

Penerapan MobileNet untuk Klasifikasi Kondisi Wajah Mengantuk pada Citra Wajah

Ghefin Indra ¹, Hazna Hamida Saputri ², Najwa Aulia Aziza ³

Telkom University Bandung, Indonesia

¹ ghefinindra@student.telkomuniversity.ac.id

² haznahamida@student.telkomuniversity.ac.id

³ najwaauliaaziza@student.telkomuniversity.ac.id

Received on 01-06-2025, revised on 06-06-2025, accepted on 30-06-2025

Abstract

Kondisi mengantuk saat berkendara merupakan salah satu penyebab utama kecelakaan lalu lintas. Penelitian ini bertujuan untuk mengembangkan model klasifikasi berbasis *Convolutional Neural Network* (CNN) guna mendeteksi kondisi mengantuk dari citra wajah pengemudi secara otomatis. Dataset yang digunakan adalah *Driver Drowsiness* Dataset yang terdiri dari lebih dari 41.000 gambar wajah berlabel "mengantuk" dan "tidak mengantuk". Model dikembangkan menggunakan arsitektur MobileNetV2 dengan pendekatan *transfer learning* dan teknik data augmentation untuk mengatasi tantangan seperti variasi ekspresi, pencahayaan, dan latar belakang. Proses pelatihan melibatkan normalisasi citra. Evaluasi performa dilakukan dengan menggunakan metrik akurasi, presisi, *recall*, dan F1-score. Hasil dari penelitian ini menunjukkan bahwa pendekatan CNN mampu mengenali kondisi mengantuk secara efektif, sehingga berpotensi diimplementasikan dalam sistem deteksi dini pada kendaraan untuk meningkatkan keselamatan berkendara.

Keywords: CNN, Kantuk, Mobilenetv2, Deteksi Wajah, *Computer Vision*, *Image Classification*.

I. PENDAHULUAN

Kondisi mengantuk merupakan salah satu faktor utama penyebab kecelakaan, terutama dalam aktivitas berkendara. Berdasarkan data dari *National Sleep Foundation*, sekitar 20% pengemudi merasa mengantuk saat mengemudi, yang berpotensi meningkatkan risiko kecelakaan secara signifikan [1]. Secara fisiologis, rasa kantuk dapat dikenali melalui perubahan ekspresi wajah seperti menutup mata yang lebih lama dari biasanya (*microsleep*) dan menguap. Deteksi otomatis terhadap kondisi ini diperlukan dalam upaya pencegahan terhadap risiko terjadinya kecelakaan. Penelitian ini bertujuan untuk membangun model klasifikasi yang mampu membedakan antara kondisi wajah mengantuk dan tidak mengantuk menggunakan pendekatan *Convolutional Neural Network* (CNN). Karena kemampuannya dalam menganalisis pola visual pada gambar secara otomatis, CNN digunakan tanpa perlu tahap ekstraksi ciri secara terpisah seperti pada metode konvensional. Model ini dilatih menggunakan kumpulan gambar wajah dengan beragam ekspresi yang merepresentasikan kondisi mengantuk dan tidak mengantuk, sehingga mampu mengenali pola secara langsung dari citra yang tersedia. Tantangan utama dalam mendeteksi kondisi mengantuk adalah adanya variasi ekspresi wajah antar individu, serta perubahan pencahayaan dan kondisi lingkungan yang dapat mempengaruhi deteksi. *Driver Drowsiness* Dataset. Dataset yang digunakan dalam penelitian ini mencakup gambar wajah dengan ekspresi dengan indikator mengantuk seperti mata yang mulai tertutup dan mata yang sudah tertutup. Meskipun dataset ini

