untuk menjaga detail fitur wajah seperti kondisi mata dan mulut.
<i>batch</i>	16	Jumlah citra yang diproses secara paralel dalam satu iterasi. Nilai 16 disesuaikan dengan kapasitas memori GPU NVIDIA Tesla T4 (16 GB VRAM).
<i>device</i>	0 (GPU)	Mengarahkan seluruh komputasi ke GPU primer (indeks 0) agar proses <i>training</i> berjalan menggunakan akselerasi <i>hardware</i> secara penuh.
<i>plots</i>	True	Mengaktifkan pembangkitan grafik evaluasi otomatis (<i>kurva loss</i> , <i>mAP</i> , <i>confusion matrix</i>) di akhir proses <i>training</i> untuk keperluan analisis.
<i>save</i>	True	Mengaktifkan penyimpanan bobot model terbaik (<i>best.pt</i>) dan terakhir (<i>last.pt</i>) selama proses pelatihan berlangsung.

Hasil akhir dari proses pemodelan ini menghasilkan berkas bobot optimal bernama *best.pt* yang memiliki ukuran sangat ringan. Berkas bobot inilah yang kemudian diekstrak dan disimpan sebagai komponen kecerdasan utama untuk diuji keandalannya pada tahapan evaluasi berikutnya.

3.5 Evaluation

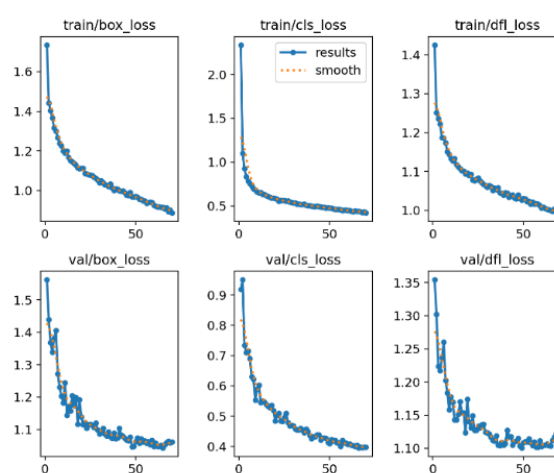
Evaluasi performa model YOLOv8n dilakukan secara objektif menggunakan subset data validasi eksternal yang belum pernah dilatih sebelumnya. Hasil *Confusion Matrix* pada Gambar 3 menunjukkan korelasi hasil antara nilai aktual (*ground truth*) dan prediksi model untuk ketiga kelas target. Berdasarkan data absolut tersebut, model menunjukkan kapabilitas generalisasi yang kuat dengan akumulasi nilai yang dominan pada diagonal utama matriks.

Untuk menjamin transparansi data, hasil pengujian matriks dinarasikan secara eksplisit berdasarkan total sampel aktual per kelas. Sebanyak 372 citra dari total 374 sampel aktual mata terbuka (*open-eyes*) berhasil diklasifikasikan dengan tepat. Sementara itu, kelas *pre-drowsy* (gejala menguap) mencatatkan performa sempurna dengan 183 citra dari total 183 sampel terdeteksi secara 100% akurat tanpa adanya kesalahan klasifikasi silang (*misclassification*) ke kelas lainnya.



Gambar 3. Hasil *Confusion Matrix*

Meskipun performa keseluruhan sangat tinggi, terdapat deviasi minor pada pengenalan fase kantuk berat. Dari total 347 sampel aktual pada kelas *drowsy*, sebanyak 329 citra berhasil diklasifikasikan dengan tepat. Adapun akumulasi kesalahan klasifikasi pada kelas ini berjumlah 18 citra, yang rinci atas 8 sampel salah diprediksi sebagai *open-eyes* akibat kemiripan visual pada durasi kedipan mata yang sangat singkat, serta 10 sampel terdeteksi sebagai *background* akibat faktor oklusi wajah atau intensitas cahaya ekstrem di bawah ambang batas *confidence score*. Namun, tingkat margin kesalahan tersebut masih berada dalam batas toleransi penelitian karena persentase akurasi total kelas *drowsy* tetap bertahan tinggi di angka 94,8%.

