

# Dampak Penggunaan Data Augmentasi Terhadap Akurasi MobileNetV2 Dalam Deteksi Mikrosleep Berbasis Rasio Aspek Mata

Isa Iant Maulana<sup>1</sup>, Muhammad Fatah Abiyyu Riadi<sup>1</sup>, Farrikh Alzami<sup>2</sup>, Muhammad Naufal<sup>1,\*</sup>,  
Harun Al Azies<sup>1</sup>, Ricardus Anggi Pramunendar<sup>1</sup>, Ruri Suko Basuki<sup>2</sup>

<sup>1</sup> Fakultas Ilmu Komputer, Teknik Informatika, Universitas Dian Nuswantoro, Semarang, Indonesia

<sup>2</sup> Fakultas Ilmu Komputer, Sistem Infomasi, Universitas Dian Nuswantoro, Semarang, Indonesia

Email: <sup>1</sup>111202214416@mhs.dinus.ac.id, <sup>2</sup>111202214409@mhs.dinus.ac.id, <sup>3</sup>alzami@dsn.dinus.ac.id,

<sup>4,\*</sup>m.naufal@dsn.dinus.ac.id, <sup>5</sup>harun.alazies@dsn.dinus.ac.id, <sup>6</sup>ricardus.anggi@dsn.dinus.ac.id, <sup>7</sup>ruri.basuki@dsn.dinus.ac.id

Email Penulis Korespondensi: m.naufal@dsn.dinus.ac.id

Submitted: 17/11/2025; Accepted: 16/12/2025; Published: 16/12/2025

**Abstrak**—Deteksi kondisi *microsleep* merupakan hal penting dalam mencegah kecelakaan akibat penurunan kewaspadaan, khususnya pada aktivitas yang membutuhkan konsentrasi tinggi seperti mengemudi. Penelitian ini bertujuan untuk mengembangkan model deteksi *microsleep* berbasis citra area mata dengan memanfaatkan metode ekstraksi mata menggunakan MediaPipe FaceMesh. Nilai EAR hanya digunakan untuk proses labeling yang menjadi dasar pembentukan dataset. Permasalahan utama yang diteliti adalah bagaimana menghasilkan model klasifikasi yang mampu membedakan kondisi mata normal dan kondisi *microsleep* secara akurat menggunakan data citra yang diambil dari potongan area mata. Untuk menjawab permasalahan tersebut, penelitian ini menerapkan serangkaian tahapan mulai dari pembentukan dataset, pemrosesan awal berupa penyesuaian ukuran citra, normalisasi, serta peningkatan kualitas melalui augmentasi data, hingga pelatihan model menggunakan arsitektur MobileNetV2 dengan teknik pembelajaran transfer dan *fine tuning*. Hasil eksperimen menunjukkan bahwa penggunaan strategi augmentasi data berpengaruh signifikan dalam meningkatkan performa model, dengan konfigurasi terbaik menghasilkan akurasi pengujian sebesar 87,54 persen, dengan metrik performa lainnya yang juga tinggi, yaitu Precision sebesar 88,64 persen, Recall (Sensitivity) sebesar 87,14 persen, dan F1-Score sebesar 87,34 persen. Temuan ini membuktikan bahwa pendekatan berbasis citra area mata yang dipadukan dengan model jaringan saraf konvolusi mampu memberikan performa yang menjanjikan dalam mendeteksi kondisi *microsleep*. Penelitian ini diharapkan dapat menjadi dasar pengembangan sistem deteksi *microsleep* yang lebih efektif dan dapat diimplementasikan pada lingkungan nyata.

**Kata Kunci:** Deteksi Kantuk; Eye Aspect Ratio; MobileNetV2; Fine Tuning; Data Augmentasi

**Abstract**—Detecting *microsleep* is important in preventing accidents caused by decreased alertness, especially in activities that require high concentration such as driving. This study aims to develop an image-based *microsleep* detection model using the MediaPipe FaceMesh. The EAR value is only used for the tagging process that forms the basis for dataset creation. The main problem investigated is how to produce a classification model that can accurately distinguish between normal eye conditions and *microsleep* conditions using image data taken from eye area snippets. To address this issue, this study applies a series of stages, starting from dataset formation, initial processing in the form of image size adjustment, normalization, and quality improvement through data augmentation, to model training using the MobileNetV2 architecture with transfer learning and fine-tuning techniques. The results of the experiment show that the use of data augmentation strategies has a significant effect on improving model performance, with the best configuration producing a test accuracy of 87.54 percent, with other high performance metrics, namely Precision of 88.64 percent, Recall (Sensitivity) of 87.14 percent, and F1-Score of 87.34 percent. These findings prove that an eye area image-based approach combined with a convolutional neural network model is capable of providing promising performance in detecting *microsleep* conditions. These findings prove that an approach based on eye area images combined with a convolutional neural network model can deliver promising performance in detecting *microsleep*. This research is expected to form the basis for the development of a more effective *microsleep* detection system that can be implemented in real world environments.

**Keywords:** Drowsiness Detection; Eye Aspect Ratio; MobileNetV2; Fine Tuning; Data Augmentation

## 1. PENDAHULUAN

*Microsleep* selama mengemudi telah menjadi masalah keselamatan yang serius dan kompleks di Indonesia, dengan implikasi yang membahayakan bagi semua pengguna jalan. Berdasarkan data dari Korps Lalu Lintas Kepolisian Republik Indonesia, *microsleep* pengemudi menyumbang sekitar 22% dari total kecelakaan lalu lintas yang terjadi setiap tahunnya, dengan angka kejadian yang terus mengalami peningkatan dalam lima tahun terakhir. Fenomena *microsleep* atau kantuk sesaat yang berlangsung hanya dalam hitungan 3-15 detik dapat menyebabkan konsekuensi yang fatal, dimana kendaraan mampu meluncur tanpa kendali pada jarak yang berbahaya dan meningkatkan risiko kecelakaan secara dramatis [1]. Kondisi ini semakin diperparah oleh berbagai faktor eksternal seperti padatnya lalu lintas, panjangnya jarak tempuh, tingginya intensitas mobilitas masyarakat Indonesia, serta budaya kerja yang menuntut produktivitas tinggi dengan waktu istirahat yang terbatas. Dalam konteks global, World Health Organization (WHO) mencatat bahwa *microsleep* pengemudi merupakan kontributor utama dalam kecelakaan lalu lintas di seluruh dunia, dengan perkiraan 20% dari semua kecelakaan lalu lintas disebabkan oleh kelelahan pengemudi.

Perkembangan teknologi *deep learning* dan kecerdasan buatan dalam beberapa tahun terakhir telah membuka babak baru yang revolusioner dalam sistem deteksi *microsleep* pengemudi. Berbagai pendekatan dan metodologi inovatif telah dikembangkan secara masif, mulai dari teknik *machine learning* konvensional yang mengandalkan fitur manual hingga model *deep learning* yang lebih kompleks dan mampu mengekstraksi fitur secara otomatis dengan

presisi tinggi. Sebuah tinjauan komprehensif terhadap teknik *machine learning* dan pengukuran perilaku telah dilakukan secara mendalam untuk mendeteksi *microsleep* pengemudi, menyoroti berbagai metode dan pendekatan yang tersedia [2]. *Systematic review* terkini yang mengkaji perkembangan model *deep learning* untuk deteksi *microsleep* menunjukkan dengan jelas bahwa pendekatan berbasis *convolutional neural network* (CNN) mampu memberikan hasil yang akurat dan konsisten dalam berbagai skenario [3]. Beberapa penelitian terkini telah berhasil mengembangkan arsitektur CNN novel dengan akurasi yang sangat tinggi, namun sayangnya model-model tersebut memerlukan sumber daya komputasi yang besar dan konsumsi daya yang tidak efisien [4].

Sistem *real-time* dengan analisis kedipan mata berdasarkan *computer vision* dan *deep learning* yang canggih juga telah diimplementasikan dalam berbagai penelitian [5], meskipun kinerjanya masih terbatas pada kondisi pencahayaan optimal dan kurang mampu beradaptasi dengan variasi lingkungan yang dinamis. Upaya optimasi khusus untuk perangkat *edge* telah dilakukan secara intensif melalui pengembangan model berbasis *TinyML* yang ringan dan efisien [6] serta implementasi pendekatan *bidirectional LSTM* untuk deteksi *microsleep* berbasis fitur wajah yang menunjukkan hasil optimal [7]. Implementasi kombinasi *MobileNet* dengan *Single Shot Multibox Detector* (SSD) untuk deteksi mata dan wajah telah menunjukkan hasil yang sangat menjanjikan dalam berbagai pengujian [8], sementara modifikasi *MobileNetV2* untuk deteksi menguap yang lebih spesifik juga telah berhasil dikembangkan dengan tingkat keberhasilan yang memuaskan [9].