menyediakan variasi ekspresi wajah yang relevan, tantangan terkait variasi antar individu dan kondisi lingkungan seperti pencahayaan yang berbeda tetap menjadi faktor yang harus diatasi. Oleh karena itu, penelitian ini juga akan mempertimbangkan penggunaan teknik data *augmentation* dan *preprocessing* untuk meningkatkan keberagaman data dan mengatasi perubahan pencahayaan dan latar belakang, sehingga dapat meningkatkan akurasi model dalam mendeteksi kondisi mengantuk pada wajah. Metodologi yang digunakan dalam penelitian ini meliputi pengumpulan dan pemrosesan data citra wajah, pembangunan arsitektur CNN, pelatihan model, serta evaluasi performa menggunakan metrik seperti akurasi, presisi, dan *recall*. Evaluasi ini diharapkan dapat memberikan gambaran tentang efektivitas model dalam mengenali kondisi mengantuk pada wajah pengemudi, yang dapat menjadi langkah awal dalam pengembangan sistem deteksi mengantuk berbasis komputer. Proses dalam penelitian ini meliputi pengumpulan dan *preprocessing* data citra wajah, pembangunan arsitektur CNN, pelatihan model, serta evaluasi performa menggunakan metrik seperti akurasi, presisi, dan *recall*. Hasil evaluasi diharapkan dapat mengukur efektivitas model dalam mengenali indikator visual kondisi mengantuk pada wajah.

II. TINJAUAN PUSTAKA

A. Penelitian Terkait

Deteksi kondisi mengantuk berdasarkan citra wajah telah menjadi area penelitian yang berkembang pesat dalam upaya meningkatkan keselamatan berkendara. Pendekatan berbasis *Convolutional Neural Network* (CNN) menunjukkan dominasi dalam bidang ini karena kemampuannya mengekstraksi fitur visual secara otomatis tanpa memerlukan *feature engineering* manual. Arsitektur CNN yang kuat seperti ResNet50V2 telah menunjukkan hasil yang sangat menjanjikan, dengan sistem yang dikembangkan mampu mengidentifikasi kondisi mata terbuka dan tertutup sebagai indikator utama kelelahan pengemudi. Dalam implementasinya, ekstraksi wajah dan area mata dilakukan menggunakan *MediaPipe*, yang kemudian dianalisis oleh model CNN dan menghasilkan akurasi sebesar 99,71% serta menunjukkan ketahanan sistem terhadap variasi pencahayaan [2].

Keberhasilan implementasi CNN pada perangkat dengan sumber daya terbatas juga telah dibuktikan melalui pengembangan model deteksi kantuk berbasis CNN untuk aplikasi mobile. Model ini dilatih menggunakan data wajah dalam berbagai kondisi seperti penggunaan kacamata dan pencahayaan siang atau malam, dengan bantuan teknik *facial landmark detection*. Hasilnya menunjukkan akurasi rata-rata sebesar 83,33% serta memiliki ukuran dan kebutuhan komputasi yang efisien, sehingga cocok diimplementasikan pada perangkat mobile dengan sumber daya terbatas [3]. Sejalan dengan pendekatan efisiensi tersebut, framework MobileNet-SSD dalam arsitektur CNN juga dikembangkan khusus untuk mendeteksi dan melokalisasi kondisi mata terbuka dan tertutup secara *real-time* melalui *video stream*, dengan hasil pengujian menunjukkan nilai *Mean Average Precision* (mAP) sebesar 0.837 dan terbukti dapat diimplementasikan pada perangkat seperti Raspberry Pi 3 atau *smartphone* Android [4].

Untuk meningkatkan *robustness* dan akurasi deteksi, pendekatan *ensemble* yang menggabungkan beberapa model CNN telah menunjukkan efektivitas yang signifikan. Kombinasi model seperti AlexNet, VGG-FaceNet, FlowImageNet, dan ResNet digunakan untuk mengekstraksi fitur dari ekspresi wajah, gerakan tangan, dan elemen lingkungan. Data wajah diklasifikasikan menjadi empat kondisi: tidak mengantuk, mengantuk dengan kedipanmata, menguap, dan mengganggu, dengan sistem mencapai akurasi hingga 85% yang menunjukkan efektivitas pendekatan *ensemble* dalam mendeteksi kondisi mengantuk secara lebih *robust* di berbagai situasi [5].

Pengembangan lebih lanjut dalam arsitektur CNN mencakup pendekatan *stacked deep* CNN yang fokus pada analisis kondisi mata pengemudi. Dalam metode ini, wajah dideteksi menggunakan algoritma Viola-Jones, kemudian area mata diekstraksi dan dianalisis dengan *stacked deep* CNN. Model ini mampu mencapai akurasi hingga 96,42%, melampaui performa dari CNN konvensional, serta berpotensi diterapkan secara efisien pada perangkat mobile dan sistem tertanam [6].