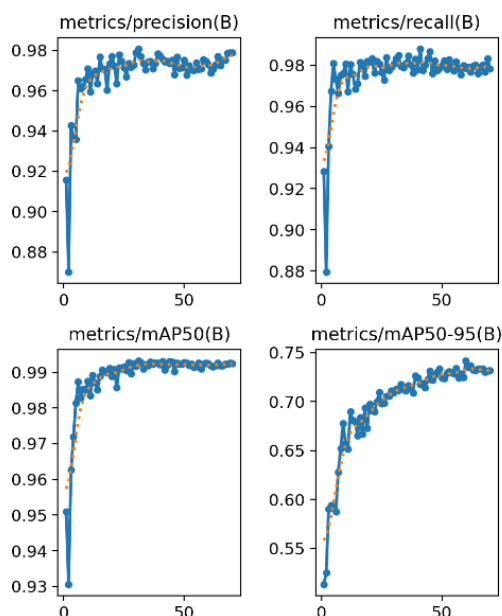


Gambar 4. Grafik *loss function*

Stabilitas konvergensi model selama pelatihan terekam melalui kurva penurunan fungsi kerugian (*loss function*) pada Gambar 4. Untuk memvalidasi klaim optimalisasi model, dilakukan analisis komparatif performa nilai *loss* antara epoch awal dan akhir. Pada

epoch 1, nilai *train/box_loss* tercatat tinggi sebesar 1,72 dan *train/cls_loss* sebesar 2,34. Model YOLOv8n mencapai konvergensi dan titik optimalnya setelah melewati epoch ke-70. Dengan keberadaan mekanisme *Early Stopping* berbasis nilai *patience* 10, proses pelatihan secara otomatis dihentikan pada *epoch* ke-72 karena metrik fungsi kerugian pada data validasi telah stabil sejak epoch ke-62.

Pada *epoch* terakhir (*epoch* ke-72), nilai *train/box_loss* berhasil ditekan hingga berakhir rendah di kisaran 0,88 yang menandakan tingkat presisi tinggi penentuan posisi *bounding box*. Penurunan nilai yang signifikan juga terjadi pada *train/cls_loss* (menjadi 0,43) dan *train/dfl_loss* (menjadi 0,99), yang membuktikan minimalnya kesalahan klasifikasi serta kemampuan model dalam menangani variasi bentuk objek yang dinamis. Hasil pengujian data luar menunjukkan nilai *val/box_loss* stabil di angka 1,06 dan *val/cls_loss* di angka 0,39, mengonfirmasi tidak adanya celah (*gap*) lebar antara kurva pelatihan dan validasi (*overfitting*).



Gambar 5. Grafik Recall, Precision, dan mAP

Berdasarkan akumulasi data kuantitatif performa yang dirinci pada Tabel 4, model menghasilkan keseimbangan akurasi yang optimal ditandai dengan rata-rata nilai *F1-Score* sebesar 0,976. Nilai *Precision* keseluruhan mencapai 0,973 dengan tingkat sensitivitas *Recall* di angka 0,979, membuktikan ketajaman model dalam mendeteksi fitur wajah tanpa menghasilkan prediksi *false positive*. Kelas *open-eyes* menjadi kelas paling unggul pada pengujian ketat mAP50-95 dengan skor 0,824, sedangkan kelas *drowsy* yang secara visual lebih kompleks tetap mencatatkan performa mAP50 yang tinggi sebesar 0,992. Ditambah dengan kecepatan inferensi yang sangat responsif sebesar 4,1 milidetik per citra, hasil evaluasi ini memberikan validasi penuh bahwa model siap diintegrasikan untuk sistem pemantauan perkuliahan daring secara *real-time*.