Penelitian terbaru mulai mengintegrasikan teknik augmentasi data yang canggih dan optimasi model yang komprehensif untuk kondisi pencahayaan rendah dan menantang [10], menunjukkan upaya serius dan berkelanjutan untuk meningkatkan robustness sistem dalam menghadapi kondisi *real-world* yang kompleks. Pendekatan EfficientNet-KNN untuk deteksi *real-time microsleep* juga telah memberikan kontribusi signifikan dalam meningkatkan kecepatan proses deteksi tanpa mengorbankan akurasi [11]. Integrasi teknologi Internet of Things (IoT) dalam sistem deteksi *microsleep* telah membuka peluang baru untuk pengembangan sistem yang lebih komprehensif [12] berhasil mengintegrasikan deteksi *microsleep* dengan sistem alerting berbasis IoT, menciptakan ekosistem yang terhubung untuk meningkatkan keselamatan berkendara.

Sistem deteksi berbasis *computer vision* yang mengandalkan perilaku mata juga banyak diteliti. Salah satu parameter fisiologis yang dianggap *gold standard* dalam mendeteksi indikasi *microsleep* adalah Eye Aspect Ratio (EAR) [13]. EAR digunakan secara luas sebagai indikator perubahan bentuk mata untuk mengidentifikasi kedipan lambat, mata terpejam, atau fase penurunan kewaspadaan [14]. Namun, meskipun EAR efektif sebagai fitur manual, kinerjanya dapat menurun pada kondisi pencahayaan buruk atau bentuk wajah yang bervariasi. Hal ini mendorong pengembangan pendekatan berbasis citra mata menggunakan *deep learning*, yang memungkinkan model mengekstraksi fitur secara otomatis dengan robust pada kondisi lingkungan yang beragam.

Berdasarkan analisis mendalam dan komprehensif terhadap penelitian-penelitian terdahulu, teridentifikasi beberapa celah penelitian yang signifikan dan memerlukan perhatian khusus. Optimasi parameter *alpha* dan *freeze ratio* pada arsitektur *MobileNetV2* untuk deteksi *microsleep* belum dieksplorasi secara komprehensif dan sistematis, padahal parameter-parameter ini sangat krusial dalam menentukan efisiensi model pada perangkat *edge* dengan sumber daya terbatas. Dampak augmentasi data terhadap kemampuan generalisasi model juga belum diukur secara statistik dengan pendekatan yang *rigor* dan metodologis [15], sehingga kurang memiliki dasar ilmiah yang kuat dalam menentukan strategi augmentasi yang optimal. Selain itu, masih sering ditemukan masalah *overfitting* yang persisten ketika model diuji pada *dataset* dengan variasi subjek yang terbatas [16].

Deteksi *microsleep* pada pengemudi menggunakan *convolutional neural networks* (CNN) telah menunjukkan potensi yang signifikan [17], namun implementasi pada kondisi cahaya rendah masih menjadi tantangan besar. Penelitian terbaru yang mengintegrasikan Time Distributed *MobileNetV2* dengan Auto-CLAHE untuk deteksi *microsleep* pada region mata dalam kondisi cahaya rendah [18] memberikan solusi inovatif untuk mengatasi keterbatasan ini. Analisis komprehensif terhadap teknik deteksi *microsleep* berbasis analisis wajah secara *real-time* juga memberikan landasan kuat untuk pengembangan sistem yang responsif dan akurat [19], sementara perbandingan antara metode Haar Cascade dan YOLOFace untuk klasifikasi region of interest dalam deteksi *microsleep* memberikan wawasan berharga dalam pemilihan teknik deteksi wajah yang optimal [20].

Kelemahan dalam hal validasi juga menjadi gap penting yang perlu diisi. Sejumlah penelitian terdahulu melakukan pengujian hanya pada dataset laboratorium atau dataset publik yang cenderung memiliki kondisi yang terkontrol dan homogen [3], [19]. Hal ini menyebabkan model kurang representatif ketika digunakan pada kondisi mengemudi di dunia nyata yang penuh variasi [20]. Padahal, sistem deteksi kantuk yang akan diterapkan pada perangkat *edge* dan digunakan langsung oleh pengemudi membutuhkan tingkat generalisasi yang tinggi pada berbagai kondisi pencahayaan, sudut wajah, dan lingkungan kendaraan. Ketiadaan validasi pada dataset *real-world* menjadi celah penting yang menurunkan kesiapan implementasi model dalam skenario nyata.

Berdasarkan analisis tersebut, penelitian ini berusaha mengisi kesenjangan yang belum tercakup dalam penelitian sebelumnya melalui empat kontribusi utama. Pertama, melakukan optimasi arsitektur *MobileNetV2* yang menyeluruh dengan mengeksplorasi variasi parameter *alpha* dan *freeze ratio* secara sistematis untuk menemukan konfigurasi optimal yang menyeimbangkan antara akurasi dan efisiensi komputasi. Kedua, memanfaatkan EAR sebagai acuan fisiologis (*gold standart*) dalam proses pelabelan serta mengimplementasikan *pipeline* augmentasi data yang komprehensif dan terintegrasi dengan koreksi *brightness* yang adaptif dan transformasi geometris-fotometrik yang beragam untuk meningkatkan robustness model dalam berbagai kondisi pencahayaan dan lingkungan. Ketiga, menganalisis dampak statistik augmentasi data menggunakan uji hipotesis yang tepat dan reliable untuk memastikan

validitas hasil dan memberikan landasan ilmiah yang kuat dalam pemilihan strategi augmentasi. Keempat, memvalidasi model pada *dataset real-world* yang ekstensif dan mencakup berbagai kondisi mengemudi yang representatif untuk memastikan kemampuan generalisasi model dalam skenario nyata.

Pemilihan arsitektur MobileNetV2 dalam penelitian ini didasarkan pada pertimbangan yang matang terhadap karakteristik deteksi *microsleep* pada perangkat *edge*. MobileNetV2 menawarkan kombinasi optimal antara akurasi dan efisiensi melalui penggunaan *depthwise separable convolution* dan arsitektur *bottleneck* yang dioptimalkan. Kemampuannya untuk beradaptasi dengan berbagai constraint komputasi melalui parameter *alpha* menjadikannya kandidat ideal untuk implementasi pada perangkat dengan sumber daya terbatas. Selain itu, fleksibilitas dalam mengatur *freeze ratio* memungkinkan penyesuaian tingkat transfer learning yang optimal sesuai dengan karakteristik dataset target. Pendekatan ini diharapkan dapat mengatasi keterbatasan model sebelumnya yang memerlukan sumber daya komputasi besar dengan tetap mempertahankan akurasi deteksi yang tinggi.

Dengan adanya sistem deteksi *microsleep* yang akurat, efisien, dan andal ini, diharapkan dapat berkontribusi secara signifikan dalam mengurangi angka kecelakaan akibat *microsleep* pengemudi yang semakin memprihatinkan. Hasil penelitian ini juga diharapkan dapat menjadi dasar yang kokoh bagi pengembangan sistem keselamatan kendaraan yang lebih adaptif, responsif, dan intelijen terhadap kondisi nyata di jalan raya Indonesia, sekaligus membuka peluang implementasi teknologi keselamatan yang lebih terjangkau dan dapat diakses secara luas oleh masyarakat. Pada akhirnya, penelitian ini tidak hanya bermaksud memberikan kontribusi akademis tetapi juga memiliki dampak sosial yang nyata dan berkelanjutan dalam meningkatkan keselamatan dan kenyamanan berkendara di Indonesia, serta mendukung program pemerintah dalam mengurangi angka kecelakaan lalu lintas nasional melalui penerapan teknologi kecerdasan buatan yang tepat guna dan efektif.