Pendekatan serupa dalam klasifikasi mata terbuka dan tertutup juga dikembangkan menggunakan *Deep Convolutional Neural Networks* dengan memanfaatkan kamera cahaya tampak, yang mampu mendeteksi kondisi mata dalam berbagai skenario pencahayaan dan menghasilkan akurasi deteksi lebih dari 85%, dengan arsitektur *deep residual CNN* memberikan hasil yang lebih baik dibandingkan model CNN biasa [7].

Teknik *transfer learning* telah menjadi strategi penting dalam meningkatkan efisiensi pengembangan sistem deteksi kantuk. Pendekatan berbasis CNN dan *transfer learning* yang memanfaatkan berbagai dataset wajah mampu mengenali tanda-tanda kelelahan seperti mata tertutup dan ekspresi menguap secara *real-time*, dengan hasil eksperimen menunjukkan akurasi tinggi berkisar antara 90% hingga 99,86% [8]. Efisiensi arsitektur jaringan juga ditingkatkan melalui implementasi *transfer learning* pada tiga arsitektur CNN berbeda, yaitu *Fully Designed Neural Network* (FD-NN) serta dua model berbasis VGG16 dan VGG19, yang dioptimalkan untuk mengenali kondisi mata tertutup dan menghasilkan akurasi lebih dari 98% dengan performa *real-time* yang baik [9].

Integrasi informasi temporal dalam deteksi kantuk dilakukan melalui pendekatan hibrida yang menggabungkan CNN dan Long Short-Term Memory (LSTM). Dalam arsitektur ini, CNN bertugas mengekstraksi fitur spasial dari wajah, sedangkan LSTM menganalisis fitur temporal dari data video. Sistem yang diuji menggunakan dataset pengemudi mengantuk dari kompetisi ACCV 2016 menunjukkan performa yang lebih baik dibandingkan metode sebelumnya, khususnya dalam menghadapi variasi pencahayaan dan ekspresi wajah [10].

Pengembangan lebih lanjut dalam pendekatan temporal mencakup sistem deteksi kantuk pengemudi berbasis CNN dan Bidirectional Long Short-Term Memory (BiLSTM) yang menggabungkan analisis ekspresi wajah dan gerakan mata. Dengan menggunakan dataset gambar mata terbuka dan tertutup, sistem ini mampu mendeteksi kondisi mengantuk secara *real-time* dan mencapai akurasi lebih dari 95%. Pendekatan ini menunjukkan efektivitas dalam mengklasifikasikan kondisi mengantuk dan dapat diimplementasikan pada perangkat dengan sumber daya terbatas [11]. Model hibrida LSTM-CNN juga telah diusulkan untuk mendeteksi kondisi mengantuk pengemudi dengan menganalisis sinyal electroencephalogram (EEG), yang dapat mengklasifikasikan tiga kondisi kesadaran yaitu terjaga, tidur, dan mengantuk. Hasil penelitian menunjukkan bahwa model LSTM-CNN dengan jendela input 4 detik memberikan kinerja terbaik, dengan akurasi mencapai 95,6% dalam klasifikasi dua kelas dan 86% dalam klasifikasi tiga kelas [12].

Diversifikasi dalam algoritma deteksi wajah juga menjadi fokus penting untuk meningkatkan robustness sistem deteksi kantuk. Penelitian yang memanfaatkan *deep learning* untuk menganalisis ekspresi wajah dan kondisi mata pengemudi menggunakan beberapa algoritma deteksi wajah, yaitu Viola-Jones, DLib, dan YOLOv3, yang memungkinkan deteksi kondisi pengemudi dalam situasi pencahayaan rendah. Pendekatan ini mampu mengklasifikasikan pengemudi sebagai *drowsy* atau *non-drowsy* dengan akurasi tinggi serta dapat diimplementasikan dalam aplikasi *real-time* [13]. Eksplorasi lebih luas dalam aspek deteksi juga mencakup pendekatan yang menggabungkan analisis ekspresi wajah dengan data perilaku berkendara seperti akselerasi dan kecepatan untuk menilai tingkat kelelahan, yang mampu mendeteksi emosi serta perilaku agresif saat mengemudi dengan akurasi lebih dari 90% [14].

