

PENGEMBANGAN SISTEM DETEKSI KANTUK WAKTU NYATA BERBASIS EYE ASPECT RATIO DAN MEDIAPIPE FACE MESH

Agustian Putra Sukarya¹, Achmad Rizal Wahid², Untung Rohwadi³, Felix Wuryo Handono⁴, Sigit Wibawa⁵

Email: agustianputra2108@gmail.com¹, rizal.waredmi@gmail.com²,
untung.unr@bsi.ac.id³, felix@bsi.ac.id⁴, sigit.stb@bsi.ac.id⁵

^{1,2,3,4,5}Universitas Bina Sarana Informatika

ABSTRAK

Kantuk dan kelelahan pengguna, baik pengemudi kendaraan maupun operator perangkat, merupakan kontributor utama dalam kasus kecelakaan dan penurunan produktivitas.¹ Data statistik menunjukkan bahwa kantuk dapat menyumbang sekitar 20% dari total kecelakaan lalu lintas.² Kebutuhan mendesak untuk sistem pemantauan yang non-invasif, ekonomis, dan beroperasi dalam waktu nyata mendorong pengembangan solusi berbasis visi komputer. Penelitian ini bertujuan untuk merancang dan mengimplementasikan sistem deteksi kantuk yang akurat dan responsif menggunakan metode Eye Aspect Ratio (EAR) yang diekstraksi secara efisien melalui MediaPipe Face Mesh. Sistem ini memanfaatkan MediaPipe Face Mesh untuk mengestimasi 468 titik landmark wajah 3D secara presisi.³ Logika keputusan ditetapkan ketika nilai EAR rata-rata mata jatuh di bawah ambang batas \$0.25\$ 4 selama minimal 18 consecutive frames. Durasi 18 frame ini setara dengan \$0.6\$ detik pada 30 Frames Per Second (FPS), yang efektif memfilter kedipan normal dan secara temporal mengidentifikasi kondisi microsleep.⁶ Peringatan insiden dilakukan secara multimodal melalui alarm suara non-blocking menggunakan threading, didukung oleh peringatan visual, pencatatan data (logging) ke CSV, dan pengambilan automatic screenshot untuk dokumentasi insiden. Hasil implementasi menunjukkan bahwa sinergi MediaPipe yang cepat dengan mekanisme peringatan non-blocking menghasilkan sistem yang stabil dan memadai untuk aplikasi waktu nyata, menawarkan alternatif yang hemat biaya dibandingkan metode fisiologis yang mahal.

Kata Kunci: Deteksi Kantuk, Eye Aspect Ratio (Ear), Mediapipe Face Mesh, Waktu Nyata, Peringatan Multimodal.

ABSTRACT

Drowsiness and user fatigue, whether in vehicle drivers or device operators, are major contributors to accidents and decreased productivity.¹ Statistical data indicates that drowsiness accounts for approximately 20% of total traffic accidents.² The urgent need

for a non-invasive, economical, and real-time monitoring system motivates the development of computer vision-based solutions. This research aims to design and implement an accurate and responsive drowsiness detection system utilizing the Eye Aspect Ratio (EAR) method efficiently extracted via MediaPipe Face Mesh. The system leverages MediaPipe Face Mesh to precisely estimate 468 3D facial landmarks.³ The decision logic is triggered when the mean EAR value drops below the threshold of \$0.25\$⁴ for a minimum of 18 consecutive frames. This duration of 18 frames equates to \$0.6\$ seconds at 30 Frames Per Second (FPS), effectively filtering out normal blinks and temporally identifying microsleep conditions.⁶ Incident warnings are delivered multimodally through a non-blocking audio alarm utilizing threading, supported by visual alerts, data logging to CSV, and automatic screenshot capture for incident documentation. Implementation results demonstrate that the synergy between fast MediaPipe processing and the non-blocking warning mechanism yields a stable system suitable for real-time applications, offering a cost-effective alternative to expensive physiological methods.