## 2. METODOLOGI PENELITIAN

### 2.1 Tahapan Penelitian

Pada Gambar 1 terlihat tahapan penelitian mengikuti *workflow* sistematis yang mencakup koleksi data, preprocessing, split data, augmentasi, training model, dan evaluasi. Tahapan penelitian ini dirancang untuk mengoptimalkan arsitektur MobileNetV2 dalam mendeteksi kondisi *microsleep* secara real-time pada perangkat *edge* dengan sumber daya terbatas, mengacu pada pendekatan sistematis yang telah diterapkan dalam penelitian sebelumnya [6].



Gambar 1. Diagram Alir Tahapan Penelitian

Secara umum, Gambar 1 menunjukkan gambaran alur penelitian mulai dari mengambil dataset hingga tahapan evaluasi model. Setiap tahapan akan ditunjukkan pada diagram tersebut dan dijelaskan secara lebih mendalam pada sub bab 2.2 hingga 2.5.

### 2.2 Koleksi Data

Dataset dikumpulkan menggunakan kamera smartphone dengan resolusi 1080p, frame rate 30 fps, landscape dan berdurasi  $\pm 20$  detik dengan kondisi dan pencahayaan yang berbeda. Pada tahap awal, dataset dikumpulkan dalam bentuk video per subjek dengan empat kategori utama, yaitu normal, berbicara, menguap, dan *microsleep*. Setiap subjek menghasilkan empat video terpisah yang kemudian menjadi dasar proses ekstraksi frame pada tahap selanjutnya. Struktur dataset pada fase ini masih tersusun per subjek, dan seluruh data masih berada dalam format video sebelum diubah menjadi gambar pada proses eye extraction di tahap berikutnya.

### 2.3 Ekstraksi Fitur dengan Mediapipe

Dataset video yang telah dikumpulkan kemudian diekstaksi menggunakan landmark menjadi data numerik melalui proses ekstraksi frame. Proses ini dilakukan pada setiap frame dengan menggunakan framework MediaPipe untuk mendeteksi 468 titik landmark pada wajah secara real-time [21]. Dalam konteks penelitian ini, tidak semua dari 468 titik tersebut digunakan. Fokus utama diarahkan pada titik-titik di area mata, karena parameter utama yang diamati adalah perubahan Eye Aspect Ratio (EAR) untuk mendeteksi kondisi mata terbuka dan tertutup. Nilai EAR ini menjadi indikator utama dalam membedakan kondisi mata terbuka dan tertutup, sehingga memungkinkan identifikasi keadaan kantuk (*microsleep*) pada pengemudi secara akurat.

#### 2.3.1 Eye Aspect Ratio (EAR)

Eye Aspect Ratio (EAR) digunakan untuk mengukur derajat keterbukaan mata berdasarkan posisi landmark pada kelopak mata. EAR dihitung dari perbandingan antara jarak vertikal antara kelopak atas dan bawah dengan jarak horizontal antara sudut luar dan dalam mata [22]. Untuk mendapatkan nilai EAR dapat didefinisikan seperti (1).

$$EAR = \frac{\|P_2 - P_6\| + \|P_3 - P_5\|}{2\|P_1 - P_4\|} \quad (1)$$

Pada penelitian ini, deteksi landmark wajah difokuskan pada area periocular dengan 6 titik landmark untuk masing-masing mata [33, 160, 158, 133, 153, 144] untuk mata kiri dan [362, 385, 387, 263, 373, 380] untuk mata kanan, mengadopsi teknik yang telah teruji dalam deteksi kantung [17]. Setiap frame diekstraksi dan dilakukan deteksi wajah untuk memastikan kualitas data yang konsisten.

### 2.3.2 Seleksi Frame Berdasarkan Nilai EAR

Frame selection menggunakan algoritma Eye Aspect Ratio (EAR) dengan persamaan matematis yang telah terstandarisasi. Threshold EAR ditetapkan sebesar 0.25 untuk kondisi normal dan 0.18 untuk kondisi microsleep yang ditentukan melalui analisis distribusi EAR dari dataset preliminier menggunakan fungsi perhitungan EAR yang sama persis dengan implementasi sistem [16]. Nilai tersebut dipilih berdasarkan titik separasi optimal (*ROC-like*) yang memberikan pemisahan paling stabil antara mata terbuka dan microsleep. Seleksi frame dilakukan secara adaptif dengan interval minimum 1.0 detik untuk memastikan variasi data yang cukup dan menghindari duplikasi, mengikuti praktik terbaik dalam preprocessing data video untuk deteksi kantung [5].

## 2.4 Pre-Processing Data

Untuk memastikan Data siap digunakan sebelum proses training model, dilakukan beberapa langkah pre-processing data yang mencakup resize, normalisasi, konversi citra, dan penyesuaian brightness. Tahapan ini bertujuan untuk menyamakan dimensi, memperbaiki kualitas visual, dan menyesuaikan format citra agar sesuai dengan kebutuhan arsitektur model CNN yang digunakan.

### 2.4.1 Resize Citra

Pada tahap ini, seluruh citra diubah ukurannya menjadi  $128 \times 128$  piksel agar sesuai dengan spesifikasi input arsitektur MobilenetV2 yang digunakan dalam penelitian. Penyeragaman ukuran citra ini penting untuk menjaga konsistensi dimensi input, mengurangi kompleksitas komputasi, serta memastikan proses ekstraksi fitur dapat berjalan secara optimal sebagaimana dijelaskan pada referensi arsitektur MobilenetV2 [23].

### 2.4.2 Konversi Citra

Setelah itu, semua gambar dikonversi menjadi tiga kanal (RGB) agar CNN dapat memproses informasi warna secara penuh. Setiap filter konvolusi memerlukan kedalaman yang sama dengan jumlah kanal input, sehingga tiga kanal memungkinkan model mengekstrak fitur tekstur dan warna secara bersamaan, sesuai dengan desain CNN untuk citra berwarna [24].

### 2.4.3 Brightness

Kemudian, Penyesuaian brightness dilakukan secara adaptif berdasarkan rata-rata kecerahan gambar setiap gambar dievaluasi, dan apabila rata-rata brightness berada di bawah ambang batas tertentu, brightness ditingkatkan menggunakan transformasi linear sederhana. Pendekatan ini bertujuan untuk menangani variasi pencahayaan yang umum terjadi pada kondisi pengemudi dalam situasi nyata, sehingga model dapat belajar fitur wajah secara lebih konsisten di berbagai kondisi cahaya [15]. Peningkatan kecerahan ini membantu model CNN dalam mengenali pola ekspresi wajah secara lebih akurat, terutama pada kondisi pencahayaan rendah yang sebelumnya dapat menyebabkan kehilangan detail penting [24].

### 2.4.4 Normalisasi

Selanjutnya, setiap gambar dinormalisasi ke rentang  $[-1, 1]$  untuk menyesuaikan input dengan ekspektasi model, terutama saat menggunakan bobot pretrained dari ImageNet, sehingga meningkatkan kemampuan transfer learning, stabilitas training, dan konvergensi model [23].

$$x_{norm} = \frac{x}{127.5} - 1 \quad (2)$$

Rumus tersebut digunakan untuk mengubah nilai intensitas piksel asli yang berada pada rentang  $[0, 255]$  menjadi rentang  $[-1, 1]$ . Nilai 127.5 merupakan setengah dari nilai maksimum piksel (255), sehingga pembagian dengan 127.5 akan menskalakan data ke rentang  $[0, 2]$ , dan pengurangan 1 kemudian menggesernya menjadi  $[-1, 1]$ .

### 2.4.5 Split Dataset

Setelah seluruh tahapan pre-processing selesai, dataset kemudian dibagi menjadi 2 subset utama, yaitu Data train sebanyak 70%, dan data testing sebanyak 30% menggunakan stratified split. Strategi ini dirancang untuk menghindari bias pelatihan, mengukur kemampuan generalisasi model, serta mengevaluasi performa akhir secara objektif [25].

## 2.5 Model Training

Tahapan pelatihan model dilakukan menggunakan arsitektur MobileNetV2 dengan pendekatan transfer learning, yang telah terbukti efisien untuk pengenalan citra pada perangkat dengan sumber daya terbatas [26]. Konsep *transfer learning* memungkinkan model memanfaatkan bobot hasil pelatihan sebelumnya atau imagenet, sehingga mempercepat proses konvergensi dan meningkatkan akurasi meskipun menggunakan dataset yang lebih kecil [27].

Model dikustomisasi dengan menambahkan beberapa lapisan fully connected dan dropout setelah base model untuk menyesuaikan output dengan jumlah kelas pada dataset, yaitu kelas normal dan microsleep.