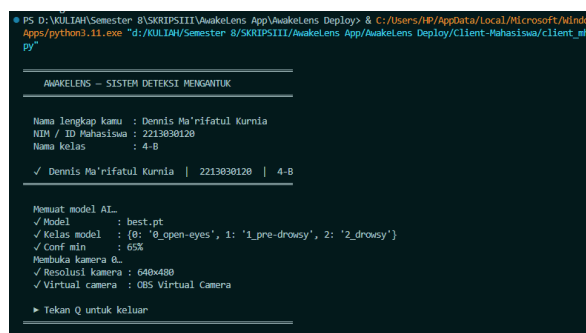
Tabel 4 : Hasil Akhir Evaluasi Model

Kelas	Precisi on (P)	Recall (R)	mAP50	mAP50 -95	F1- Score
Rata-rata (All)	0.973	0.979	0.992	0.741	0.976
0_open- eyes	0.984	0.972	0.993	0.824	0.978
1_pre- drowsy	0.976	0.978	0.992	0.628	0.977
2_drowsy	0.959	0.985	0.992	0.770	0.972

3.6 Deployment

Tahap akhir *deployment* diimplementasikan melalui arsitektur perangkat lunak bernama aplikasi AwakeLens. Sistem ini mengintegrasikan model inferensi terbaik (*best.pt*) ke dalam program lokal berbasis Python dengan menetapkan nilai ambang batas kepercayaan (*confidence threshold*) minimal sebesar 65%. Justifikasi pemilihan *threshold* 65% ini ditentukan secara empiris melalui serangkaian pengujian awal (*fine-tuning trade-off*). Batas kepercayaan yang terlalu tinggi ($\geq 80\%$) ditemukan menyebabkan tingginya angka gagal deteksi (*miss-detection*) saat mahasiswa melakukan gerakan kemiringan kepala ekstrem. Sebaliknya, batas yang terlalu rendah ($\leq 50\%$) rentan memicu galat prediksi (*false positive*) akibat distorsi latar belakang objek. Nilai 65% terbukti menjadi titik optimal yang menyeimbangkan metrik *Precision* dan *Recall* selama aliran video berlangsung.

Pada saat inisialisasi awal, pengguna diwajibkan melakukan input data identitas meliputi Nama, NIM, dan Kelas melalui antarmuka terminal seperti Gambar 6 untuk kebutuhan sinkronisasi *database server*. Aliran video dari kamera diatur pada resolusi standar 640 x 480 px dan dikonfigurasi menggunakan fitur *virtual camera* agar hasil deteksi dapat langsung disematkan pada *platform* konferensi video seperti Zoom atau Google Meet.

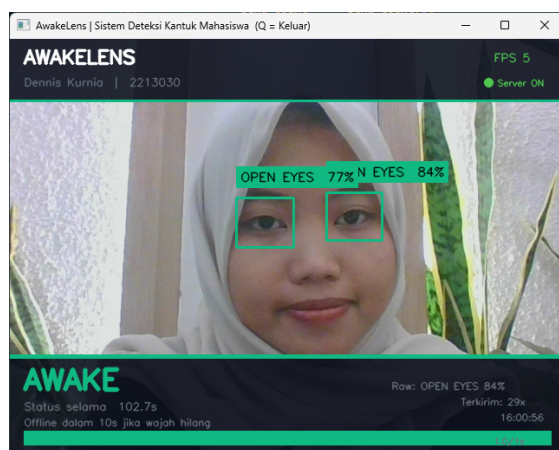


Gambar 6. Antarmuka Terminal Inisialisasi Sistem AwakeLens

Untuk memvalidasi keandalan operasionalnya, aplikasi AwakeLens telah diuji secara langsung dalam skenario simulasi perkuliahan daring tatap maya yang melibatkan subjek mahasiswa aktif. Berdasarkan hasil pengujian pada kondisi video langsung (*live video streaming*), sistem berhasil memproses data visual secara konstan dengan menghasilkan performa kecepatan penayangan bingkai pada rentang 4 hingga 7 FPS. Dalam perspektif pemantauan perilaku

