

KLASIFIKASI TINGKAT FOKUS PESERTA DIDIK BERDASARKAN HEAD POSE ESTIMATION MENGGUNAKAN MOBILENETV3

Dika Adi Prasetya, Muhammad Najibulloh Muzaki, Aidina Ristyawan

Sistem Informasi, Universitas Nusantara PGRI Kediri

Jl. Ahmad Dahlan No.76, Mojoroto, Kec. Mojoroto, Kota Kediri, Jawa Timur 64112

dikaadi040@gmail.com

ABSTRAK

Perkembangan pembelajaran daring dan hibrida menuntut adanya sistem pemantauan fokus peserta didik yang lebih objektif dan otomatis. Namun, proses pemantauan fokus saat ini masih dilakukan secara manual melalui pengamatan dosen sehingga bersifat subjektif dan kurang efektif pada pembelajaran berbasis video conference. Penelitian ini bertujuan membangun sistem klasifikasi tingkat fokus peserta didik berbasis citra wajah menggunakan arsitektur MobileNetV3 agar mampu berjalan secara real-time pada perangkat dengan keterbatasan sumber daya. Metode penelitian yang digunakan adalah CRISP-DM dengan memanfaatkan dataset AFLW2000-3D yang memiliki anotasi orientasi kepala berupa yaw dan pitch dalam kondisi in-the-wild. Tahapan penelitian meliputi preprocessing citra, pelabelan otomatis berdasarkan threshold sudut kepala, pelatihan model menggunakan transfer learning, serta implementasi sistem monitoring berbasis Streamlit. Selain itu, dilakukan perbandingan performa antara MobileNetV3 dan YOLOv8 pada pengujian real-time. Hasil penelitian menunjukkan bahwa model MobileNetV3 memperoleh accuracy sebesar 93%, precision 0,84, recall 0,87, dan F1-score 0,85. Pada implementasi real-time, MobileNetV3 menunjukkan performa yang lebih stabil dan ringan dibandingkan YOLOv8 pada perangkat CPU. Berdasarkan hasil tersebut, MobileNetV3 dinilai efektif digunakan sebagai sistem monitoring fokus peserta didik secara otomatis dan real-time.

Kata kunci : Head Pose Estimation, MobileNetV3, Computer Vision, Real-time Monitoring, Tingkat Fokus.

1. PENDAHULUAN

Perkembangan teknologi pembelajaran daring (*online learning*) dan *hybrid learning* saat ini semakin meningkat dan berperan penting dalam mendukung proses belajar mengajar. Salah satu faktor yang memengaruhi keberhasilan pembelajaran adalah tingkat fokus peserta didik selama proses pembelajaran berlangsung, karena fokus berkaitan langsung dengan pemahaman materi dan pencapaian hasil belajar. Namun, proses pemantauan fokus peserta didik masih sering dilakukan secara manual melalui pengamatan dosen selama kegiatan pembelajaran, sehingga penilaian yang dihasilkan cenderung subjektif dan kurang efektif, terutama pada kelas dengan jumlah peserta didik yang banyak-[1], [2].

Salah satu indikator visual yang dapat digunakan untuk mengukur tingkat fokus adalah orientasi kepala (*head pose*) berdasarkan parameter yaw dan pitch. Akan tetapi, kondisi fokus tidak selalu ditunjukkan dengan posisi wajah yang menghadap lurus ke depan karena aktivitas seperti membaca atau menulis juga termasuk perilaku fokus. Oleh karena itu, diperlukan sistem yang mampu mengklasifikasikan tingkat fokus secara lebih fleksibel dan objektif berdasarkan arah orientasi kepala peserta didik[2], [3].

Beberapa penelitian sebelumnya telah menggunakan kombinasi model seperti YOLO, MTCNN, dan MobileNet untuk mendeteksi fokus secara real-time[4]. Pendekatan tersebut mampu menghasilkan performa deteksi yang baik, namun memiliki kompleksitas arsitektur dan kebutuhan

komputasi yang cukup tinggi sehingga kurang optimal diterapkan pada perangkat dengan keterbatasan sumber daya. Selain itu, sebagian penelitian sebelumnya masih menggunakan dataset dengan kondisi lingkungan yang terkontrol sehingga kurang merepresentasikan kondisi nyata (*in-the-wild*)[5].