Selain pendekatan berbasis citra, integrasi sinyal fisiologis seperti electroencephalogram (EEG) dengan CNN juga menunjukkan potensi yang besar. Sistem berbasis EEG yang terdiri dari tiga tahap utama (pengumpulan sinyal EEG, proses deteksi kantuk, dan implementasi strategi peringatan dini) menggunakan modul Inception dan arsitektur AlexNet yang dimodifikasi berhasil mencapai akurasi klasifikasi hingga 95,59% [15]. Pendekatan serupa menggunakan sinyal EEG satu saluran dari perangkat wearable dengan analisis cross-validation terhadap dataset publik menunjukkan akurasi lebih dari 90% dan mengungguli pendekatan tradisional yang menggunakan fitur manual [16].

B. Dasar Teori

1) Drowsiness Detection

Deteksi kantuk (*drowsiness detection*) adalah proses identifikasi kondisi mengantuk pada individu, yang dapat mengakibatkan penurunan kewaspadaan dan respons, serta meningkatkan risiko kecelakaan, terutama dalam konteks berkendara. Menurut Nasri dkk., deteksi kantuk dapat dilakukan

dengan menganalisis sinyal fisiologis, fitur wajah, dan pola mengemudi, yang masing-masing memiliki kelebihan dan keterbatasan tersendiri [17].

2) *Convolutional Neural Network (CNN)*

Convolutional Neural Network (CNN) adalah algoritma deep learning yang banyak digunakan dalam pengolahan citra. CNN memanfaatkan lapisan konvolusi untuk mengekstraksi fitur dari gambar, dan lapisan-lapisan selanjutnya, seperti lapisan aktivasi dan pooling, untuk meningkatkan kemampuan model dalam mengidentifikasi pola visual. CNN banyak digunakan dalam berbagai aplikasi, termasuk dalam deteksi *drowsiness* (kantuk) dengan menganalisis ekspresi wajah dan gerakan tubuh. CNN memiliki kemampuan dalam mengidentifikasi pola-pola visual secara otomatis, sangat efektif dalam mendeteksi kondisi mengantuk pengemudi melalui analisis ekspresi wajah, gerakan tangan, dan gerakan kepala, dengan mencapai akurasi hingga 85% dalam klasifikasi kondisi mengantuk [5].

3) *Pengolahan Citra Wajah*

Pengolahan citra wajah memainkan peran penting dalam berbagai aplikasi, termasuk deteksi ekspresi wajah dan pengenalan wajah. Teknik ini melibatkan transformasi citra wajah dari *input* ke *output* yang berbeda, tanpa melibatkan klasifikasi semantik tingkat tinggi seperti pengenalan wajah. Beberapa aplikasi utama termasuk penghapusan kacamata, sintesis ekspresi wajah, dan penghapusan mata merah. Proses ini membutuhkan teknik pemrosesan gambar yang cermat karena wajah sangat sensitif terhadap perubahan kecil dalam penampilan. Misalnya, penghapusan kacamata memerlukan teknik rekonstruksi citra menggunakan *Principal Component Analysis (PCA)* untuk menghasilkan citra wajah bebas kacamata yang realistis. Selain itu, pemrosesan ekspresi wajah digunakan dalam pembuatan karikatur dan simulasi perubahan berat badan wajah, yang semuanya berfokus pada peningkatan kualitas citra wajah dan memfasilitasi tugas pengenalan wajah lebih lanjut [18].

III. METODOLOGI

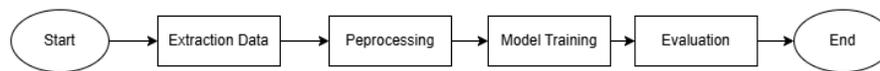


Fig 1. Alur Penelitian 1

A. Dataset

Dataset yang digunakan dalam penelitian ini adalah *Driver Drowsiness Dataset (DDD)* yang tersedia di Kaggle. Dataset ini terdiri dari gambar wajah pengemudi yang diambil dari video menggunakan perangkat lunak VLC. Proses ekstraksi gambar dilakukan dengan memanfaatkan algoritma Viola-Jones untuk mengidentifikasi dan memotong wajah pengemudi. Dataset ini memiliki lebih dari 41.790 gambar dengan dua kelas, yaitu *Drowsy* (kantuk) dan *Non-Drowsy* (tidak kantuk), dengan resolusi gambar 227 x 227 piksel. Gambar-gambar ini berformat RGB dan memiliki ukuran total sekitar 2.32 GB.