Keywords: *Drowsiness Detection, Eye Aspect Ratio (Ear), Mediapipe Face Mesh, Real-Time, Multimodal Warning.*

1. PENDAHULUAN

Kantuk, yang didefinisikan sebagai tingkat kelelahan ekstrem yang menyebabkan penurunan kewaspadaan dan waktu reaksi, merupakan faktor risiko yang signifikan dalam kecelakaan industri dan transportasi. Data mencatat bahwa kelelahan dapat meningkatkan risiko kecelakaan hingga tiga kali lipat. Akibatnya, pengembangan sistem deteksi kantuk pengemudi (DDD) yang akurat dan responsif telah menjadi area penelitian kritis dalam visi komputer.

Pendekatan untuk deteksi kantuk secara tradisional dibagi menjadi metode fisiologis, seperti penggunaan sensor elektroensefalogram (EEG) atau elektrokardiogram (EKG), dan metode perilaku atau berbasis citra. Meskipun metode fisiologis memberikan pengukuran objektif, biaya implementasinya yang tinggi dan sifat invasif atau contact membuatnya kurang praktis untuk penggunaan sehari-hari. Oleh karena itu, sistem berbasis visi komputer non-kontak menjadi pilihan utama karena sifatnya yang ekonomis dan kemudahan integrasinya menggunakan kamera standar untuk memantau indikator kelelahan visual, seperti tingkat penutupan mata (EAR), rasio bukaan mulut (MAR), atau posisi kepala.

Tantangan utama dalam merancang sistem berbasis citra yang efektif adalah menjaga kinerja real-time (tinggi FPS) sementara mengintegrasikan kemampuan deteksi

yang akurat. Latensi pemrosesan yang tinggi dapat menyebabkan frame terlewatkan, mengurangi akurasi deteksi temporal kantung. Penelitian ini mengatasi tantangan ini dengan memanfaatkan kerangka kerja MediaPipe yang sangat efisien untuk ekstraksi fitur dan merancang mekanisme respons multimodal yang tidak memblokir alur pemrosesan video utama

2. TINJAUAN PUSTAKA

Konsep dasar dalam deteksi kantung berbasis citra adalah metrik Eye Aspect Ratio (EAR), yang berfungsi sebagai kuantifikasi geometris sederhana dari tingkat keterbukaan mata. Metode EAR banyak digunakan dalam studi lokal karena sensitivitasnya yang tinggi terhadap penutupan mata yang dipicu kelelahan [Firdaus et al., 2023; Kadi et al., 2023; Marshella et al., 2024].

Dalam memilih teknologi untuk ekstraksi landmark, penelitian terbaru menyoroti keunggulan MediaPipe Face Mesh. MediaPipe dirancang untuk kinerja real-time yang optimal, bahkan pada perangkat dengan sumber daya komputasi terbatas. MediaPipe Face Mesh mampu mengestimasi 468 titik landmark wajah dalam ruang 3D. Kemampuan untuk menghasilkan landmark 3D ini memberikan ketahanan (robustness) yang lebih baik terhadap variasi orientasi kepala dibandingkan detektor landmark 2D tradisional. Kecepatan dan presisi yang ditawarkan oleh MediaPipe menjadikannya alat yang lebih unggul dibandingkan metode seperti Dlib Shape Predictor atau Haar Cascade untuk aplikasi deteksi kantung waktu nyata.

Penelitian sebelumnya di Indonesia telah mengeksplorasi respons sistem deteksi kantung, mulai dari integrasi dengan sensor tambahan [Marshella et al., 2024] hingga penggunaan notifikasi eksternal melalui GSM atau Telegram untuk memberikan peringatan jarak jauh atau lokasi kecelakaan. Penelitian ini menyempurnakan aspek respons dengan fokus pada solusi in-application yang mengutamakan kecepatan respons melalui non-blocking alarm audio dan menyediakan dokumentasi insiden yang objektif, yang jarang menjadi fokus utama dalam implementasi EAR sederhana.

Kontribusi dan Tujuan Penelitian

Kontribusi penelitian ini adalah mendemonstrasikan sinergi yang efisien antara pemrosesan high-performance MediaPipe Face Mesh dengan logika deteksi temporal yang tervalidasi secara fisiologis. Secara khusus, sistem yang dikembangkan menyajikan:

1. Pemanfaatan MediaPipe Face Mesh untuk ekstraksi landmark mata yang stabil dan presisi.
2. Penetapan logika keputusan kritis ($EAR \leq 0.25$ selama 18 frame) yang memiliki justifikasi kausal yang jelas terhadap durasi microsleep berbahaya (> 0.5 detik).
3. Implementasi sistem peringatan multimodal yang memanfaatkan threading untuk alarm suara (simpleaudio), memastikan operasi non-blocking dan menjaga laju FPS sistem tetap tinggi selama insiden.
4. Penyediaan fitur dokumentasi insiden otomatis melalui automatic screenshot dan pencatatan log data (CSV) yang detail, mendukung tujuan forensik dan analisis data di masa mendatang.