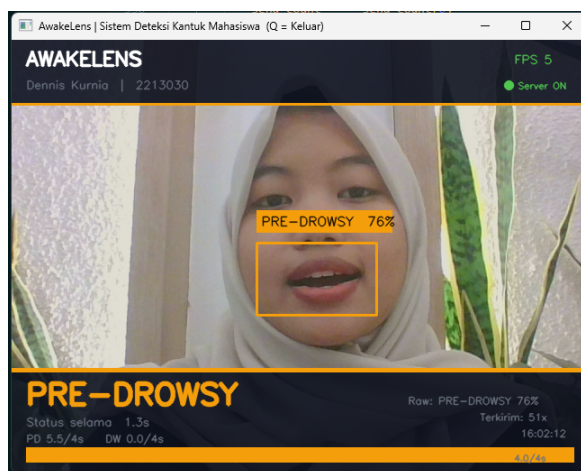
komputasi, capaian 4–7 FPS ini dinyatakan sangat memadai dan andal untuk kebutuhan ruang kelas daring. Hal ini dikarenakan tanda-tanda penurunan kesiagaan belajar (seperti durasi mata terpejam pada microsleep atau pergerakan rahang saat menguap) dikategorikan sebagai transisi gerakan makro/lambat, sehingga tidak membutuhkan *frame rate* tinggi layaknya pelacakan objek berkecepatan tinggi (*high-speed action tracking*).

Mekanisme inferensi visual model secara *real-time* terbagi menjadi tiga kondisi klasifikasi ekspresi wajah mahasiswa dengan skema *feedback* warna antarmuka yang adaptif. Saat mahasiswa dalam kondisi terjaga dan fokus mengikuti pembelajaran, model mendeteksi kelas *0_open-eyes* dengan skor kepercayaan tinggi berkisar antara 77%–84% yang ditandai dengan indikator status berwarna hijau bertuliskan "*AWAKE*" seperti pada Gambar 7.



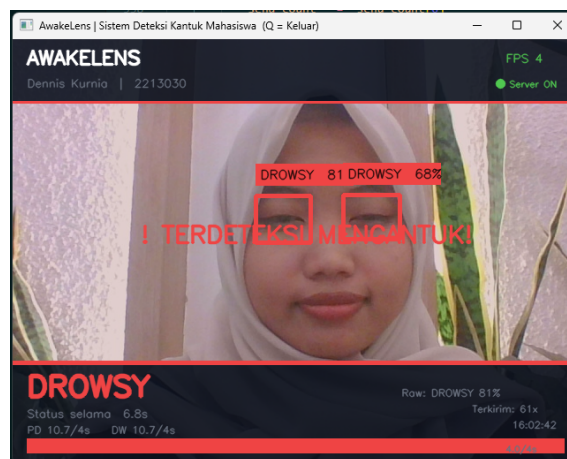
Gambar 7. Visualisasi Status *Awake* (Terjaga)

Fitur peringatan dini (*early warning*) diaktifkan ketika model mendeteksi aktivitas mulut menguap (*1_pre-drowsy*) dengan skor kepercayaan mencapai 76%, yang secara responsif mengubah antarmuka menjadi warna jingga bertuliskan "*PRE-DROWSY*" sebagaimana diilustrasikan pada Gambar 8.



