

Analisis dan Implementasi Metode Viola-Jones dan CNN pada Sistem Deteksi Kantuk Real-Time

Alven Safik Ritonga*¹, Isnaini Muhandhis²

¹²Teknik Informatika, Universitas Wijaya Putra

Jl. Pd. Benowo Indah No.1-3, Babat Jerawat, Kec. Pakal, Surabaya, Jawa Timur 60197

Email : ¹alvensafik@uwp.ac.id, ²isnainimuhandhis@uwp.ac.id

Abstract

Drowsiness detection is a crucial step in enhancing safety and productivity, especially in the transportation and industrial sectors. Drowsiness can lead to fatal accidents, particularly for drivers and industrial workers. This study develops a computer vision-based drowsiness detection system using a combination of the Viola-Jones method and Convolutional Neural Network (CNN). Viola-Jones is used for rapid face and eye detection as an initial step, while CNN is trained to classify eye conditions (open or closed) that indicate drowsiness. The system is designed to operate in real-time and can be accessed via a website. The dataset used consists of 3,232 images, divided into two groups: 1,432 images of focused conditions and 1,800 images of drowsiness conditions, with an 80% training and 20% testing split. Results show that the system achieves 91% accuracy, with precision, recall, and F1-score all reaching 0.91, indicating high performance in classifying drowsiness and focused states. During testing, the training process lasted 22 minutes and 39 seconds for 80 epochs with a learning rate of 0.001. This model provides accurate and responsive drowsiness detection across various lighting conditions and poses, helping reduce accident risk due to drowsiness.

Keywords: computer vision, convolutional neural network, drowsiness detection, eye classification, Viola-Jones.

Abstraksi

Deteksi kantuk merupakan langkah penting dalam meningkatkan keselamatan dan produktivitas, khususnya di sektor transportasi dan industri. Kantuk dapat memicu kecelakaan fatal, terutama bagi pengemudi dan pekerja industri. Penelitian ini mengembangkan sistem deteksi kantuk berbasis visi komputer menggunakan kombinasi metode Viola-Jones dan Convolutional Neural Network (CNN). Viola-Jones digunakan untuk mendeteksi wajah dan mata secara cepat sebagai langkah awal, sementara CNN dilatih untuk mengklasifikasikan kondisi mata (terbuka atau tertutup) yang menunjukkan kantuk. Sistem ini dirancang untuk berjalan secara real-time dan dapat diakses melalui website. Dataset yang digunakan terdiri dari 3.232 gambar, dibagi menjadi dua kelompok: 1.432 gambar kondisi fokus dan 1.800 gambar kondisi kantuk, dengan pembagian data 80% untuk training dan 20% untuk testing. Hasil menunjukkan bahwa sistem berhasil mencapai akurasi 91% dengan precision, recall, dan F1-score sebesar 0,91, mengindikasikan kinerja yang tinggi dalam mengklasifikasikan kondisi kantuk dan fokus. Pada saat pengujian, proses training berlangsung selama 22 menit 39 detik untuk 80 epoch dengan learning rate 0,001. Model ini memberikan deteksi kantuk yang akurat dan responsif dalam berbagai kondisi

pencapaian dan pose, sehingga dapat digunakan untuk meminimalisir risiko kecelakaan akibat kantuk.

Kata Kunci: *convolutional neural network, deteksi kantuk, klasifikasi mata, komputer Vision, Viola-Jones.*

1. PENDAHULUAN

Perkembangan teknologi digital telah membawa berbagai inovasi dalam bidang keselamatan dan kesehatan, khususnya untuk mengatasi permasalahan kantuk yang dapat berdampak buruk pada produktivitas dan meningkatkan risiko kecelakaan, terutama di sektor transportasi dan industri. Kecelakaan akibat kantuk telah menjadi salah satu faktor utama yang menyebabkan kematian dan cedera serius di berbagai sektor, terutama dalam transportasi dan industri. Data dari Organisasi Kesehatan Dunia (WHO) menunjukkan bahwa setiap tahun, sekitar 1,35 juta orang kehilangan nyawa akibat kecelakaan lalu lintas, di mana kantuk menjadi salah satu faktor signifikan [1]. Di Indonesia, Korlantas Polri mencatat bahwa sekitar 10% dari total kecelakaan lalu lintas disebabkan oleh kantuk, yang sering kali berakibat fatal terutama pada kendaraan besar seperti truk dan bus.

Kecelakaan akibat kantuk juga memiliki dampak signifikan di sektor industri, terutama dalam bidang yang membutuhkan tingkat kewaspadaan tinggi, seperti sektor manufaktur, pertambangan, dan konstruksi. Menurut data dari Badan Keselamatan dan Kesehatan Kerja (OSHA) di Amerika Serikat, kelelahan dan kantuk berkontribusi pada lebih dari 13% kecelakaan kerja serius setiap tahunnya. Kecelakaan di lingkungan industri yang disebabkan oleh kantuk dapat berujung pada cedera parah atau bahkan kehilangan nyawa, di samping menyebabkan kerugian ekonomi besar akibat kerusakan peralatan dan penurunan produktivitas [2].