Tujuan utama penelitian ini adalah merancang dan mengimplementasikan sistem perangkat lunak yang secara akurat mendeteksi kantuk di bawah ambang batas EAR 0.25 yang berlangsung selama 18 frame, serta memberikan respons peringatan yang cepat, non-blocking, dan terdokumentasi

3. METODE PENELITIAN

Desain Arsitektur Sistem

Sistem deteksi kantuk ini dirancang sebagai pipeline pemrosesan citra real-time. Komponen utama meliputi akuisisi video, pemrosesan landmark menggunakan MediaPipe Face Mesh, perhitungan Eye Aspect Ratio (EAR), logika keputusan temporal, dan sistem respons multimodal. Implementasi perangkat lunak dilakukan menggunakan Python, OpenCV untuk antarmuka video, NumPy untuk perhitungan matematis vektor, dan pustaka simpleaudio serta pyttsx3 untuk alarm audio.

Ekstraksi Landmark Wajah Menggunakan MediaPipe Face Mesh

MediaPipe Face Mesh digunakan untuk mengidentifikasi wajah dan mengestimasi 468 titik landmark 3D. Untuk mengoptimalkan kinerja dan akurasi, modul diinisialisasi dengan konfigurasi `static_image_mode=False` untuk mode video, dan `refine_landmarks=True` untuk meningkatkan ketepatan estimasi, terutama di area kontur mata dan bibir.

Tabel Indeks Landmark MediaPipe untuk Perhitungan EAR

Komponen EAR	Titik Kritis pada Mata Kiri	Indeks Kiri (LEFT_EYE_IDX)	Titik Kritis pada Mata Kanan	Indeks Kanan (RIGHT_EYE_IDX)
Horizontal 1 (P1, P4)	Temporal, Nasal	33, 133	Temporal, Nasal	263, 362
Vertikal 1 (P2, P6)	Atas, Bawah	160, 144	Atas, Bawah	387, 373
Vertikal 2 (P3, P5)	Atas, Bawah	158, 153	Atas, Bawah	385, 380

Perhitungan Eye Aspect Ratio (EAR)

EAR adalah metrik yang digunakan untuk mengkuantifikasi keterbukaan mata. Jarak antar titik landmark dihitung menggunakan rumus jarak Euclidean [code euclidean]:

$$d(a, b) = \sqrt{(a_x - b_x)^2 + (a_y - b_y)^2}$$

Di mana \$a\$ dan \$b\$ adalah koordinat dua titik di bidang 2D. Formula EAR kemudian dihitung berdasarkan jarak antara titik-titik kunci vertikal dan horizontal:

$$EAR = \frac{d(P_2, P_6) + d(P_3, P_5)}{2 \cdot d(P_1, P_4)}$$

Penggunaan rata-rata dua jarak vertikal \$d(P_2, P_6)\$ dan \$d(P_3, P_5)\$ dimaksudkan untuk meningkatkan akurasi dan ketahanan pengukuran terhadap distorsi kelopak mata. Nilai rata-rata dari EAR mata kiri dan kanan EAR_mean digunakan sebagai indikator tunggal kondisi kantuk.

Logika Keputusan Kritis

Sistem menggunakan ambang batas $EAR_THRESHOLD = 0.25$ yang harus dicapai secara persisten. Lebih penting lagi, sistem mengimplementasikan filter temporal dengan mengatur $CONSECUTIVE_FRAMES = 18$.

Justifikasi untuk parameter **18 frame** didasarkan pada pemisahan antara kedipan mata refleksi normal dan microsleep. Kedipan normal berlangsung kurang dari **0.3 detik**, sedangkan microsleep didefinisikan sebagai penutupan mata yang berlangsung **0.5 detik atau lebih lama**. Dengan laju frame tipikal **30 FPS**, 18 frame setara dengan durasi **0.6 detik**. Durasi ini memastikan bahwa sistem secara andal mengabaikan kedipan normal (yang terjadi dalam 7–10 frame) dan hanya memicu alarm pada penutupan mata yang memenuhi kriteria temporal microsleep. *Frame counter* direset segera setelah **EAR_mean** melebihi ambang batas, memastikan akurasi deteksi berkelanjutan.