B. Preprocessing

Pada tahap *preprocessing*, data gambar diproses menggunakan *ImageDataGenerator* untuk memastikan konsistensi dan efisiensi pelatihan model. Gambar-gambar yang digunakan dalam penelitian ini dinormalisasi dengan *rescaling*, mengubah nilai piksel dari rentang [0, 255] menjadi [0, 1] menggunakan parameter *rescale=1./255*. Proses ini diterapkan pada tiga set data, yaitu data pelatihan, pengujian, dan validasi.

Setiap gambar diperkecil ukurannya menjadi 224x224 piksel agar sesuai dengan input yang diharapkan oleh model. Data diproses dalam batch dengan ukuran 16 gambar per iterasi dan label diberikan dalam format biner, yaitu *drowsy* dan *non-drowsy*. Untuk memastikan bahwa urutan data tidak mempengaruhi model, setiap set data diacak dengan parameter *shuffle=True*.

C. Model Training

Arsitektur yang digunakan dalam model training adalah MobileNetV2. MobileNetV2 dapat melakukan *transfer learning* dengan efisiensi komputasi tinggi. Pada tahap awal, MobileNetV2 dimuat dengan parameter `weights='imagenet'` untuk memanfaatkan bobot yang telah dilatih dengan dataset yang besar. Dalam implementasi ini, bagian atas model MobileNetV2 dihilangkan (parameter `include_top=False`) untuk memastikan bahwa model hanya menggunakan bagian base model yang bertanggung jawab atas ekstraksi fitur gambar.

Gambar input yang diberikan ke model memiliki ukuran 224x224 piksel dengan tiga saluran warna RGB. Untuk menyempurnakan fitur yang dihasilkan oleh model MobileNetV2, sejumlah lapisan tambahan ditambahkan setelah *base model*, termasuk *Flatten*, *Dense*, dan lapisan *output*. *Flatten* digunakan untuk mengubah keluaran dua dimensi dari base model menjadi vektor satu dimensi yang bisa diproses lebih lanjut oleh lapisan *Dense*. Selanjutnya, dua lapisan *Dense* dengan 1024 dan 512 unit, masing-masing menggunakan fungsi aktivasi ReLU, ditambahkan untuk menangkap kompleksitas lebih lanjut dari representasi gambar. Lapisan terakhir adalah lapisan *Dense* dengan dua unit dan fungsi aktivasi softmax, yang digunakan untuk klasifikasi biner `yawn` dan `no_yawn` (mengantuk dan tidak mengantuk), mengeluarkan probabilitas untuk masing-masing kelas.

Model kemudian dikompilasi dengan *Adam optimizer* dengan learning rate 0.0001, dan menggunakan *sparse categorical crossentropy* sebagai fungsi kehilangan, yang cocok untuk masalah klasifikasi multikelas dengan label yang disusun dalam bentuk integer. Model dilatih selama 20 epoch menggunakan data pelatihan yang telah diproses dan data validasi untuk mengevaluasi kinerja model di setiap epoch. Untuk menghindari *overfitting*, Model dilatih untuk menghasilkan akurasi yang optimal dengan memanfaatkan teknik *transfer learning* untuk menghindari pelatihan yang berlebihan.

D. Evaluation

Evaluasi model dilakukan menggunakan beberapa metrik untuk mengukur seberapa baik model dalam mendeteksi kantuk. Metrik yang digunakan adalah akurasi, presisi, *recall*, dan F1-score. Dalam menghitung nilai tersebut maka diperlukan perumusan yang akan dijelaskan sebagai berikut:

(1). Akurasi, mengukur seberapa banyak prediksi model yang benar dibandingkan dengan jumlah total data yang diuji. Akurasi dihitung dengan rumus:

$$\text{Akurasi} = \frac{\text{Jumlah Prediksi Benar}}{\text{Total Data}} = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

Keterangan:

- TP (*True Positive*): Jumlah prediksi benar untuk kelas positif (wajah kantuk).
- TN (*True Negative*): Jumlah prediksi benar untuk kelas negatif (wajah tidak mengantuk).
- FP (*False Positive*): Jumlah prediksi salah untuk kelas positif (wajah tidak mengantuk diprediksi sebagai wajah kantuk).
- FN (*False Negative*): Jumlah prediksi salah untuk kelas negatif (wajah kantuk) diprediksi sebagai wajah tidak mengantuk.