Deteksi kantuk secara *real-time* menghadapi tantangan besar, mengingat kondisi fisik pekerja atau pengemudi yang mengantuk sulit dikenali secara konsisten. Variasi dalam pola tidur, ekspresi wajah, dan kondisi lingkungan yang dinamis semakin mempersulit deteksi [3]. Tantangan ini semakin besar dalam aplikasi *real-time*, yang membutuhkan akurasi dan kecepatan tinggi untuk menjaga keselamatan tanpa mengganggu aktivitas bekerja atau mengemudi.

Penelitian ini bertujuan untuk mengembangkan sistem deteksi kantuk *real-time* dengan memanfaatkan metode Viola-Jones dan Convolutional Neural Network (CNN).

Viola-Jones digunakan untuk mendeteksi wajah dan mata pengemudi atau pekerja dalam kondisi *real-time*, karena kecepatan dan akurasi dalam deteksi objek. Setelah objek terdeteksi, CNN digunakan untuk menganalisis pola kantuk dari ekspresi wajah atau kondisi mata yang tertutup. Dengan mengintegrasikan kedua metode ini, penelitian ini diharapkan mampu menciptakan sistem deteksi kantuk yang andal dan efisien.

Secara spesifik, Viola-Jones berfungsi sebagai metode deteksi awal yang cepat untuk menemukan posisi wajah dan mata dalam video, sementara CNN menangani klasifikasi tingkat kantuk berdasarkan pola mata (misalnya, mata yang tertutup dalam periode tertentu) dan ekspresi wajah lainnya [4]. Kombinasi kedua metode ini diharapkan dapat memberikan hasil deteksi yang akurat dan responsif, bahkan dalam kondisi lingkungan yang bervariasi.

Deteksi objek adalah salah satu tugas utama dalam visi komputer yang melibatkan identifikasi dan klasifikasi objek dalam gambar atau video. Algoritma deteksi objek biasanya menggabungkan teknik-teknik seperti konvolusi untuk ekstraksi fitur dan klasifikasi untuk pengenalan objek.

Metode deteksi objek telah berkembang dari teknik berbasis fitur dan pembelajaran mesin sederhana menjadi teknik canggih berbasis *deep learning* yang mampu mendeteksi objek dengan akurasi dan kecepatan yang luar biasa. Pendekatan-pendekatan seperti Haar-like *features* dan SVM masih digunakan dalam aplikasi tertentu tetapi metode *deep learning* seperti R-CNN, YOLO, dan SSD telah menjadi standar dalam penelitian dan aplikasi praktis karena keunggulan mereka dalam menangani variasi objek dan kondisi gambar yang kompleks [5].

Dalam konteks aplikasi pendeteksi kantuk, pemilihan metode deteksi objek yang tepat sangat penting untuk memastikan kecepatan dan akurasi yang optimal, terutama dalam lingkungan yang dinamis dan *real-time* seperti deteksi kantuk pada pengemudi [6].

Teori Viola-Jones Algoritma Viola-Jones, diperkenalkan oleh Paul Viola dan Michael Jones pada tahun 2001, menjadi terobosan dalam deteksi wajah karena kemampuannya melakukan deteksi secara *real-time* dengan akurasi yang memadai. Algoritma ini mengandalkan beberapa teknik utama, yaitu fitur Haar-like, integral *image*, dan *cascade classifiers*, yang bekerja bersama untuk mendeteksi wajah secara efisien [7].

Penelitian oleh [8] menggunakan metode Viola-Jones untuk deteksi mata dan mulut sebagai indikator kantuk. Meskipun metode ini menunjukkan efisiensi dalam kondisi ideal, penelitian ini juga menggarisbawahi keterbatasan Viola-Jones dalam menghadapi variasi pencahayaan dan pose [8]. Penelitian [9] menggunakan Viola-Jones untuk mendeteksi wajah dan kemudian menggunakan teknik pelacakan kepala untuk mengidentifikasi perubahan posisi yang berhubungan dengan kantuk. Hasilnya menunjukkan bahwa pola gerakan kepala dapat mengindikasikan tingkat kantuk [9].