### 2.5.1 Fine Tuning dan Layer Freezing

Proses fine-tuning dilakukan dengan membekukan sebagian lapisan pada jaringan dasar (base model) dan hanya melatih lapisan akhir. Pendekatan ini bertujuan agar fitur-fitur umum yang telah dipelajari dari dataset ImageNet dapat dipertahankan, sementara lapisan akhir menyesuaikan diri terhadap karakteristik data baru, seperti pola wajah pengemudi pada kondisi microsleep dan normal. Teknik ini dapat meningkatkan efisiensi pelatihan serta mencegah overfitting pada dataset berukuran terbatas [27]. Dalam penelitian ini, dilakukan variasi freeze ratio sebesar 25%, 50%, 75%, hingga 100% dari total layer base model untuk menganalisis pengaruhnya terhadap performa klasifikasi.

### 2.5.2 Konfigurasi Arsitektur dan Hyperparameter

Setelah base model, ditambahkan Global Average Pooling untuk meratakan feature map menjadi vektor satu dimensi, dua lapisan Dense masing-masing 128 neuron dengan aktivasi ReLU dan regularisasi L2, serta lapisan Dropout berturut-turut 0,5 dan 0,4 untuk meningkatkan generalisasi model sebagaimana direkomendasikan pada arsitektur jaringan saraf konvolusi modern [24]. Lapisan output menggunakan fungsi aktivasi softmax sesuai jumlah kelas. Selain itu, variasi parameter alpha pada MobileNetV2 diuji dengan nilai 0,75 dan 1,0 untuk mengevaluasi pengaruh lebar jaringan terhadap akurasi dan efisiensi komputasi, sesuai dokumentasi resmi MobileNetV2 yang menjelaskan bahwa nilai alpha mempengaruhi ukuran dan performa model [23]. Model dioptimasi menggunakan Adam optimizer dengan Cosine Decay Learning Rate, sedangkan fungsi loss menggunakan Sparse Categorical Crossentropy, sesuai format label integer untuk klasifikasi multi-kelas [27].

## 2.6 Strategi Pelatihan dan Evaluasi

Pelatihan model dilakukan menggunakan Stratified K-Fold Cross-Validation sebanyak 7 fold, di mana setiap fold bergiliran menjadi subset validasi, sementara fold lainnya digunakan sebagai data train. Setiap fold dilatih selama 30 epoch dengan batch size 16, menggunakan early stopping untuk mencegah overfitting. Data augmentasi diterapkan hanya pada subset train fold, meliputi random horizontal flip, random rotation, dan random zoom, untuk meningkatkan variasi data dan kemampuan generalisasi model. Setelah pelatihan tiap fold selesai, model dievaluasi pada subset test yang terpisah untuk mendapatkan performa akurasi dan loss secara objektif.

Validasi model dilakukan dengan pembagian stratified 70:30 antara data train dan test pada setiap fold, sehingga evaluasi mencerminkan performa model secara komprehensif pada dataset terbatas dan meminimalkan bias [6]. Performansi tiap fold dicatat, kemudian dihitung rata-rata dan standar deviasi  $\pm$  std untuk tiap epoch, sehingga dapat dianalisis konsistensi model antar fold. Visualisasi performa menggunakan plot mean  $\pm$  std untuk akurasi train dan validation, memberikan indikasi stabilitas model selama proses pelatihan. Metrik evaluasi utama mengikuti standar dalam penelitian deteksi kantuk [12].

$$\text{Accuracy: } \frac{TP+TN}{TP+TN+FP+FN} \quad (3)$$

$$\text{Precision: } \frac{TP}{TP+FP} \quad (4)$$

$$\text{Recall: } \frac{TP}{TP+FN} \quad (5)$$

$$\text{F1-Score: } 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (6)$$

Analisis tambahan dilakukan melalui confusion matrix dan ROC curve untuk memahami perilaku model dalam mendeteksi kondisi microsleep secara mendalam [27]. Pendekatan ini bertujuan menghasilkan model yang stabil, mampu mengekstraksi fitur wajah secara optimal, dan memiliki performa klasifikasi tinggi dengan kompleksitas komputasi rendah. Evaluasi dilakukan secara menyeluruh terhadap kombinasi freeze ratio dan alpha, sehingga diperoleh konfigurasi model yang paling optimal untuk mendeteksi kondisi microsleep pada pengemudi.

## 3. HASIL DAN PEMBAHASAN

Bagian ini menyajikan hasil evaluasi model berdasarkan serangkaian eksperimen yang telah dilakukan dengan variasi parameter alpha, freeze ratio, serta penerapan data augmentations. Pengujian dilakukan untuk membandingkan performa model pada dua skenario utama, yaitu dengan augmentasi data dan tanpa augmentasi data. Setiap konfigurasi dianalisis berdasarkan metrik akurasi train, validasi, dan pengujian, nilai loss, serta waktu pelatihan rata-rata. Melalui penyajian ini, pembahasan difokuskan pada bagaimana setiap kombinasi parameter memengaruhi kemampuan generalisasi model, peningkatan performa, serta stabilitas selama proses pelatihan, sehingga dapat memberikan pemahaman mendalam mengenai efektivitas strategi yang digunakan dalam penelitian ini.

### 3.1 Data Collection

Pada tahap ini, seluruh video yang telah dikumpulkan sebagaimana dijelaskan pada Bab 2.2 proses melalui metode *eye extraction* menggunakan MediaPipe FaceMesh untuk memperoleh nilai Eye Aspect Ratio (EAR). Proses ini dilakukan pada setiap frame untuk mendeteksi area mata secara konsisten dan menghasilkan potongan citra mata yang valid sebagai dasar pembentukan dataset.

**Tabel 1.** Jumlah Dataset

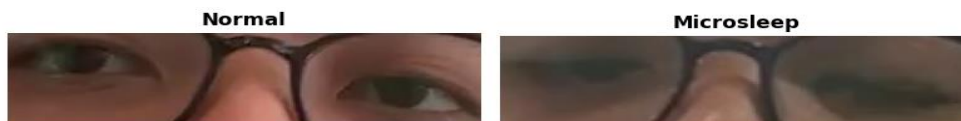
Kelas	Jumlah Partisipan
Normal	72
Berbicara	72
Menguap	72
Microsleep	72

Terlihat di Tabel 1, Jumlah dataset sama yaitu 72 partisipan untuk masing – masing kelas. Dataset tersebut mencakup mahasiswa Universitas Dian Nuswantoro serta beberapa partisipan non mahasiswa. Namun, pada penelitian ini hanya menggunakan data dari kondisi normal dan microsleep untuk process pelatihan dan pengujian model.

**Tabel 2.** Jumlah Dataset setelah Proses Eye Extractions

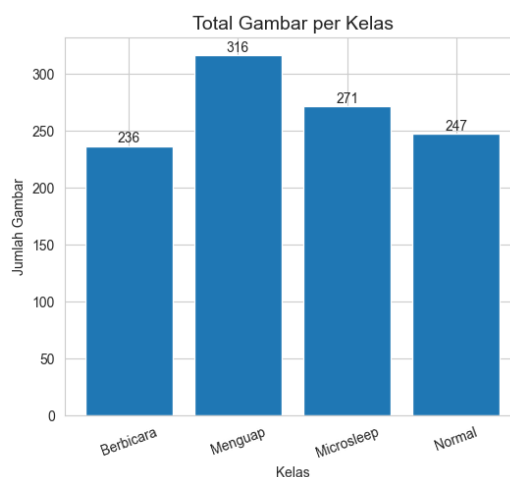
Kelas	Jumlah Partisipan
Normal	63
Berbicara	63
Menguap	63
Microsleep	63

Terlihat pada Tabel 2, dari total 72 subjek awal, hanya 63 subjek yang berhasil lolos proses ekstraksi karena pada sebagian subjek sistem tidak dapat mendeteksi area mata secara konsisten untuk perhitungan Eye Aspect Ratio (EAR). Setelah proses eye extraction diterapkan pada seluruh video, tidak semua subjek dapat dipertahankan dalam dataset akhir. Pada tahap deteksi EAR, sistem juga menghitung rata-rata koordinat landmark mata untuk setiap kelas. Nilai ini digunakan sebagai indikator apakah posisi mata berhasil terdeteksi secara konsisten pada setiap subjek.