Sistem Peringatan Multimodal dan Dokumentasi

Sistem ini memastikan respons peringatan yang cepat dan komprehensif.

Peringatan Audio Non-Blocking: Untuk mempertahankan kinerja real-time (FPS tinggi), alarm suara (`play_alarm_file`) dijalankan menggunakan modul threading. Langkah ini sangat penting; jika pemutaran audio dijalankan pada main thread, operasi I/O file audio dapat memblokir pemrosesan video, menyebabkan latensi dan frame drop. Pustaka `pyttsx3` digunakan sebagai mekanisme fail-safe (`ensure_alarm_file`) untuk menghasilkan file WAV peringatan secara lokal jika file tersebut tidak ada, menghilangkan ketergantungan pada sumber daya eksternal untuk alarm.

Dokumentasi Insiden: Sistem ini mengintegrasikan lapisan dokumentasi yang objektif. Fitur automatic screenshot (`SAVE_SCREENSHOT_ON_DETECT = True`) mengambil gambar frame insiden saat alarm pertama kali terpicu dengan nama file **`drowsy_{timestamp}.jpg`**, memberikan bukti visual dengan timestamp yang akurat. Selain itu, fitur logging data (`LOG_DETECTIONS = True`) mencatat secara berkelanjutan **EAR_mean** setiap frame, hitungan frame, dan status deteksi ke file CSV. Data log ini memfasilitasi analisis kinerja sistem jangka panjang dan dapat digunakan sebagai set data pelatihan untuk model deep learning yang memerlukan data sekuensial.

3. HASIL DAN PEMBAHASAN

Kinerja dan Efisiensi Pemrosesan

Implementasi menggunakan MediaPipe Face Mesh menunjukkan kinerja pemrosesan yang efisien, yang krusial untuk aplikasi waktu nyata. Meskipun model MediaPipe mengestimasi 468 landmark, arsitektur yang ringan dan dioptimalkan (dengan akselerasi GPU/TPU yang tersedia) memungkinkan sistem mempertahankan FPS tinggi. Kinerja yang cepat ini adalah alasan utama mengapa sistem dapat secara akurat menghitung dan merespons 18 frame penutupan mata dalam waktu **0.6 detik**.

Stabilitas deteksi landmark juga menjadi keunggulan MediaPipe. Bahkan ketika terjadi variasi minor pada pencahayaan atau pergerakan wajah, fitur *refine_landmarks* yang dikombinasikan dengan estimasi 3D meminimalkan jitter pada titik-titik mata, memastikan perhitungan EAR yang konsisten. Keandalan ini merupakan perbaikan signifikan dibandingkan sistem yang bergantung pada detektor wajah non-landmark atau model 2D murni yang lebih sensitif terhadap perubahan lingkungan.

Analisis Kuantitatif dan Validasi Temporal

Pemilihan ambang batas $EAR \leq 0.25$ secara empiris terbukti memadai dalam memicu hitungan frame ketika mata menunjukkan tanda-tanda kelelahan. Pembahasan utama terletak pada validasi temporal 18 frame (Theresia et al., 2021; Thoriq et al., 2024).

Dengan asumsi **30 FPS**, konfigurasi 18 frame secara eksplisit menetapkan ambang batas temporal deteksi pada **0.6 detik**. Penentuan ini penting karena menyediakan justifikasi berbasis fisiologis yang kuat. Dengan durasi kedipan normal berada di bawah **0.3 detik**, sistem secara otomatis memfilter kedipan yang tidak terkait dengan kantuk. Alarm hanya dipicu ketika durasi penutupan mata memenuhi atau melebihi definisi microsleep (lebih dari **0.5 detik**). Logika keputusan temporal ini memastikan bahwa alarm yang dihasilkan benar-benar mencerminkan insiden kelelahan yang signifikan, sehingga meminimalkan *false alarm* (Perdana et al., 2019).