Penelitian [10] membuat aplikasi deteksi kantuk berbasis *webcam* untuk pengguna komputer. Menggunakan metode Viola-Jones digunakan untuk deteksi wajah dan mata [10]. Sistem ini memantau durasi mata tertutup dan perubahan posisi kepala. Hasil yang diperoleh adalah sebuah sistem yang dapat mendeteksi kantuk dengan akurasi tinggi dalam lingkungan terkendali. Tantangan utama adalah variabilitas pencahayaan dan pose kepala. Mengembangkan sistem peringatan kantuk untuk pengemudi kendaraan. Metode: Menggunakan Viola-Jones untuk deteksi wajah dan mata, serta algoritma tambahan untuk analisis lebih lanjut. Sistem ini memantau waktu penutupan mata dan frekuensi kedipan mata. Hasil: Sistem mampu mendeteksi tanda-tanda kantuk dengan tingkat kesalahan yang rendah. Namun, algoritma memerlukan pencahayaan yang konsisten untuk performa optimal.

CNN merupakan salah satu jenis algoritma *deep learning* yang sangat efektif untuk analisis citra, termasuk dalam mendeteksi tanda-tanda kantuk. CNN memiliki kemampuan untuk mengenali pola visual yang kompleks melalui lapisan konvolusi, *pooling*, dan *fully connected*, yang semuanya memungkinkan ekstraksi fitur otomatis dari data citra. Dalam konteks pendeteksian kantuk, CNN dapat digunakan untuk menganalisis berbagai elemen wajah yang menjadi indikator kantuk, seperti mata tertutup, frekuensi kedipan mata, dan pergerakan kepala.

CNN terdiri dari beberapa lapisan, termasuk lapisan konvolusi, lapisan *pooling*, dan lapisan *fully connected*. Setiap lapisan memiliki peran khusus dalam ekstraksi fitur dari gambar input [11]. Lapisan konvolusi bertanggung jawab untuk mendeteksi fitur-fitur rendah seperti tepi, tekstur, atau warna dari gambar. Dalam konteks deteksi kantuk, lapisan ini dapat mengidentifikasi karakteristik tertentu seperti bentuk mata atau posisi kepala [11]. Lapisan *pooling*: bertujuan untuk mengurangi dimensi gambar secara spasial sambil

tetap mempertahankan informasi penting. Dalam deteksi kantuk, ini berarti ukuran gambar diperkecil tanpa kehilangan fitur utama, sehingga mempercepat proses tanpa mengorbankan akurasi [11]. Lapisan *Fully Connected*, pada lapisan ini, fitur yang telah diekstraksi diklasifikasikan berdasarkan pola yang diidentifikasi oleh CNN. Misalnya, CNN dapat menentukan apakah kondisi kantuk terdeteksi berdasarkan pola dari data gambar, seperti mata tertutup dalam waktu lama atau kedipan mata yang cepat [11].

Penelitian menggunakan kombinasi Viola-Jones dan CNN untuk deteksi masker pada wajah, menemukan bahwa Viola-Jones secara efektif dapat mendeteksi wajah dan fitur wajah lainnya dengan cepat, tetapi tidak mampu melakukan klasifikasi detail seperti mendeteksi keberadaan masker [12]. Untuk mengatasi keterbatasan ini, CNN ditambahkan untuk melakukan klasifikasi, yang menghasilkan akurasi yang tinggi pada pengenalan objek khusus, seperti masker. Hasil ini menunjukkan kekuatan pendekatan gabungan Viola-Jones untuk deteksi awal dan CNN untuk klasifikasi spesifik pada objek wajah [12].

Penelitian [13] mengembangkan sistem deteksi kantuk *real-time* berbasis CNN, yang menunjukkan akurasi tinggi tetapi memiliki keterbatasan waktu proses yang lebih lambat saat seluruh wajah diproses oleh CNN. Oleh karena itu, Viola-Jones digunakan untuk mendeteksi area wajah lebih dahulu, yang kemudian memfokuskan proses klasifikasi CNN pada area mata. Strategi ini mempercepat proses deteksi, membuatnya lebih efisien untuk aplikasi *real-time*. Studi ini menggarisbawahi pentingnya segmentasi awal sebelum klasifikasi, serupa dengan penelitian kami yang menggunakan Viola-Jones untuk mendeteksi area wajah dan mata sebelum melibatkan CNN [13].

Penelitian tentang sistem pengenalan keadaan mata dengan CNN menggunakan transfer *learning* juga menyebutkan bahwa penggunaan Viola-Jones untuk mendeteksi area mata sebelum klasifikasi CNN membantu meningkatkan akurasi secara keseluruhan. Pendekatan ini sangat efisien dalam mengidentifikasi mata tertutup atau terbuka dengan latensi yang rendah. Namun, penelitian ini difokuskan pada gambar statis, sehingga belum optimal untuk aplikasi yang membutuhkan pemrosesan *real-time* pada gambar bergerak [14].