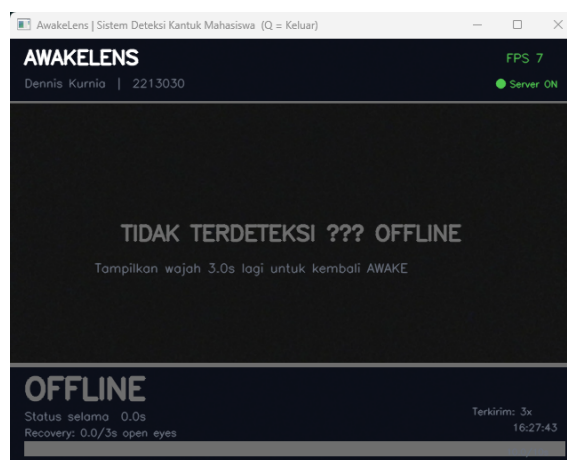
Gambar 8. Visualisasi Status *Pre-Drowsy* (Gejala Kantuk) saat Mahasiswa Menguap

Fase paling kritis ditangani melalui logika pewaktu (*timer logic*), di mana jika model mendeteksi mata tertutup (*2_drowsy*) dengan skor kepercayaan 81% selama durasi akumulatif 4 detik, sistem akan memicu alarm visual warna merah bertuliskan "*TERDETEKSI MENGANTUK !!*" yang diperlihatkan pada Gambar 9.



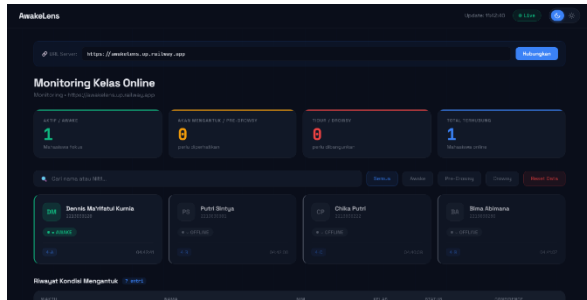
Gambar 9. Visualisasi Peringatan saat kondisi Status *Drowsy* (Mengantuk)

Fitur Validasi Kehadiran dan Integritas Status *Offline* Untuk menjaga integritas dan validitas data selama jalannya perkuliahan daring, aplikasi dilengkapi dengan modul pengawasan eksistensi objek wajah dalam bingkai (*frame*) kamera. Jika mahasiswa meninggalkan perangkat sehingga objek wajah tidak terdeteksi oleh sensor kamera dalam durasi ambang batas (*buffer duration*) lebih dari 10 detik, sistem secara otomatis menghentikan proses klasifikasi kelas ekspresi. Kondisi ini akan langsung memicu transisi sistem ke status "*OFFLINE*" yang ditandai dengan perubahan visual antarmuka menjadi warna abu-abu seperti yang ditampilkan pada Gambar 10. Implementasi mekanisme ini sangat krusial untuk mengeliminasi anomali data atau kesalahan prediksi klasifikasi (*false detection*) yang disebabkan oleh faktor oklusi latar belakang ketika area tangkapan kamera kosong.



Gambar 10. Visualisasi antarmuka *AwakeLens* saat mahasiswa berada di luar jangkauan kamera (Status *Offline*).

Seluruh aliran data dari aplikasi *client* mahasiswa dikirimkan secara berkala menuju server terpusat berbasis *cloud* untuk divisualisasikan pada *Dashboard Monitoring Kelas Online* milik dosen yang tertera pada Gambar 11. Dasbor berbasis *web* ini mengintegrasikan ringkasan statistik kelas secara kolektif yang menampilkan grafik riil jumlah mahasiswa berstatus *awake*, *pre-drowsy*, *drowsy*, maupun *offline*. Melalui antarmuka terpusat ini, pengajar dapat mengidentifikasi mahasiswa yang memerlukan intervensi secara cepat lewat kartu statistik khusus berkategori "Perlu Diperhatikan" atau "Perlu Dibangunkan".



Gambar 11. *Dashboard Monitoring Kelas Online* untuk Pengawasan Dosen

Integrasi data historis di dalam dasbor juga menyajikan rekapitulasi individu secara mendetail yang memuat data Nama, NIM, Kelas, stempel waktu (*timestamp*), serta skor kepercayaan model pada setiap entri data. Adanya fitur pencarian spesifik dan indikator status *Live* pada pojok kanan atas memastikan aliran data yang diterima oleh dosen bersifat aktual. Implementasi menyeluruh dari kerangka kerja CRISP-DM ini membuktikan bahwa kombinasi model YOLOv8n yang ringan dengan sistem pemantauan terdistribusi mampu memberikan solusi instruksional yang efisien, objektif, dan terukur dalam menjaga kualitas pembelajaran daring.