Berdasarkan permasalahan tersebut, penelitian ini mengusulkan sistem klasifikasi tingkat fokus peserta didik berbasis *computer vision* menggunakan arsitektur MobileNetV3. Model ini dipilih karena memiliki karakteristik ringan (*lightweight*), efisien, dan mendukung implementasi real-time pada perangkat dengan spesifikasi terbatas [6], [7]. Dataset yang digunakan adalah AFLW2000-3D yang memiliki anotasi orientasi kepala berupa yaw dan pitch pada kondisi nyata (*in-the-wild*)[8]. Selain itu, penelitian ini juga membandingkan performa MobileNetV3 dengan YOLOv8 untuk menentukan model yang paling optimal dalam sistem monitoring fokus peserta didik secara real-time.

2. TINJAUAN PUSTAKA

2.1. Computer Vision

Computer vision merupakan salah satu bidang dalam kecerdasan buatan (Artificial Intelligence) yang memungkinkan komputer memperoleh dan memahami informasi dari gambar maupun video digital. Dalam penelitian ini, *computer vision* digunakan untuk mendeteksi wajah peserta didik dan menganalisis orientasi kepala secara real-time menggunakan kamera laptop[8].

2.2. Convolutional Neural Network dan MobileNetV3

Convolutional Neural Network (CNN) merupakan metode *deep learning* yang banyak digunakan pada pengolahan citra karena mampu melakukan ekstraksi fitur secara otomatis melalui proses konvolusi[9]. Pada penelitian ini digunakan arsitektur MobileNetV3 yang memiliki karakteristik ringan dan efisien sehingga cocok diterapkan pada perangkat dengan keterbatasan sumber daya. Selain itu, MobileNetV3 mendukung implementasi *transfer learning* menggunakan bobot awal (*pre-trained weights*) dari ImageNet sehingga proses pelatihan model menjadi lebih optimal. Secara umum, CNN terdiri dari beberapa lapisan utama, yaitu:

- Convolution Layer, untuk mengekstraksi fitur dari gambar.

$$Y(i,j) = \sum_m \sum_n x(i+m, j+n) \cdot k(m,n) \quad (1)$$

Keterangan:

X: citra input

K: kernel / filter

Y: feature map hasil konvolusi

Operasi konvolusi digunakan untuk mengekstraksi fitur dari citra dengan cara menggeser kernel pada citra input.

- Activation Function (ReLU), untuk menambahkan non-linearitas

$$f(x) = \max(0, x) \quad (2)$$

Fungsi aktivasi ReLU digunakan untuk menambahkan sifat non-linear pada model CNN. Sedangkan Softmax Layer (Klasifikasi Multikelas) digunakan pada lapisan output untuk mengubah nilai hasil keluaran jaringan saraf menjadi probabilitas pada setiap kelas.

- Softmax

$$\text{softmax}(z_i) = \frac{e^{z_i}}{\sum_{j=1}^n e^{z_j}} \quad (3)$$

Keterangan:

z_i : nilai output ke-i dari neuron

n: jumlah kelas

e: bilangan eksponensial

Kelas dengan nilai probabilitas tertinggi akan dipilih sebagai hasil prediksi model

2.3. YOLOv8n

YOLOv8n merupakan varian Nano dari algoritma YOLOv8 yang dirancang untuk kebutuhan deteksi objek secara *real-time* dengan komputasi ringan. Model ini memiliki jumlah parameter yang lebih kecil dibandingkan YOLOv8s, YOLOv8m, maupun YOLOv8l sehingga lebih optimal digunakan pada perangkat dengan keterbatasan sumber daya seperti laptop tanpa GPU. Model YOLOv8n menggunakan format file *.pt* sebagai model *pretrained* berbasis PyTorch[10].

2.4. Evaluasi

Penerlitan ini menggunakan akurasi, precision, recall, dan F1-score agar hasil evaluasi lebih komprehensif dan objektif.

akurasi prediksi, akurasi digunakan untuk mengukur tingkat ketepatan prediksi model secara keseluruhan, yaitu perbandingan antara jumlah prediksi yang benar dengan total seluruh data uji. Berikut ini untuk rumus penentuan dari accuracy, precision, recall, dan f1-score[11].