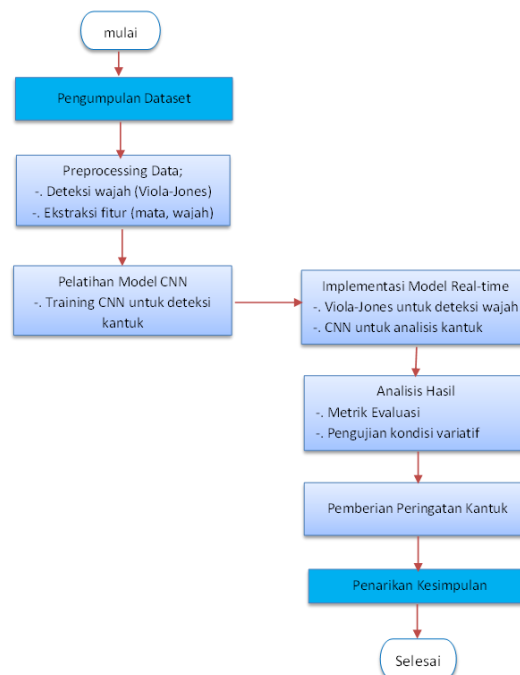
Kombinasi Viola-Jones dan CNN telah diterapkan dalam berbagai penelitian, namun masih terdapat perbedaan terutama pada efisiensi pemrosesan *real-time* serta ketahanan akurasi dalam kondisi lingkungan yang beragam. Penelitian ini mengisi melengkap hal

tersebut dengan memaksimalkan kekuatan Viola-Jones dalam mendeteksi area wajah dan mata sebagai input yang terbatas untuk CNN. Algoritma CNN dapat fokus hanya pada bagian mata yang sudah terdeteksi, yang mengoptimalkan proses pemrosesan dalam konteks *real-time* serta meningkatkan akurasi dalam berbagai posisi wajah dan kondisi lingkungan.

Penelitian ini diharapkan mampu memberikan kontribusi signifikan pada pengembangan teknologi deteksi kantuk yang dapat diterapkan di dunia nyata, baik dalam sektor transportasi maupun industri di Indonesia, untuk meningkatkan keselamatan kerja dan mengurangi risiko kecelakaan.

2. METODE PENELITIAN

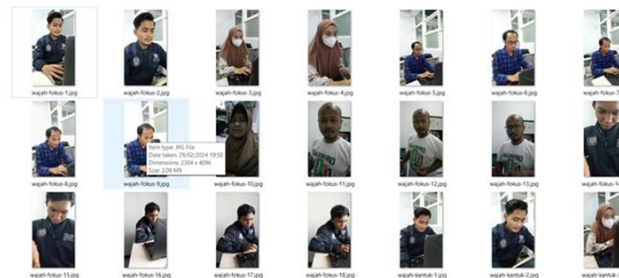
Penelitian ini menggunakan pendekatan eksperimental dengan fokus pada pengembangan dan evaluasi sistem deteksi kantuk berbasis visi komputer. Jenis penelitian ini bertujuan untuk menguji efektivitas kombinasi metode Viola-Jones dan Convolutional Neural Network (CNN) dalam mendeteksi kondisi kantuk secara *real-time*. Uji coba dilakukan pada sistem deteksi kantuk yang diimplementasikan untuk mengukur akurasi dan kinerja model dalam kondisi dunia nyata.



Gambar 1. Flowchart Penelitian

2.1. Data yang Digunakan

Jenis data yang digunakan dalam penelitian ini adalah data gambar atau video yang diambil dari kondisi wajah manusia, khususnya fokus pada mata, yang merupakan indikator utama untuk mendeteksi kantuk. Data ini mencakup berbagai kondisi pencahayaan, pose wajah, dan ekspresi wajah (seperti mata tertutup atau terbuka). Sumber datanya ada 2, yaitu, sumber internal yang didapatkan melalui proses fotografi menggunakan *smartphone* peneliti, seperti yang di gambar 2, dan data sumber eksternal didapatkan melalui situs *universe.roboflow.com* dan *CEW dataset*, seperti di gambar 3.



Gambar 2. Contoh Data Internal



Gambar 3. Contoh Data Eksternal

2.2. Rincian Informasi Dataset

Dataset yang digunakan dalam penelitian ini terdiri dari 3.232 gambar yang terbagi menjadi dua kelompok, 1.432 gambar kondisi fokus (mata terbuka dan ekspresi wajah yang menunjukkan kewaspadaan) dan 1.800 gambar kondisi kantuk (mata tertutup atau kedipan yang menunjukkan tanda-tanda kelelahan). *Dataset* ini mencakup variasi kondisi pencahayaan dan posisi wajah untuk meningkatkan keakuratan model dalam mendeteksi *drowsiness* pada lingkungan yang dinamis dan *real-time*.