4. Kesimpulan

Penelitian ini berhasil mengembangkan sistem deteksi kantuk mahasiswa pada kelas online menggunakan arsitektur YOLOv8 Nano melalui analisis ekspresi wajah pada area mata dan mulut. Pengujian empiris pada tahap evaluasi menunjukkan bahwa model yang dilatih dengan total 8.573 citra pasca-augmentasi menghasilkan performa yang sangat stabil dengan rata-rata nilai *F1-Score* sebesar 0,976, *Precision* sebesar 0,973, *Recall* sebesar 0,979, serta nilai mAP50 mencapai 0,992. Pengujian menggunakan *Confusion Matrix* juga membuktikan fakta bahwa tingkat akurasi klasifikasi per kelas berada pada level yang sangat tinggi, yaitu mencapai 100% untuk pengenalan gejala awal kantuk (*pre-drowsy*) dan 95% untuk fase kantuk berat (*drowsy*). Hasil empiris ini membuktikan bahwa arsitektur model mampu menjawab permasalahan deteksi objek visual wajah secara akurat.

Implikasi dari penelitian ini mewujudkan pada kesiapan aplikasi AwakeLens untuk diterapkan secara riil pada skenario perkuliahan daring yang sebenarnya. Ditopang oleh kecepatan inferensi model yang responsif sebesar 4,1 milidetik per citra, sistem ini memiliki kemungkinan aplikasi yang luas sebagai *virtual camera* pada platform konferensi video tanpa membebani kapasitas komputasi perangkat lokal. Keberadaan fitur intervensi dini berbasis alarm visual dan pengawasan status *offline* terbukti mampu meminimalkan celah manipulasi data kehadiran serta mengoptimalkan pengawasan instruksional dosen secara terpusat melalui integrasi modul kendali keberadaan wajah (*face presence control*) yang otomatis memutus deteksi saat mahasiswa meninggalkan perangkat.

Meskipun memiliki performa metrik yang tinggi, penelitian ini memiliki keterbatasan utama pada aspek variabilitas data primer, di mana dataset yang digunakan seutuhnya masih bergantung pada platform publik sekunder Roboflow sehingga belum menguji keunikan struktur antropometri wajah mahasiswa Indonesia secara masif. Selain itu, saat model diimplementasikan pada aplikasi video langsung menggunakan perangkat kamera web (*webcam*) standar tingkat konsumen, kecepatan penayangan bingkai masih terbatas pada kisaran 4 hingga 7 FPS akibat batasan spesifikasi perangkat keras lokal.

Sebagai langkah pengembangan untuk mengatasi keterbatasan tersebut, penelitian selanjutnya disarankan untuk mengumpulkan dataset primer secara langsung dari mahasiswa aktif Indonesia dalam konteks pembelajaran daring nyata. Pengumpulan data tersebut wajib mencakup variasi etnis nusantara, kondisi pencahayaan ruangan kamar yang khas, serta tingkat resolusi kamera yang beragam guna meningkatkan validitas eksternal dan kemampuan generalisasi model terhadap populasi target yang sesungguhnya.