Nilai Sistem Peringatan Multimodal dan Dokumentasi

Kinerja alarm adalah aspek kritis dari sistem DDD. Penggunaan threading untuk memutar alarm suara merupakan elemen desain yang penting. Dengan memisahkan proses audio (I/O) dari *main video loop*, sistem menjamin bahwa deteksi frame tidak terblokir. Ini memastikan bahwa sekalipun alarm berbunyi dalam durasi yang lama,

sistem tetap mampu mendeteksi kapan pengguna membuka mata kembali dan segera mereset *frame_counter* tanpa jeda pemrosesan. Respons non-blocking ini merupakan prasyarat untuk efektivitas sistem real-time.

Fitur dokumentasi otomatis—khususnya automatic screenshot dan pencatatan CSV—memberikan nilai tambah yang signifikan. Tangkapan layar dengan timestamp berfungsi sebagai bukti kejadian yang terverifikasi, berguna untuk analisis insiden setelah penggunaan. Sementara itu, file CSV menyediakan data deret waktu mentah yang memuat tren **EAR_mean** dan akumulasi *frame_counter*. Data ini dapat digunakan untuk analisis lebih lanjut, seperti mengamati peningkatan frekuensi deteksi dalam sesi pemantauan panjang, atau sebagai data berlabel untuk melatih model deep learning di masa depan yang dapat memprediksi kantuk sebelum mencapai ambang batas (Ananta et al., 2025).

Batasan dan Prospek Pengembangan

Batasan inheren sistem berbasis EAR adalah kepekaannya terhadap oklusi, seperti penggunaan kacamata gelap atau tindakan pengguna menutup mata secara sengaja, yang dapat menghasilkan *false positive*. Sistem saat ini juga belum mengintegrasikan indikator kantuk selain dari aspek mata.

Untuk meningkatkan robustness sistem, penelitian lanjutan dapat berfokus pada:

- **Ekstraksi Fitur Multimodal:** Mengintegrasikan EAR dengan Mouth Aspect Ratio (MAR) untuk mendeteksi menguap, serta menganalisis kemiringan kepala untuk mendeteksi *head nodding* (Nizar, 2023). Kombinasi fitur ini akan menghasilkan diagnosis kantuk yang lebih komprehensif (Firdaus et al., 2023).
- **Transisi ke Deep Learning Temporal:** Menerapkan arsitektur seperti CNN-LSTM untuk memproses data sekuensial EAR yang telah direkam, memungkinkan deteksi pola kantuk yang lebih kompleks dibandingkan pendekatan berbasis ambang batas.
- **Porting ke Perangkat Tertanam:** Menguji sistem pada perangkat embedded seperti Raspberry Pi untuk memvalidasi kinerja dan kelayakan implementasi dalam sistem kendaraan atau aplikasi portabel lainnya (Perdana et al., 2019).

4. KESIMPULAN

Sistem deteksi kantuk waktu nyata yang memanfaatkan kerangka kerja MediaPipe

Face Mesh untuk ekstraksi landmark mata dan metrik Eye Aspect Ratio (EAR) telah berhasil dikembangkan dengan fokus pada efisiensi real-time dan respons multimodal. Penggunaan MediaPipe Face Mesh menjamin pemrosesan frame yang cepat dan stabil, sementara parameter kunci $EAR \leq 0.25$ dan 18 frame berturut-turut terbukti efektif serta memiliki justifikasi ilmiah yang kuat dalam membedakan kedipan mata normal dari durasi microsleep berbahaya (> 0.5 detik).

Selain itu, desain peringatan multimodal yang memanfaatkan threading untuk alarm suara memastikan operasi non-blocking, menjaga integritas laju pemrosesan frame. Fitur dokumentasi otomatis, melalui screenshot insiden dan logging data CSV yang rinci, memberikan lapisan verifikasi objektif yang berharga.