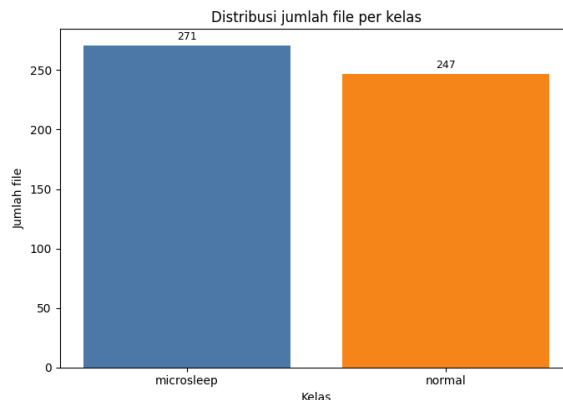
**Gambar 2.** Sampel Dataset setelah seleksi Frame

Terlihat di Gambar 2, setelah nilai EAR terdeteksi dan memenuhi kriteria seleksi, frame hasil pemotongan disimpan secara otomatis ke dalam folder direktori. Tetapi, subjek yang gagal terdeteksi dengan baik kemudian dikeluarkan dari dataset untuk menjaga kualitas data.



**Gambar 3.** Total Gambar setelah proses Eye Extractions

Pada Gambar 3 ditunjukkan distribusi data setelah proses *eye extraction*. Jumlah citra yang diperoleh adalah 236 untuk kelas Berbicara, 316 untuk kelas Menguap, 271 untuk kelas Microsleep, dan 247 untuk kelas Normal. Secara keseluruhan terdapat 1.070 citra hasil ekstraksi mata yang valid, dengan rata-rata 17.26 citra per subjek. Meskipun pada tahap awal struktur folder masih dipisahkan berdasarkan subjek, seluruh citra dari berbagai subjek kemudian digabungkan ke dalam dua kategori utama, yaitu normal dan microsleep, untuk keperluan pelatihan model.

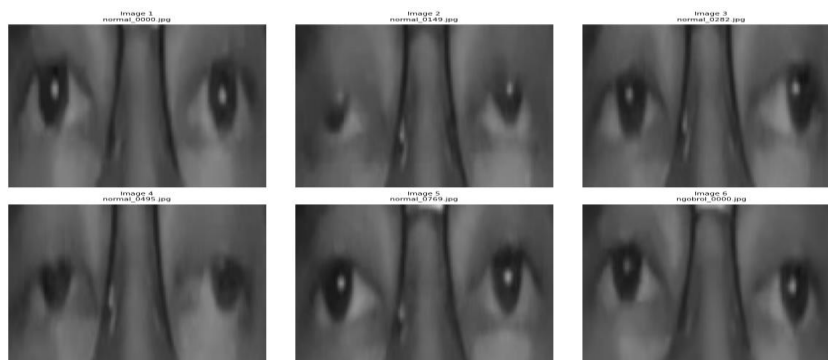


**Gambar 4.** Hasil Dataset setelah Proses Penggabungan

Pada Gambar 4, proses penggabungan dilakukan untuk mengurangi ketergantungan terhadap identitas subjek serta memastikan distribusi data yang lebih representatif. Jumlah citra pada kelas Microsleep sebanyak 271 dan kelas Normal sebanyak 247, menunjukkan distribusi yang telah cukup seimbang sehingga tidak diperlukan teknik penyeimbangan data tambahan seperti SMOTE. Untuk memperkuat bukti keseimbangan ini, jumlah total citra per kelas yang digunakan pada tahap *training* dan *testing* juga disajikan secara kuantitatif pada Gambar 7.

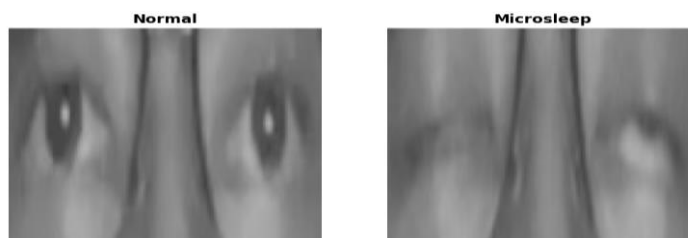
### 3.2 Hasil Pre-Processing Data

Tahapan pre-processing yang dijelaskan pada Bab 2.4 telah berhasil diterapkan pada seluruh dataset, mencakup proses *resize*, *normalisasi*, *konversi citra*, dan *penyesuaian brightness*. Hasil dari setiap tahap menunjukkan bahwa data telah berada dalam kondisi optimal untuk digunakan pada proses pelatihan model CNN.



**Gambar 5.** Sampel setelah dilakukan *resize*, dan konversi citra

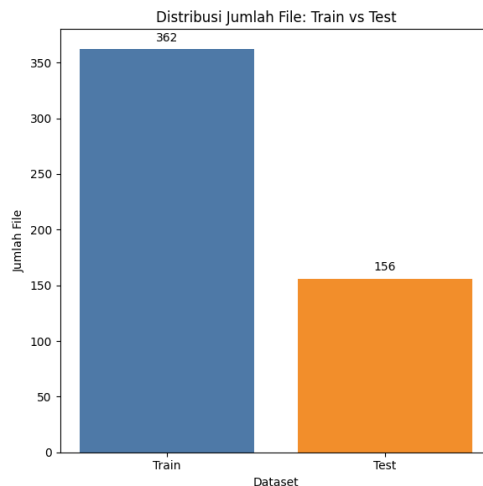
Di Gambar 5, citra berhasil di-*resize* menjadi ukuran  $128 \times 128$  piksel, sehingga memiliki dimensi yang seragam sesuai dengan kebutuhan arsitektur MobileNetV2. Selanjutnya, seluruh gambar berhasil dikonversi menjadi tiga kanal (RGB). Konversi ini memastikan bahwa informasi warna dan tekstur dapat diproses secara optimal oleh setiap filter konvolusi pada CNN, serta menjaga kompatibilitas dengan arsitektur jaringan yang memerlukan input tiga kanal.



**Gambar 6.** Sampel setelah dilakukan *Brightness*

Terlihat Gambar 6, bahwa gambar yang dihasilkan lebih terang dan memiliki tingkat pencahayaan yang lebih merata, sehingga fitur wajah seperti mata dan mulut menjadi lebih jelas untuk diekstraksi oleh model. Tahap penyesuaian *brightness* menghasilkan peningkatan kualitas pencahayaan pada citra dengan tingkat kecerahan rendah. Hasil visual menunjukkan bahwa citra menjadi lebih terang dan homogen, sehingga fitur wajah seperti mata, kontur wajah, dan area mulut terlihat lebih jelas. Perbaikan ini penting terutama untuk kondisi pencahayaan tidak merata, sehingga meningkatkan reliabilitas fitur yang dipelajari oleh model. Kemudian, dilakukan proses *normalisasi* ke

rentang nilai [-1,1] untuk menyesuaikan karakteristik input dengan standar pretrained ImageNet, sehingga mendukung stabilitas proses training.



Gambar 7. Hasil Data Train dan Test

Pada Gambar 7, dilakukan Split data dengan bobot 70% untuk data train, dan 30% untuk data testing. Terlihat bahwa distribusi data menunjukkan jumlah train sebanyak 362 gambar, sedangkan test sebanyak 156 gambar, sehingga proporsi pembagian data telah sesuai dengan bobot yang ditetapkan dan tetap mempertahankan keseimbangan distribusi kelas pada kedua subset tersebut.

### 3.3 Model Training

Pada tahap ini, model MobileNetV2 dilatih menggunakan konfigurasi arsitektur dan strategi pelatihan yang telah dijelaskan pada Bab 2. Pelatihan dilakukan untuk mengevaluasi kemampuan model dalam mengklasifikasikan dua kelas utama, yaitu normal dan microsleep, berdasarkan citra area mata yang telah diekstraksi pada tahap preprocessing.

Proses pelatihan mencakup penggunaan teknik transfer learning dan fine-tuning dengan variasi freeze ratio 25%, 50%, 75%, dan 100% dari jumlah layer serta dua nilai alpha 0.75 dan 1.0 yang menentukan lebar jaringan MobileNetV2. Selain itu, pelatihan model dilakukan dalam dua skenario, yaitu dengan menggunakan data augmentation dan tanpa data augmentation, untuk menganalisis perbandingan pengaruhnya terhadap performa model. Seluruh pelatihan dilakukan menggunakan pendekatan 7-Fold Stratified Cross-Validation, sehingga evaluasi performa model lebih stabil dan representatif terhadap dataset yang terbatas.