(2). , mengukur seberapa banyak prediksi positif yang benar dari total data yang seharusnya positif. Recall dihitung dengan rumus:

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

(3). Presisi, mengukur seberapa banyak prediksi positif yang benar dari total prediksi positif yang dilakukan oleh model. Presisi dihitung dengan rumus:

(3)

$$\text{Presisi} = \frac{TP}{TP + FP}$$

(4). F1-score, harmonik rata-rata antara presisi dan recall. F1-score memberikan gambaran lebih seimbang tentang kinerja model ketika ada ketidakseimbangan antara presisi dan recall. F1-score dihitung dengan rumus:

$$\text{F1-Score} = 2 \times \frac{\text{Presisi} \times \text{Recall}}{\text{Presisi} + \text{Recall}} = 2 \times \frac{TP}{2TP + FP + FN}$$

IV. EVALUASI

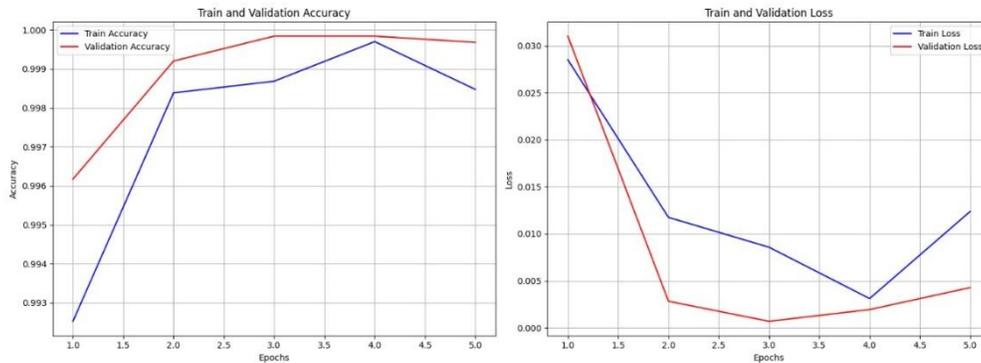


Fig 2. Grafik akurasi dan Loss 1

Berdasarkan hasil pelatihan pada fig 2, dapat disimpulkan bahwa model menunjukkan performa yang baik dalam hal akurasi dan kehilangan (loss). *Train accuracy* meningkat tajam selama beberapa epoch pertama, mencapai hampir 1.0 (100%) pada epoch ke-4 dan ke-5. Hal ini menunjukkan bahwa model mampu mempelajari pola dari data latih dengan sangat baik. *Validation accuracy* juga mengalami peningkatan serupa, dan stabil mendekati 1.0 pada epoch ke-4 dan ke-5, yang mengindikasikan kemampuan generalisasi model terhadap data yang belum pernah dilihat sebelumnya.

Namun, meskipun akurasi tinggi tercapai, terdapat indikasi *overfitting*, yang ditandai dengan munculnya gap antara *train accuracy* dan *validation accuracy* pada epoch ke-4 dan ke-5. Hal ini menunjukkan bahwa performa model pada data validasi mulai tertinggal dibandingkan dengan data latih, menandakan bahwa model mulai terlalu menyesuaikan diri terhadap data pelatihan.

Lebih lanjut, *train loss* terlihat menurun drastis pada epoch pertama hingga ke-3 dan mencapai nilai yang sangat rendah pada epoch ke-4, mencerminkan berkurangnya kesalahan pada data pelatihan. Peningkatan *validation loss* ini merupakan sinyal kuat bahwa model mengalami *overfitting*—di mana kemampuan adaptasinya terhadap data baru mulai menurun karena terlalu terpaku pada pola data pelatihan.