2.3. Pembagian Data Latih dan Data Uji

Dataset dibagi menjadi dua untuk tujuan pelatihan dan pengujian model: data latih (*Training Data*), 80% dari total *dataset* (sekitar 2.586 gambar) digunakan untuk melatih model CNN. Data latih ini digunakan untuk mengajarkan model mengenali pola mata terbuka dan tertutup, serta berbagai ekspresi wajah yang mengindikasikan tingkat kewaspadaan atau kantuk. Data uji (*Testing Data*), 20% dari total *dataset* (sekitar 646 gambar) digunakan untuk menguji performa model setelah proses pelatihan. Data uji ini membantu mengevaluasi akurasi, *precision*, *recall*, dan F1-score dari sistem deteksi kantuk. Metrik evaluasi yang disebutkan di atas, yaitu Akurasi, *Precision*, *Recall*, dan F1-Score. Berikut adalah penjelasan detail dan rumus untuk masing-masing *metric* [15]:

a. Akurasi

Akurasi menunjukkan seberapa tepat model dalam melakukan prediksi secara keseluruhan ditunjukkan pada rumus 1 berikut.

$$Akurasi = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

Di mana TP (*True Positives*) adalah prediksi benar untuk kelas positif (kantuk). TN (*True Negatives*) adalah prediksi benar untuk kelas negatif (fokus). FP (*False Positives*) adalah prediksi salah untuk kelas positif (kantuk), seharusnya negatif (fokus). FN (*False Negatives*) adalah prediksi salah untuk kelas negatif (fokus), seharusnya positif (kantuk).

b. *Precision*

Precision digunakan untuk melihat ketepatan model dalam mengidentifikasi kondisi "kantuk" tanpa memprediksi secara salah pada rumus 2 berikut.

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

c. *Recall*

Metrik ini menunjukkan sensitivitas model atau kemampuannya untuk mendeteksi kondisi "kantuk" yang benar-benar terjadi, pada rumus 3 berikut.

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

d. F1-Score

F1-Score digunakan saat keseimbangan antara *precision* dan *recall* penting untuk performa model secara keseluruhan, pada rumus 4 berikut.

$$F1 - score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (4)$$

2.4. Preprocessing Data

Pada tahap *preprocessing*, beberapa langkah dilakukan untuk mempersiapkan data agar dapat diproses lebih lanjut oleh model. Langkah-langkah *preprocessing* yang digunakan meliputi:

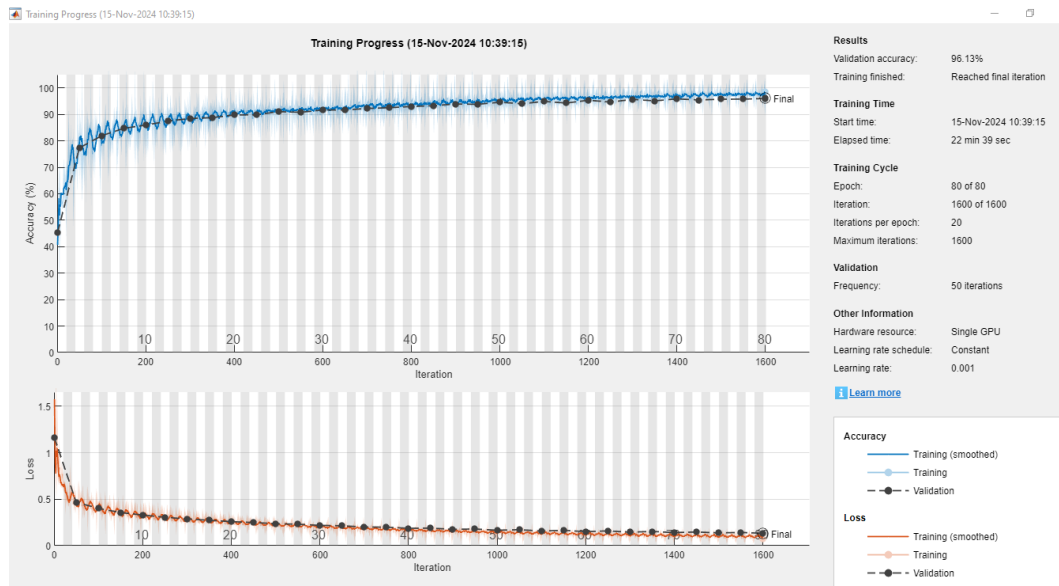
- a. Deteksi Wajah Menggunakan Viola-Jones: Algoritma Viola-Jones pertama-tama digunakan untuk mendeteksi wajah dalam setiap gambar. Kemudian, posisi wajah yang terdeteksi digunakan untuk mengekstraksi Region of Interest (ROI), yang lebih lanjut dianalisis oleh model CNN.
- b. Normalisasi Gambar: Setelah deteksi wajah dan mata, gambar diubah ke format standar dengan ukuran tetap dan normalisasi intensitas piksel untuk memastikan konsistensi input gambar pada model CNN.
- c. Penyaringan dan Pemrosesan Lanjutan: Menggunakan teknik *Image Augmentation* untuk memperbesar variasi *dataset* dengan rotasi, perubahan pencahayaan, dan *flipping* gambar. Hal ini bertujuan untuk meningkatkan keberagaman data dan mencegah model *overfitting*.
- d. Konversi ke *Grayscale* (Opsional): Jika diperlukan, gambar dapat dikonversi ke format *grayscale* untuk mengurangi kompleksitas komputasi dan fokus hanya pada fitur geometris wajah, seperti bentuk mata dan posisi.