#### 3.3.1 Arsitektur MobileNetV2

Arsitektur model yang digunakan pada penelitian ini dibangun berdasarkan MobileNetV2 sebagai feature extractor dengan bobot awal (pretrained weights) dari ImageNet. MobileNetV2 dipilih karena memiliki struktur yang ringan dan efisien, sehingga mampu melakukan ekstraksi fitur secara optimal pada dataset berukuran kecil hingga menengah [27]. Pada tahap fine-tuning, model menerapkan mekanisme freeze ratio, yaitu persentase lapisan awal yang dibekukan untuk mencegah hilangnya pengetahuan representasional dari pretrained model.

Tabel 3. Arsitektur MobileNetV2

Bagian	Detail
Backbone	MobileNetV2 (ImageNet), input (128, 128, 3), alpha = 0.75 / 1.0
Freeze Ratio	25%, 50%, 75%, atau 100% lapisan awal dibekukan
Pooling	GlobalAveragePooling2D
Dense 1	128 neuron, ReLU, L2 (5e-4)
Dropout 1	0.5
Dense 2	128 neuron, ReLU, L2 (5e-4)
Dropout 2	0.4
Output Layer	Softmax, jumlah neuron = banyak kelas
Data Augmentation	Flip horizontal, rotasi 0.2, zoom 0.2
Optimizer	Adam + CosineDecay LR
Loss	Sparse Categorical Crossentropy
Training	7-Fold Stratified K-Fold

Terlihat pada Tabel 3, arsitektur dilengkapi dengan beberapa fully connected layers untuk proses klasifikasi. Bagian akhir jaringan terdiri dari lapisan GlobalAveragePooling2D, diikuti dua lapisan Dense dengan 128 neuron yang menggunakan aktivasi ReLU serta regularisasi L2 untuk mengurangi overfitting. Setiap lapisan Dense disertai

dengan dropout 0.5 dan 0.4 guna meningkatkan generalisasi model. Lapisan output menggunakan softmax untuk menghasilkan probabilitas pada masing-masing kelas.

### 3.3.2 Hasil Pelatihan Dengan Augmentasi Data

Skenario pertama, dilakukan pelatihan dengan menggunakan data augmentasi yang terdiri dari random horizontal flip, random rotation, dan random zoom berkontribusi positif terhadap peningkatan kemampuan generalisasi model. Augmentasi yang diterapkan hanya pada data pelatihan setiap fold terbukti efektif dalam mencegah overfitting, terutama mengingat ukuran dataset yang terbatas, sebagaimana dijelaskan pada bagian metodologi.

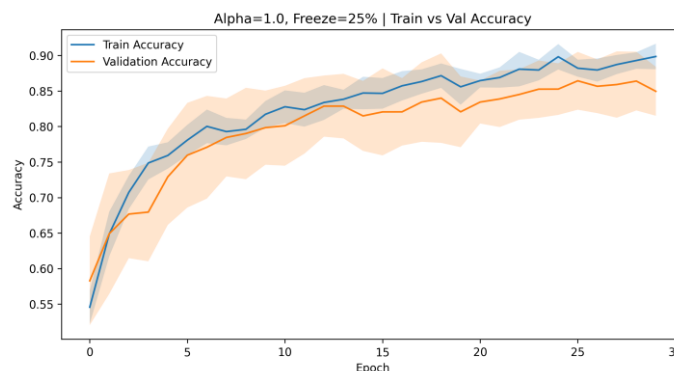
**Tabel 5.** Hasil Training menggunakan Augmentasi Data

Alpha	Freeze Ratio	Average Train Accuracy	Average Val Accuracy	Average Test Accuracy	Average Val Loss	Average Test Loss	Average Train Time (s)	Average Precision	Average Recall	Average F1 - Score
0.75	0.25	87,70	82,03	86,08	60,56	53,09	338,16	73,3260	72,1254	70,8827
0.75	0.5	87,10	79,34	79,57	77,81	59,36	363,63	83,4377	78,7386	78,0750
0.75	0.75	86,97	79,85	80,58	67,18	59,93	593,03	84,2677	79,7533	79,5660
0.75	1.0	87,88	78,73	79,30	70,21	61,42	566,69	84,0452	78,3643	77,8159
1.0	0.25	89,68	84,25	87,54	59,02	50,67	348,72	88,6374	87,1363	87,3384
1.0	0.5	87,24	84,24	85,71	59,44	53,01	406,49	87,2826	85,2340	85,3287
1.0	0.75	88,90	83,15	86,44	64,74	51,72	636,06	87,5112	86,0345	86,2146
1.0	1.0	87,34	84,80	85,89	58,46	53,33	653,66	86,9295	85,4930	85,6538

Secara keseluruhan, penggunaan data augmentation memberikan peningkatan performa yang signifikan pada seluruh konfigurasi. Terlihat di Tabel 5, model mampu mencapai akurasi test tertinggi sebesar 87,54% pada konfigurasi alpha = 1.0 dan freeze ratio = 0.25, yang merupakan hasil terbaik di antara seluruh variasi pengujian. Selain akurasi, konfigurasi ini juga menghasilkan metrik performa lain yang sangat baik, yaitu Precision sebesar 88,64%, Recall (*Sensitivity*) sebesar 87,14%, dan F1-Score sebesar 87,34%. Capaian ini menunjukkan bahwa kombinasi arsitektur MobileNetV2 dengan augmentasi data mampu meningkatkan ketahanan model terhadap variasi citra area mata, sehingga menghasilkan prediksi yang lebih akurat untuk kelas normal dan microsleep.

Selain itu, Nilai loss yang tercantum pada tabel bukan merupakan nilai loss mentah dari fungsi Sparse Categorical Crossentropy, tetapi merupakan nilai loss rate yang telah dikalikan 100 agar memudahkan interpretasi dan penyajian pada tabel. Dengan kata lain, nilai loss sebesar 50,67 pada tabel sebenarnya adalah 0,5067 dalam skala aslinya. Proses scaling ini dilakukan secara konsisten untuk seluruh nilai loss, baik pada data training, validation, maupun testing. Karena itu, nilai loss terlihat lebih besar dari rentang normal loss SCCE, tetapi secara matematis tetap sesuai dengan karakteristik fungsi loss yang digunakan. Sementara, rata-rata waktu pelatihan juga menunjukkan kecenderungan meningkat seiring bertambahnya freeze ratio dan kompleksitas model. Pada konfigurasi terbaik yaitu alpha = 1.0, freeze ratio = 0.25, waktu pelatihan berada pada kisaran 348,72 detik, yang relatif lebih efisien dibandingkan konfigurasi lain yang mencapai lebih dari 600 detik. Hal ini mengindikasikan bahwa model dengan freeze ratio lebih rendah tidak hanya memberikan akurasi yang lebih tinggi, tetapi juga memiliki waktu pelatihan yang lebih optimal, sehingga lebih cocok digunakan untuk skenario pelatihan berulang atau eksperimen berskala besar.

Pada Gambar 8, terlihat pada konfigurasi MobileNetV2 dengan alpha = 1.0 dan freeze ratio = 25% selama 30 epoch. Area berwarna di sekitar kurva merupakan rentang variasi ( $\pm$  standar deviasi) dari hasil 7-Fold Cross-Validation. Secara umum, akurasi pelatihan dan validasi mengalami peningkatan stabil seiring bertambahnya epoch. Kurva train dan validation terlihat berjalan berdekatan, menunjukkan bahwa model mampu belajar secara efektif tanpa mengalami overfitting yang signifikan. Pada awal pelatihan epoch 1–5, perbedaan kedua kurva masih relatif besar, tetapi mulai mengerucut setelah memasuki epoch ke-10, menandakan proses adaptasi model terhadap fitur citra berjalan baik.



**Gambar 8.** Hasil Akurasi dengan Data Augmentasi

### 3.3.3 Hasil Pelatihan Tanpa Augmentasi Data

Skenario kedua, dilakukan pelatihan tanpa menggunakan data augmentasi pada data pelatihan. Pada skenario ini, model dilatih langsung menggunakan citra asli tanpa penambahan variasi sintesis, sehingga performa yang dihasilkan mencerminkan kemampuan model dalam mempelajari pola dari data apa adanya.