Secara keseluruhan, meskipun performa awal model sangat baik, munculnya indikasi *overfitting* pada tahap akhir pelatihan menandakan perlunya penerapan teknik regulasi tambahan. Beberapa pendekatan yang dapat digunakan untuk mencegah *overfitting* antara lain *dropout*, *data augmentation* yang bertujuan untuk menjaga keseimbangan antara performa pada data pelatihan dan kemampuan generalisasi terhadap data validasi. *Recall*

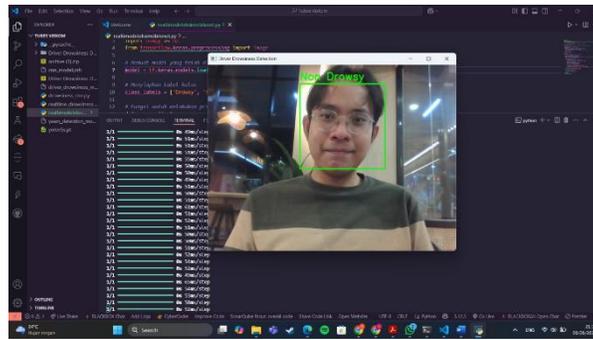


Fig 3. Tampilan Webcam Kondisi Non Drowsy

Gambar pada fig 3 hasil deteksi menggunakan webcam yang menggunakan *Haar Cascade* dari OpenCV. Dalam gambar ini, wajah pengguna terdeteksi dengan kotak hijau yang menandakan status "Non Drowsy" atau tidak mengantuk. Kondisi ini terlihat dari ekspresi wajah yang fokus, mata terbuka, dan posisi kepala yang tegak. Warna hijau pada kotak deteksi memperkuat interpretasi bahwa model mengenali pengguna dalam keadaan waspada.

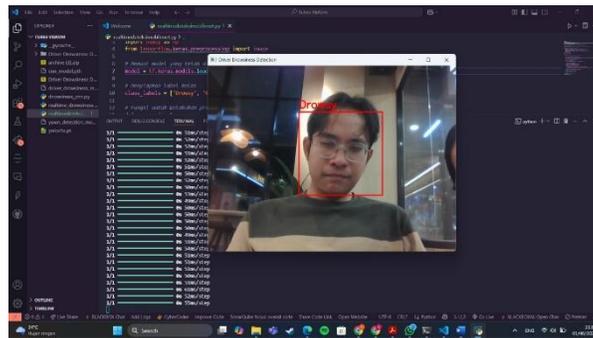


Fig 4. Tampilan Webcam Kondisi Drow

Gambar menampilkan hasil deteksi yang sama dengan situasi berbeda, yaitu kondisi pengguna yang terdeteksi mengantuk (*Drowsy*). Wajah pengguna diberi kotak merah sebagai indikator kondisi kantuk, ditandai dengan ekspresi mata yang tertutup dan posisi kepala sedikit menunduk. Model deteksi berhasil mengklasifikasikan kondisi ini secara tepat dengan label "Drowsy" yang muncul di pojok kiri atas kotak.

V. Kesimpulan

Model klasifikasi kondisi mengantuk pada citra wajah pengemudi yang dikembangkan menggunakan arsitektur MobileNetV2 dengan pendekatan *transfer learning* dan teknik data augmentation menunjukkan performa yang baik. Hasil pelatihan memperlihatkan peningkatan akurasi yang tajam pada beberapa epoch pertama dan stabil mendekati 1.0 pada epoch ke-4 dan ke-5, baik pada data pelatihan maupun validasi.

Meskipun demikian, terdapat indikasi *overfitting* yang ditandai dengan perbedaan antara akurasi pelatihan dan validasi pada tahap akhir pelatihan serta peningkatan nilai *validation loss*. Hal ini menunjukkan perlunya penerapan teknik regulasi tambahan seperti *dropout*, dan data *augmentation*.

Implementasi model dalam deteksi *real-time* menggunakan webcam berhasil mengklasifikasikan kondisi wajah pengemudi sebagai "drowsy" maupun "non-drowsy" dengan akurasi yang cukup baik, terlihat dari kotak deteksi berwarna merah dan hijau yang menandakan kondisi mengantuk dan tidak mengantuk secara berturut-

turut. Akan tetapi implementasi deteksi real-time yang dilakukan belum sepenuhnya sempurna karena beberapa indikator seperti sudut dari dataset yang kurang variatif, dataset yang diambil dalam bentuk *capture* video juga membuat indikasi *leak* data karena terdapat beberapa foto dengan wajah orang yang sama sedangkan proses *splitting* data dilakukan secara random menggunakan *library* sehingga hasil *real-time* kurang baik sedangkan proses training nilainya baik. Dengan demikian, walau model yang dikembangkan masih memiliki kekurangan akan tetapi tetap berpotensi untuk digunakan dalam sistem deteksi dini kantuk pada pengemudi guna meningkatkan keselamatan berkendara.