3. HASIL DAN PEMBAHASAN

3.1. Hasil Training Model

Grafik pada gambar 4 di bawah ini menunjukkan perkembangan akurasi model terhadap iterasi pelatihan. Dari grafik, Akurasi pelatihan dan validasi meningkat pesat pada awal pelatihan (sekitar 200 iterasi pertama), yang menunjukkan bahwa model dengan cepat belajar mengenali pola dalam data. Model mempelajari fitur-fitur sederhana dan penting dalam data, sehingga *error* atau *loss* menurun drastis di awal. Setelah itu, akurasi cenderung stabil mendekati 96.13% untuk validasi dan sekitar 98% untuk pelatihan, menunjukkan bahwa model telah mencapai titik konvergensi. Akurasi validasi 96,13% menunjukkan performa yang baik dan tingkat kesalahan yang rendah pada data yang tidak

dilatih. Waktu *training* 22 menit 39 detik, menunjukkan bahwa pelatihan diselesaikan dalam waktu yang cukup efisien dengan menggunakan GPU tunggal, *epoch* maksimal 80 dan *learning rate* sebesar 0.001.



Gambar 4. Grafik proses *training*

Hasil ini dapat dibandingkan dengan penelitian-penelitian terdahulu yang juga menggunakan kombinasi arsitektur Viola-Jones dan CNN untuk klasifikasi atau deteksi objek. Sebagian besar penelitian sebelumnya menunjukkan bahwa akurasi validasi pada model CNN biasanya berada dalam rentang 85-95%, tergantung pada kompleksitas dataset dan parameter pelatihan. Dalam penelitian ini, akurasi validasi yang dicapai (96.13%) sedikit lebih tinggi, yang menunjukkan bahwa metode yang digunakan berhasil meningkatkan performa model secara keseluruhan.

Hasil proses *training* di atas, menunjukkan hasil yang optimal karena beberapa hal, yaitu; model mencapai konvergensi yang baik tanpa *overfitting*, parameter pelatihan dan arsitektur model sudah disesuaikan dengan baik terhadap *dataset* yang digunakan.

3.2. Pengujian Model

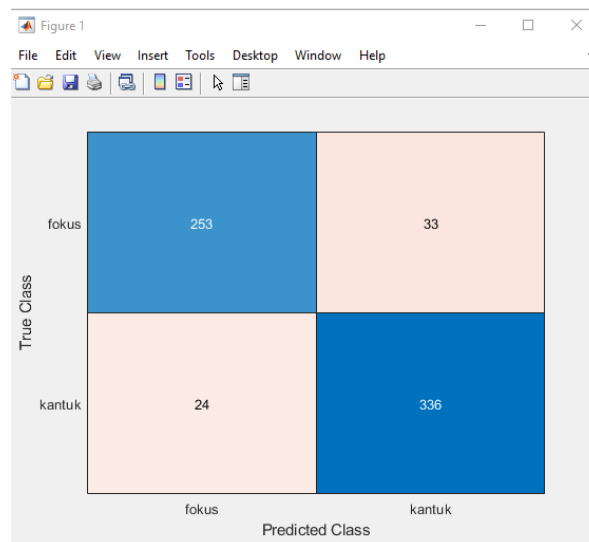
Tabel 1 berikut ini adalah hasil proses *training* menggunakan jumlah maksimum *epoch* yang berbeda. Berdasarkan tabel ini, *epoch* maksimum 80 adalah titik optimal yang memberikan akurasi tertinggi 91.17% dengan waktu pelatihan yang tidak terlalu lama, melampaui jumlah *epoch* ini hingga maksimum 100, sehingga model justru sedikit menurun

dalam performa. Ini disebabkan *overfitting*, di mana model terlalu menyesuaikan diri pada data pelatihan dan mulai kehilangan generalisasi pada data validasi.