Sistem ini menawarkan solusi yang akurat, ringan, dan hemat biaya sebagai alternatif efektif dibandingkan metode deteksi kantuk berbasis sensor fisiologis

DAFTAR PUSTAKA

- Ananta, W. G., Wali, A., Bahari, S., & Safitri, P. H. (2025). *DETEKSI KANTUK MENGGUNAKAN EAR DAN MAR BERBASIS DEEP LEARNING*. 12(5), 8221–8229.
- Asvin, A., Suradi, M., Tri, T., Manguma, F., Alam, S., & Nurul, A. N. (2025). *Deteksi dan Peringatan Jarak Wajah Otomatis Menggunakan MediaPipe dan Computer Vision untuk Kesehatan Pengguna Komputer*. 7, 127–135.
- Daffa, F. M., Pramono, S., & Sukarna, E. (2025). *Sistem monitoring kantuk dan pengiriman lokasi kecelakaan berbasis GSM pada helm*. 5(2), 145–154.
- Firdaus, A., Utamingrum, F., & Widasari, E. R. (2023). *Sistem Pendeteksi Kantuk Pengemudi berbasis Eye Aspect Ratio dan Mouth Opening Ratio menggunakan Algoritme C-LSTM*. 7(2), 927–933.
- Kadi, A. B., Munadi, R., Fitriyanti, N., Eleketro, F. T., Telkom, U., Pendahuluan, I., Vision, A. C., Landmarks, B. F., & Ratio, C. E. A. (2023). *DETEKSI KANTUK PENGEMUDI BUS TRANS METRO BANDUNG DENGAN PENDEKATAN*. 2(2), 60–64.
- Kurniawan, C., Nggiku, U., & Rabi, A. (2023). *Deteksi Kantuk Untuk Keamanan Berkendara Berbasis Pengolahan JEETech*. 48–56.
- Marshella, S. N., Hasanah, R., & Habinuddin, E. (2024). *Prototype alat pendeteksi kantuk*

menggunakan sensor MAX30102 dan kamera dengan metode eye aspect ratio. 4(1), 45–56.

Maslikah, S., Alfita, R., & Ibadillah, A. F. (n.d.). *Sistem Deteksi Kantuk Pada Pengendara Roda Empat Menggunakan Eye Blink Detection. 123–128.*

Nizar, T. N. (2023). *Deteksi Aktivitas Mata , Mulut Dan Kemiringan Kepala Sebagai Fitur Untuk Deteksi Kantuk Pada Pengendara Mobil Detection of eye activity , mouth and head tilt as a feature for detecting drowsiness in car drivers. 12(148), 91–99.*
<https://doi.org/10.34010/komputika.v12i1.9688>

Oleh, D. (2025). *Implementasi sistem deteksi rasa kantuk pada pengemudi kendaraan roda empat menggunakan opencv.*

Perdana, A. H. A. P., Tri, S., & Heri, R. (2019). *Implementasi Sistem Deteksi Mata Kantuk Berdasarkan Facial Landmarks Detection Menggunakan Metode Regression Trees. 1(1), 1–9.*

Samsinar, R., Isyanto, H., Almanda, D., & Amrullah, F. (n.d.). *Sistem Monitoring Mendeteksi Mata Lelah Pada Pengemudi Kendaraan Besar Berbasis Pengolahan Citra. 7(1), 77–84.*

Shalihah, A., Nuryani, N., & Sutomo, D. (2020). *Deteksi Kantuk Menggunakan Jaringan Saraf Tiruan Radial Basis Function dan Particle Swarm Optimization dengan RR Interval Elektrokardiogram. 10(1), 74–80.*

Telecommunication, I., Tengah, J., Makalah, I., & Learning, D. (2023). *Sistem Peringatan Tersepat untuk Pengemudi Mengantuk Embedded Alert System for Drowsy Drivers. 5(1).*

Theresia, N., Pasaribu, B., Halim, T., & Prijono, A. (2021). *EEG signal classification for drowsiness detection using wavelet transform and support vector machine. 10(2), 501–509.* <https://doi.org/10.11591/ijai.v10.i2.pp501-509>

Thoriq, M. F., Ramdhana, M. F., & Rahma, D. N. (2024). *Implementasi Sistem Deteksi Kantuk Secara Real-Time Bagi Pengemudi Menggunakan Metode Eye Aspect Ratio. 70–85.*

Yahya, M. A., Yunanda, A. B., Sidqon, M., Kridoyono, A., Informatika, T., & Timur, J. (2024). *DENGAN NOTIFIKASI LOKASI MELALUI TELEGRAM. 13(3), 229–240.*

Zikri, A. (2025). *Jurnal Sains Informatika Terapan (JSIT). 508–516.*