REFERENCES

- [1] National Sleep Foundation, *Drowsy driving: A National Sleep Foundation position statement and call to action*, 2023. Accessed: May 29, 2025.
- [2] R. Florez, F. Palomino-Quispe, R. J. Coaquira-Castillo, J. C. Herrera-Levano, T. Paixão, and A. B. Alvarez, *A CNN-based approach for the driver drowsiness detection by real-time eye state identification*, Preprints, 202306:0267, 2023.
- [3] R. Jabbar, M. Shinoy, M. Kharbeche, K. Al-Khalifa, M. Krichen, and K. Barkaoui, *Driver drowsiness detection model using convolutional neural networks techniques for android application*, Neural Computing and Applications, vol. 32, no. 20, pp. 16499–16512, 2020.
- [4] M. F. Shakeel, N. A. Bajwa, A. M. Anwaar, A. Sohail, A. Khan, and H. ur Rashid, *Detecting driver drowsiness in real time through deep learning based object detection*, SN Computer Science, vol. 1, p. 289, 2020.
- [5] M. Dua, Shakshi, R. Singla, S. Raj, and A. Jangra, *Deep CNN models-based ensemble approach to driver drowsiness detection*, Neural Computing and Applications, vol. 32, no. 20, pp. 16499–16512, 2020.
- [6] V. R. Reddy, U. S. Reddy, and V. K. K. Kolli, *Deep CNN: A machine learning approach for driver drowsiness detection based on eye state*, Revue d'Intelligence Artificielle, vol. 33, no. 6, pp. 461–466, 2019.
- [7] K. W. Kim, H. G. Hong, G. P. Nam, and K. R. Park, *A study of deep CNN-based classification of open and closed eyes using a visible light camera sensor*, Sensors, vol. 17, no. 7, p. 1534, 2017.
- [8] D. Salem and M. Waleed, *Drowsiness detection in real-time via convolutional neural networks and transfer learning*, Journal of Engineering and Applied Science, vol. 71, p. 122, 2024.
- [9] M. Hashemi, A. Mirrashid, and A. B. Shirazi, *Driver safety development: Real-time driver drowsiness detection system based on convolutional neural network*, SN Computer Science, vol. 1, no. 1, p. 289, 2020.
- [10] J.-M. Guo and H. Markoni, *Driver drowsiness detection using hybrid convolutional neural network and long short-term memory*, Multimedia Tools and Applications, vol. 77, no. 7, pp. 9295–9314, 2018.
- [11] S. P. Rajamohana, E. G. Radhika, S. Priya, and S. Sangeetha, *Driver drowsiness detection system using hybrid approach of convolutional neural network and bidirectional long short term memory (CNN-BiLSTM)*, Materials Today: Proceedings, vol. 26, pp. 1217–1223, 2020.
- [12] C. Lee and J. An, *LSTM-CNN model of drowsiness detection from multiple consciousness states acquired by EEG*, Expert Systems with Applications, vol. 213, p. 119032, 2023.
- [13] A. Sinha, A. R. P., and S. K. Gopal, *Drowsiness detection system using deep learning*, in *Proceedings of the 2021 IEEE International Conference on Smart Technologies for Sustainability (ICSTS)*, pp. 15–20, 2021.
- [14] V. Chand and J. Karthikeyan, *CNN-based driver drowsiness detection system using emotion analysis*, Intelligent Automation & Soft Computing, vol. 31, no. 2, pp. 711–728, 2022.
- [15] M. Zhu, J. Chen, H. Li, F. Liang, L. Han, and Z. Zhang, *Vehicle driver drowsiness detection method using wearable EEG based on convolution neural network*, Neural Computing and Applications, vol. 33, pp. 1–10, 2021.
- [16] V. P. Balam, V. U. Sameer, and S. Chinara, *Automated classification system for drowsiness detection using convolutional neural network and electroencephalogram*, IET Intelligent Transport Systems, vol. 15, pp. 514–524, 2021.
- [17] I. Nasri, M. Karrouchi, K. Kassmi, and A. Messaoudi, *A review of driver drowsiness detection systems: Techniques, advantages and limitations*, 2022.
- [18] X. Jiang and Y.-F. Chen, *Facial image processing*, in *Studies in Computational Intelligence*, pp. 29–48, Springer-Verlag Berlin Heidelberg, 2008.