Tabel 1. Hasil proses *training* (*learning rate*=0,001)

No	MaksEpoch	Hasil Akurasi	Waktu Training (Detik)
1	25	87,15%	6 menit 58 detik
2	40	87,77%	11 menit 16 detik
3	60	86,69%	15 Menit 11 detik
4	80	91,17%	22 menit 39 detik
5	100	90,09%	26 Menit 22 detik

Proses *training* model dengan jumlah *epoch* maksimal 80, *learning* rata 0,001 dan menggunakan GPU *single*, diperoleh nilai-nilai metrik berikut ini; nilai metrik *precision* sekitar 0,91, menunjukkan bahwa model memiliki ketepatan yang tinggi dalam mengklasifikasikan kondisi "kantuk" (atau "mata tertutup") dengan benar. Nilai metrik *recall* yang menunjukkan performa model sekitar 0,91, menunjukkan bahwa model cukup baik dalam menemukan kondisi "kantuk" yang sebenarnya terjadi dalam data uji. Nilai metrik F1-score sebesar 0,91 hampir mendekati 1. Ini berarti model berhasil dalam mengidentifikasi kondisi kantuk (mata tertutup) dan kondisi fokus (mata terbuka) dengan cukup akurat dan konsisten. Nilai-nilai metrik tersebut di atas bisa juga di tampilkan dalam bentuk *confusion* matriks yang dapat dilihat pada gambar 5 berikut ini.

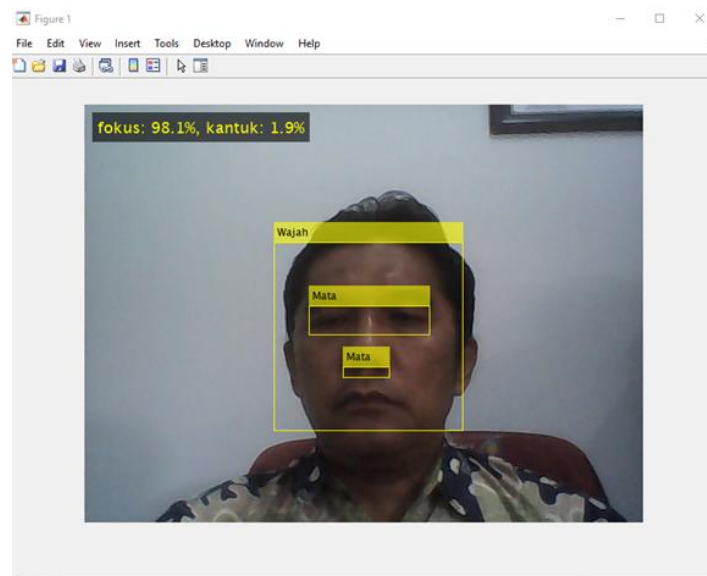


Gambar 5. *Confusion* matriks hasil *training*

Dari *confusion* matriks pada gambar yang Anda bagikan, kita dapat menghitung beberapa metrik evaluasi untuk melihat performa model dalam mendeteksi kondisi "fokus"

dan "kantuk." Penjelasan *Confusion* matriks: *True Positives* (TP): 253 - Model memprediksi "kantuk" dengan benar; *False Positives* (FP): 33 - Model memprediksi "kantuk" tetapi seharusnya "fokus"; *True Negatives* (TN): 336 - Model memprediksi "fokus" dengan benar; *False Negatives* (FN): 24 - Model memprediksi "fokus" tetapi seharusnya "kantuk."

Dari *confusion* matriks ini, bisa menyimpulkan bahwa model memiliki performa yang baik dalam mendeteksi kantuk dengan *precision*, *recall*, dan F1-score yang tinggi. Meskipun model memiliki kinerja yang cukup baik secara keseluruhan, tetapi hasil *confusion* matriks menunjukkan bahwa terdapat beberapa kesalahan prediksi antara kelas "fokus" dan "kantuk". Hal ini kemungkinan disebabkan oleh faktor seperti kemiripan antara data "fokus" dan "kantuk", *noise* dalam data, atau parameter model yang belum sepenuhnya optimal.



Gambar 6. Tampilan Webcam Deteksi Viola-Jones

Gambar 6 di atas menunjukkan hasil deteksi dan klasifikasi kondisi seseorang (fokus atau kantuk) menggunakan model pengenalan wajah dan mata. Prediksi kondisi "fokus" yang sangat tinggi 98,1% dan kantuk 1,9% menunjukkan bahwa model ini memiliki kepercayaan yang cukup tinggi terhadap prediksi fokus dalam gambar tersebut. Hal ini bisa disebabkan beberapa hal, yaitu; keadaan mata yang terbuka optimal, posisi kepala yang tegak, dan pencahayaan yang baik.

4. KESIMPULAN

Berdasarkan analisis dan hasil yang diperoleh pada sebelumnya bisa diambil kesimpulan, sebagai berikut: Penelitian menghasilkan model yang mampu mendeteksi objek dengan akurasi validasi tinggi, yaitu 96.13%, menggunakan kombinasi metode Viola-Jones dan CNN. Hasil ini menunjukkan bahwa model yang dibangun memiliki performa yang baik dalam mengenali pola dan data baru tanpa tanda-tanda *overfitting*; Model ini memiliki potensi aplikasi dalam berbagai bidang, seperti sistem keamanan kendaraan untuk mendeteksi objek di sekitar kendaraan atau memonitor kondisi pengemudi, serta dalam kesehatan kerja untuk mendeteksi postur atau aktivitas pekerja dalam rangka meningkatkan keselamatan kerja, dengan akurasi tinggi; Keterbatasan pada penelitian ini adalah, penggunaan *dataset* yang terbatas serta kurangnya pengujian dalam kondisi nyata. Pengujian lebih lanjut dengan data yang lebih beragam dan dalam lingkungan operasional sebenarnya diperlukan untuk mengukur performa model secara lebih komprehensif.