Daftar Rujukan

- [1] M. Lepore and I. Delfino, "Optical Sensors Technology and Applications," *Sensors*, vol. 22, no. 20, p. 7905, Oct. 2022, doi: 10.3390/s2207905.
- [2] T. Tanantong and P. Yongwattana, "A convolutional neural network framework for classifying inappropriate online video contents," *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 12, no. 1, p. 124, Mar. 2023, doi: 10.11591/ijai.v12.i1.pp124-136.
- [3] M. K. Reddy, Y. M. Keerthana, and P. Alku, "End-to-End Pathological Speech Detection Using Wavelet Scattering Network," *IEEE Signal Process. Lett.*, vol. 29, pp. 1863–1867, 2022, doi: 10.1109/LSP.2022.3199669.
- [4] D. I. Mulyana and D. Iskandar, "Implementasi Deteksi Objek Real Time Jenis Sampah Dengan Algoritma Yolov8," *INTECOMS: Journal of Information Technology and Computer Science*, vol. 7, no. 3, pp. 631–638, May 2024, doi: 10.31539/intecom.v7i3.8985.
- [5] D. Triyanto, M. Zidan, M. Wahyudi, L. Pujiastuti, and S. Sumanto, "Pengembangan Sistem Deteksi Objek Botol Real-Time dengan YOLOv8 untuk Aplikasi Vision," *Indonesian Journal Computer Science*, vol. 3, no. 1, pp. 44–50, Apr. 2024, doi: 10.31294/ijcs.v3i1.6070.

-
- [6] R. S. Saputri, A. Apriliani, R. S. Amar, and L. Y. Astri, "Deteksi Dini Microsleep Pada Pengemudi Kendaraan Roda Empat Melalui Citra Mata Menggunakan Algoritma YOLOv8," *Jurnal PROCESSOR*, vol. 20, no. 2, Oct. 2025, doi: 10.33998/processor.2025.20.2.2529.
- [7] Edmund Ucok Armin, Anggun Purnama Edra, Fakhri Ikhwanul Alifin, Ikhwanussafa Sadidan, Indri Purwita Sary, and Ulinnuha Latifa, "Performa Model YOLOv8 untuk Deteksi Kondisi Mengantuk pada pengendara mobil," *BRAHMANA: Jurnal Penerapan Kecerdasan Buatan*, vol. 5, Dec. 2023, doi: 10.30645/brahmana.v5i1.279.
- [8] S. Sugeng and T. N. Nizar, "Deteksi Aktivitas Mata, Mulut Dan Kemiringan Kepala Sebagai Fitur Untuk Deteksi Kantuk Pada Pengendara Mobil," *Komputika : Jurnal Sistem Komputer*, vol. 12, no. 1, pp. 83–91, May 2023, doi: 10.34010/komputika.v12i1.9688.
- [9] N. N. Pandey and N. B. Muppalaneni, "A novel drowsiness detection model using composite features of head, eye, and facial expression," *Neural Comput. Appl.*, vol. 34, no. 16, pp. 13883–13893, Aug. 2022, doi: 10.1007/s00521-022-07209-1.
- [10] A. Ardiansyah, A. Sri Widagdo, K. Nuresa Qodri, D. Hidayani, and M. Romadhani, "Implementasi Deteksi Tumor Otak Menggunakan YOLOv11 dan Flask," *JURNAL FASILKOM*, vol. 15, no. 2, pp. 415–422, Sep. 2025, doi: 10.37859/jf.v15i2.9703.
- [11] R. Amelia, "HUBUNGAN KUALITAS TIDUR TERHADAP PRESTASI AKADEMIK MAHASISWA KEDOKTERAN UNIVERSITAS BAITURRAHMAH," *EDUCATOR : Jurnal Inovasi Tenaga Pendidik dan Kependidikan*, vol. 2, no. 1, pp. 28–34, Mar. 2022, doi: 10.51878/educator.v2i1.1002.
- [12] T. A. J. JABAR, R. Heriansyah, and E. Purnamasari, "Pengenalan Pola Huruf Bahasa Isyarat Menggunakan Framework You Only Look Once (YOLO)," *JURNAL FASILKOM*, vol. 15, no. 2, pp. 327–335, Aug. 2025, doi: 10.37859/jf.v15i2.10018.
- [13] A. Nugroho and D. Harini, "Teknik Random Forest untuk Meningkatkan Akurasi Data Tidak Seimbang," *JSITIK*, vol. 2, no. 2, 2024, doi: 10.53624/jsitik.v2i2.XX.