**Tabel 5.** Hasil Training tanpa Augmentasi

Alpha	Freeze Ratio	Average Train Accuracy	Average Val Accuracy	Average Test Accuracy	Average Val Loss	Average Test Loss	Average Train Time (s)	Average Precision	Average Recall	Average F1 - Score
0.75	0.25	78,91	40,57	50,64	133,33	94,54	263.72	54,1196	51,429	45,6719
0.75	0.5	78,95	44,49	51,83	139,99	97,69	152.10	49,1339	52,2318	45,2800
0.75	0.75	78,45	40,30	49,17	139,51	106,82	147.66	41,2301	48,6887	38,7844
0.75	1.0	79,56	41,98	49,26	137,12	96,59	150.16	47,0536	48,8794	44,8698
1.0	0.25	85,12	61,87	61,81	105,57	88,49	247.20	63,8721	60,8979	58,6351
1.0	0.5	84,43	61,32	61,81	100,11	91,40	272.13	62,2937	61,0298	59,9939
1.0	0.75	85,03	62,13	61,44	96,34	85,04	267.80	63,343	61,0957	59,4031
1.0	1.0	84,66	63,84	64,01	100,84	84,41	279.71	64,6561	63,3181	62,1080

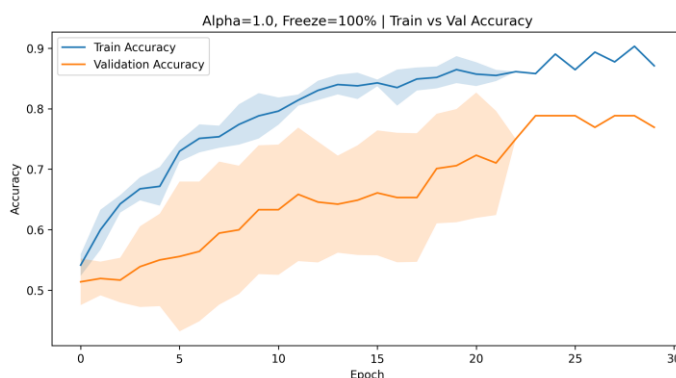
Dari Tabel 5, model menunjukkan kecenderungan akurasi yang lebih rendah dibandingkan skenario dengan augmentasi. Hal ini terlihat dari nilai akurasi validasi dan akurasi test yang cenderung stagnan serta lebih mudah mengalami overfitting, terutama pada konfigurasi freeze ratio yang tinggi. Ketiadaan variasi data menyebabkan model kurang mampu mengenali distribusi citra yang beragam, sehingga generalisasi pada data baru menjadi terbatas.

Secara khusus, pada konfigurasi dengan  $\alpha = 0.75$ , akurasi validasi hanya berada pada kisaran 40% hingga 44%, sementara akurasi test berkisar 49–51%. Nilai ini mengindikasikan bahwa model belum mampu menangkap pola penting secara efektif. Selain itu, validation loss yang cukup tinggi, hal tersebut semakin mempertegas bahwa model mengalami kesulitan dalam melakukan generalisasi dan cenderung terjebak pada pola yang hanya muncul pada data pelatihan.

Dari sisi metrik klasifikasi lainnya, performa model pada konfigurasi  $\alpha = 0.75$  juga menunjukkan hasil yang kurang memuaskan. Rata-rata Precision berada pada rentang 41,23% hingga 54,12%, Recall berkisar antara 48,69% hingga 52,23%, dan F1-Score berada pada rentang 38,78% hingga 45,67%. Nilai-nilai ini menunjukkan bahwa model tidak hanya kurang akurat, tetapi juga kurang konsisten dalam membedakan kelas normal dan microsleep, sehingga menegaskan bahwa konfigurasi ini belum optimal untuk tugas deteksi microsleep.

Jika dilihat dari rata-rata waktu pelatihan, konfigurasi tanpa augmentasi cenderung membutuhkan waktu yang lebih singkat, yaitu berkisar antara 147 hingga 279 detik, tetapi waktu yang cepat ini tidak berbanding lurus dengan performa karena model justru mengalami penurunan akurasi dan generalisasi. Hal ini menunjukkan bahwa efisiensi waktu pelatihan pada skenario tanpa augmentasi tidak mampu mengimbangi lemahnya kemampuan model dalam mempelajari fitur penting dari citra.

Berdasarkan Gambar 9, performa sedikit membaik pada konfigurasi  $\alpha = 1.0$  dengan freeze ratio = 100%, di mana akurasi validasi meningkat ke kisaran 61–63%, dan akurasi test berada pada 61–64%. Walaupun demikian, selisih antara rata-rata akurasi pelatihan (84–85%) dan akurasi validasi tetap besar, menandakan adanya overfitting yang cukup kuat. Pola ini konsisten dengan kurva pelatihan yang menunjukkan jarak signifikan antara train accuracy dan validation accuracy. Dengan kata lain, model terlalu menyesuaikan diri terhadap data pelatihan namun gagal mempertahankan performa yang stabil pada data validasi.



**Gambar 9.** Hasil Akurasi tanpa Data Augmentasi

Selain itu, nilai validation loss dan test loss yang masih relatif tinggi menunjukkan bahwa ketiadaan variasi citra membuat model kesulitan menangkap representasi fitur yang robust. Hal ini mengonfirmasi bahwa tanpa augmentasi data, model cenderung sensitif terhadap distribusi dataset yang terbatas sehingga menghasilkan performa generalisasi yang lemah. Secara keseluruhan, hasil ini menunjukkan bahwa pelatihan tanpa augmentasi tidak mampu

memberikan performa yang kompetitif, dan augmentasi terbukti berperan penting dalam meningkatkan akurasi serta mengurangi overfitting pada tugas klasifikasi microsleep.

#### 4. KESIMPULAN

Penelitian ini menyimpulkan bahwa proses deteksi microsleep berbasis citra area mata dapat dilakukan secara efektif melalui rangkaian tahapan mulai dari pembentukan dataset menggunakan metode eye extraction berbasis MediaPipe FaceMesh, proses pre-processing yang terstruktur, hingga pelatihan model menggunakan arsitektur MobileNetV2, di mana efektivitas ini dibuktikan oleh capaian akurasi tertinggi 87,54% pada pengujian, serta 84,80% pada validasi pada konfigurasi model terbaik yang menggunakan MobileNetV2 dengan  $\alpha = 1.0$  dan freeze ratio = 0.25, serta dilatih menggunakan skenario dengan *data augmentation* yang terbukti memberikan performa lebih stabil dibandingkan model tanpa augmentasi. Seluruh tahapan tersebut mampu menjawab permasalahan utama penelitian, yaitu bagaimana membangun model klasifikasi kondisi mata yang akurat menggunakan data citra yang dihasilkan dari perhitungan nilai EAR. Hasil eksperimen menunjukkan bahwa penggunaan data augmentation berpengaruh signifikan dalam meningkatkan performa model, di mana skenario dengan augmentasi menghasilkan akurasi tertinggi mencapai 87,54%, dengan metrik performa lainnya yang juga tinggi, yaitu Precision sebesar 88,64 persen, Recall (Sensitivity) sebesar 87,14 persen, dan F1-Score sebesar 87,34 persen. Sementara model tanpa augmentasi cenderung mengalami overfitting. Temuan ini menegaskan pentingnya variasi data dalam memperkuat kemampuan generalisasi model. Namun demikian, penelitian ini memiliki beberapa keterbatasan, seperti ukuran dataset yang masih terbatas, perbedaan pencahayaan pada beberapa citra yang mempengaruhi konsistensi fitur visual, serta pemrosesan yang hanya berfokus pada area mata tanpa mempertimbangkan fitur wajah lain yang mungkin turut berpengaruh dalam pendeteksian microsleep. Oleh karena itu, penelitian selanjutnya dapat diarahkan pada pengumpulan data yang lebih beragam, penggunaan metode peningkatan kualitas citra yang lebih adaptif, serta pengembangan model yang mempertimbangkan fitur wajah secara menyeluruh agar sistem deteksi microsleep dapat bekerja lebih robust dan akurat dalam berbagai kondisi nyata.