5. SARAN

Untuk pengembangan lebih lanjut, disarankan beberapa langkah yang dapat diambil dalam penelitian lanjutan, seperti untuk meningkatkan akurasi dan efisiensi, dapat mengeksplorasi kombinasi algoritma lain, seperti mengintegrasikan metode deep learning yang lebih maju, misalnya YOLO atau Faster R-CNN.

DAFTAR PUSTAKA

- [1] World Health Organization. (2018). Global Status Report on Road Safety 2018. Geneva: WHO. URL: who.int/publications.
- [2] John A. Caldwell, J. Lynn Caldwell, Lauren A., 2019. Thompson, Harris R. Lieberman, Fatigue and its management in the workplace, Neuroscience & Biobehavioral Reviews, Volume 96.
- [3] Al-Quraishi, M.S., et al., 2024. Technologies for detecting and monitoring drivers' states: A systematic review, Heliyon, vol 10.
- [4] Subekti, A.R.M., Rahajoe, D.A., dan Mandyartha, P.E., 2024. Klasifikasi Wajah Kantuk Menggunakan Parameter Wajah Dengan Algoritma Longshort Term Memory, Jurnal Informatika dan Sistem Informasi (JIFoSI), vol. 5, No. 2.
- [5] Yedidiya, N. et al., 2023, Perbandingan Metode Yolo Dan Fast R-Cnn Dalam Sistem Deteksi Pengenalan Kendaraan, JRIIN J. Ris. Inform. dan Inov., vol 1, hal 431–436.

- [6] Albadawi, Y., AlRedhaei, A. & Tahruri, M., 2023, Real-Time Machine Learning-Based Driver Drowsiness Detection Using Visual Features, *J. Imaging*, vol 9.
- [7] Sari, P.I., Ramadhani, F., Satria, A., dan Apdilah, D., 2023, Implementasi Pengolahan Citra Digital dalam Pengenalan Wajah menggunakan Algoritma PCA dan Viola Jones, *Hello World; Jurnal Ilmu Komputer*, vol.02, no. 03.
- [8] Hanafie, A., Husain, H.A., Putri, R.R., dan Kemkelo, H., 2023, Aplikasi Ekstraksi Wajah Menggunakan Algoritma Viola Jones, *ILTEK: Jurnal Teknologi*, vol. 18, no. 02.
- [9] Wati, V., Yuliana, Setyowati, N., Y., dan Qulub, M., 2023, Deteksi Wajah Menggunakan Algoritma Viola Jones Berbasis Android, *Teknimedia*, vol.04, no. 01.
- [10] Imanuddin, I., Alhadi, F., Oktafian, R. & Ihsan, A., 2019, Deteksi Mata Mengantuk pada Pengemudi Mobil Menggunakan Metode Viola Jones, *MATRIK J. Manajemen, Tek. Inform. dan Rekayasa Komput*, vol 18, hal 321–329.
- [11] Krichen, M., 2023, Convolutional Neural Networks: A Survey, *Computers*, vol. 12, no. 08, 151. <https://doi.org/10.3390/computers12080151>
- [12] Hermawati, A.F., Zai, S.R., 2021, Sistem Deteksi Pemakaian Masker Menggunakan Metode Viola-Jones dan Convolutional Neural Networks (CNN), *Konferensi Nasional Ilmu Komputer (KONIK)*, June 12.
- [13] Pushkar, P., Khandare, R., Prasad, Y., dan Kumar, V., 2023, Real Time Drowsiness Detection System Using CNN, *International Journal for Research*, vol 11, hal 1487-1490.
- [14] Kayadibi, I., Güraksin, G.E., Ergün, U. et al., 2022, An Eye State Recognition System Using Transfer Learning: AlexNet-Based Deep Convolutional Neural Network, *International Journal of Computational Intelligence Systems*, vol. 15, 49. <https://doi.org/10.1007/s44196-022-00108-2>
- [15] Florez, R., Palomino-Quispe, F., Coaquira-Castillo, R.J., Herrera-Levano, J.C., Paixão, T., dan Alvarez, A.B. A., 2023, CNN-Based Approach for Driver Drowsiness Detection by Real-Time Eye State Identification, *Applied Sciences*, vol. 13, 7849. <https://doi.org/10.3390/app13137849>