#### REFERENCES

- [1] “Benarkah Microsleep Bisa Sebabkan Kecelakaan Lalu Lintas? | tempo.co,” Tempo. Diakses: 22 Juni 2025. [Daring]. Tersedia pada: <https://www.tempo.co/otomotif/benarkah-microsleep-bisa-sebabkan-kecelakaan-lalu-lintas--71757>
- [2] M. Ngxande, J. -R. Tapamo, dan M. Burke, “Driver drowsiness detection using behavioral measures and machine learning techniques: A review of state-of-art techniques,” dalam *2017 Pattern Recognition Association of South Africa and Robotics and Mechatronics (PRASA-RobMech)*, Des 2017, hlm. 156–161. doi: 10.1109/RoboMech.2017.8261140.
- [3] T. Fonseca dan S. Ferreira, “Drowsiness Detection in Drivers: A Systematic Review of Deep Learning-Based Models,” *Applied Sciences*, vol. 15, no. 16, hlm. 9018, Jan 2025, doi: 10.3390/app15169018.
- [4] F. Majeed, U. Shafiq, M. Safran, S. Alfarhood, dan I. Ashraf, “Detection of Drowsiness among Drivers Using Novel Deep Convolutional Neural Network Model,” *Sensors (Basel)*, vol. 23, no. 21, hlm. 8741, Okt 2023, doi: 10.3390/s23218741.
- [5] F. Safarov, F. Akhmedov, A. B. Abdusalomov, R. Nasimov, dan Y. I. Cho, “Real-Time Deep Learning-Based Drowsiness Detection: Leveraging Computer-Vision and Eye-Blink Analyses for Enhanced Road Safety,” *Sensors (Basel)*, vol. 23, no. 14, hlm. 6459, Jul 2023, doi: 10.3390/s23146459.
- [6] N. N. Alajlan dan D. M. Ibrahim, “DDD TinyML: A TinyML-Based Driver Drowsiness Detection Model Using Deep Learning,” *Sensors (Basel)*, vol. 23, no. 12, hlm. 5696, Jun 2023, doi: 10.3390/s23125696.
- [7] A. R. Hartono, M. Naufal, dan F. Alzami, “Optimalisasi Arsitektur LSTM dengan Pendekatan Bidirectional untuk Deteksi Kantuk Pengemudi Berbasis Fitur Wajah,” *Building of Informatics, Technology and Science (BITS)*, vol. 7, no. 2, hlm. 1343–1354, Sep 2025, doi: 10.47065/bits.v7i2.8219.
- [8] J. S. Wijnands, J. Thompson, K. A. Nice, G. D. P. A. Aschwanden, dan M. Stevenson, “Real-time monitoring of driver drowsiness on mobile platforms using 3D neural networks,” *Neural Computing and Applications*, vol. 32, no. 13, hlm. 9731–9743, Jul 2020, doi: 10.1007/s00521-019-04506-0.
- [9] H. Z. Ilmadina, M. Naufal, dan D. S. Wibowo, “Drowsiness Detection Based on Yawning Using Modified Pre-trained Model MobileNetV2 and ResNet50,” *MATRIK : Jurnal Manajemen, Teknik Informatika dan Rekayasa Komputer*, vol. 22, no. 3, hlm. 419–430, Jun 2023, doi: 10.30812/matrik.v22i3.2785.
- [10] F. Alzami dkk., “Bayesian-Optimized CLAHE for Enhanced Drowsiness Detection in Low-Light Conditions Using Time-Distributed MobileNetV2-GRU Architecture,” *Statistics, Optimization & Information Computing*, Okt 2025, doi: 10.19139/soic-2310-5070-3024.
- [11] S. Winarno, F. Alzami, M. Naufal, H. A. Azies, M. A. Soeleman, dan N. H. A. H. Malim, “EfficientNet-KNN for Real-Time Driver Drowsiness Detection via Sequential Image Processing,” *Ingénierie des Systèmes d’Information*, vol. 30, no. 7, hlm. 1703, Jul 2025, doi: 10.18280/isi.300703.
- [12] A.-C. Phan, T.-N. Trieu, dan T.-C. Phan, “Driver drowsiness detection and smart alerting using deep learning and IoT,” *Internet of Things*, vol. 22, hlm. 100705, Jul 2023, doi: 10.1016/j.iot.2023.100705.
- [13] S. Alimoon dkk., “Vision-Based Driver Fatigue Detection Using Eye and Mouth Aspect Ratios with Machine Learning,” vol. 9, hlm. 29–32, Apr 2025.
- [14] D. Santos, *Real-time Eye Blink Detection using Computer Vision and Machine Learning*. 2024. doi: 10.20944/preprints202410.0131.v1.
- [15] F. Alzami, S. Winarno, M. Naufal, dan H. Al Azies, “Enhancing Driver Drowsiness Detection through GMM-Optimized CLAHE,” dalam *2024 International Seminar on Application for Technology of Information and Communication (iSemantic)*, Sep 2024, hlm. 212–217. doi: 10.1109/iSemantic63362.2024.10762529.



- [16] H. Z. Ilmadina, M. Naufal, dan D. Apriliani, “Microsleep detection on drowsy-driver using convolutional neural networks (CNN),” dalam *AIP Conference Proceedings*, AIP Publishing LLC, 2024, hlm. 030005.
- [17] M. F. Shakeel, N. A. Bajwa, A. M. Anwaar, A. Sohail, A. Khan, dan Haroon-ur-Rashid, “Detecting Driver Drowsiness in Real Time Through Deep Learning Based Object Detection,” dalam *Advances in Computational Intelligence*, I. Rojas, G. Joya, dan A. Catala, Ed., Cham: Springer International Publishing, 2019, hlm. 283–296.
- [18] F. Alzami, M. Naufal, H. Al Azies, S. Winarno, dan M. A. Soeleman, “Time Distributed MobileNetV2 with Auto-CLAHE for Eye Region Drowsiness Detection in Low Light Conditions.,” *International Journal of Advanced Computer Science & Applications*, vol. 15, no. 11, hlm. 488, Nov 2024, doi: 10.14569/ijacsa.2024.0151146.
- [19] S. Essahraoui *dkk.*, “Real-Time Driver Drowsiness Detection Using Facial Analysis and Machine Learning Techniques,” *Sensors*, vol. 25, no. 3, 2025, doi: 10.3390/s25030812.
- [20] M. N. Andrian *dkk.*, “Comparing haar cascade and yoloface for region of interest classification in drowsiness detection,” *Jurnal Media Informatika Budidarma*, vol. 8, no. 1, hlm. 272–281, 2024, doi: 10.30865/mib.v8i1.7167.
- [21] J.-R. González-Rodríguez, D.-M. Cordova-Esparza, J. Terven, dan J.-A. Romero-González, “Towards a Bidirectional Mexican Sign Language–Spanish Translation System: A Deep Learning Approach,” *Technologies*, vol. 12, hlm. 7, Jan 2024, doi: 10.3390/technologies12010007.
- [22] C. Dewi, R.-C. Chen, C.-W. Chang, S.-H. Wu, X. Jiang, dan H. Yu, “Eye Aspect Ratio for Real-Time Drowsiness Detection to Improve Driver Safety,” *Electronics*, vol. 11, no. 19, hlm. 3183, Jan 2022, doi: 10.3390/electronics11193183.
- [23] K. Team, “Keras documentation: MobileNet, MobileNetV2, and MobileNetV3.” Diakses: 13 November 2025. [Daring]. Tersedia pada: <https://keras.io/api/applications/mobilenet/>
- [24] B. Kromydas, “Convolutional Neural Network: A Complete Guide.” Diakses: 13 November 2025. [Daring]. Tersedia pada: <https://learnopencv.com/understanding-convolutional-neural-networks-cnn/>
- [25] X. Jiao, Y. Luo, Y. Bi, dan T. Zhou, “Impacts of Data Splitting Strategies on Parameterized Link Prediction Algorithms,” 11 November 2025, *arXiv*: arXiv:2511.05834. doi: 10.48550/arXiv.2511.05834.
- [26] P. Młodzianowski, “Weather Classification with Transfer Learning - InceptionV3, MobileNetV2 and ResNet50,” dalam *Digital Interaction and Machine Intelligence*, C. Biele, J. Kacprzyk, W. Kopeć, J. W. Owsinski, A. Romanowski, dan M. Sikorski, Ed., Cham: Springer International Publishing, 2022, hlm. 3–11. doi: 10.1007/978-3-031-11432-8\_1.
- [27] D. Naveenprabhu, R. Naveen, B. N. Bharathi, dan M. Naveenkumar, “Comprehensive review of transfer learning,” *AIP Conference Proceedings*, vol. 2914, no. 1, hlm. 050028, Des 2023, doi: 10.1063/5.0175841.