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (4)$$

$$\text{Precision} = \frac{TP}{TP+FP} \quad (5)$$

$$\text{Recall} = \frac{TP}{TP+FN} \quad (6)$$

$$\text{F1-Score} = \frac{2 \times (\text{Precision} \times \text{Recall})}{\text{Precision} + \text{Recall}} \quad (7)$$

- TP (True Positive): jumlah data yang diprediksi benar sebagai kelas positif
- TN (True Negative): jumlah data yang diprediksi benar sebagai kelas negatif
- FP (False Positive): jumlah data yang salah diprediksi sebagai kelas positif
- FN (False Negative): jumlah data yang salah diprediksi sebagai kelas negative

2.4. Penelitian Terdahulu

Penelitian sebelumnya oleh Fitri P.T menggunakan kombinasi YOLO, MTCNN, dan MobileNetV2 untuk mendeteksi fokus peserta didik secara *real-time*[12]. Pendekatan tersebut memiliki performa yang cukup baik, namun membutuhkan komputasi yang tinggi karena menggunakan beberapa model sekaligus. Penelitian lain menggunakan metode OpenCV dan Dlib yang lebih ringan, tetapi kurang stabil terhadap perubahan pencahayaan dan variasi pose wajah[13]. Oleh karena itu, penelitian ini menggunakan MobileNetV3 karena lebih ringan, efisien, dan sesuai untuk implementasi sistem monitoring fokus secara *real-time*.

3. METODE PENELITIAN

Penelitian ini menggunakan metode CRISP-DM (*Cross Industry Standard Process for Data Mining*) yang terdiri dari tahapan *Business Understanding*, *Data Understanding*, *Data Preparation*, *Modeling*, *Evaluation*, dan *Deployment*.

Dataset yang digunakan adalah AFLW2000-3D yang memiliki anotasi orientasi kepala berupa *yaw* dan *pitch* dalam kondisi nyata (*in-the-wild*). Tahap *preprocessing* meliputi *face cropping* menggunakan Haar Cascade, *resize* citra menjadi 224×224 piksel, normalisasi data, serta augmentasi citra untuk meningkatkan kemampuan generalisasi model. Dataset kemudian dibagi menjadi data latih sebesar 80% dan data uji sebesar 20%.

Proses pelabelan dilakukan berdasarkan nilai orientasi kepala menggunakan parameter *yaw* dan *pitch*. Kategori fokus ditentukan menggunakan batas toleransi sudut tertentu sebagaimana ditunjukkan pada Tabel 1.

Tabel 1. Kategori Fokus Orientasi Kepala

Kategori	Judul pertama	Judul kedua	Nama pertama
Fokus	$\leq 30^\circ$	$\leq 20^\circ$	Orientasi kepala masih dalam batas wajar untuk membaca dan mengerjakan soal
Tidak Fokus	$> 50^\circ$	$> 30^\circ$	Orientasi kepala ekstrem (menunduk/mendongak berlebihan) yang mengindikasikan distraksi

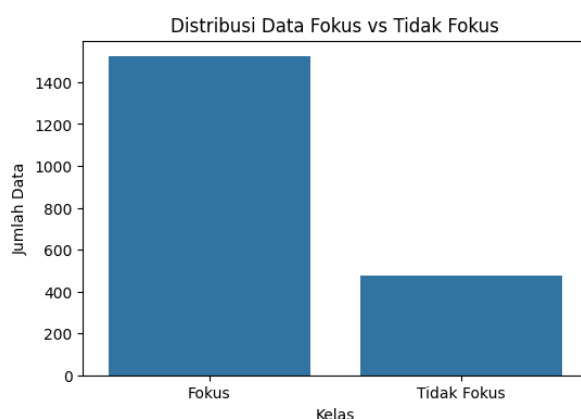
Model utama yang digunakan adalah MobileNetV3 dengan pendekatan *transfer learning* menggunakan bobot *pre-trained* ImageNet. Proses pelatihan menggunakan optimizer Adam dengan *learning rate* sebesar 0,0001 dan fungsi aktivasi Softmax pada lapisan output. Evaluasi model dilakukan menggunakan metrik *accuracy*, *precision*, *recall*, *F1-score*, serta *confusion matrix*.

4. HASIL DAN PEMBAHASAN

Hasil pelatihan menunjukkan bahwa model MobileNetV3 mampu mencapai tingkat akurasi sebesar 93%. Grafik *accuracy* dan *loss* menunjukkan bahwa model mampu mempelajari pola data dengan baik selama proses pelatihan. Nilai *loss* yang semakin menurun menunjukkan bahwa model semakin optimal dalam melakukan prediksi.

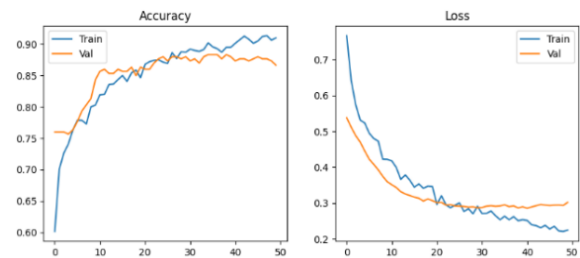
Evaluasi menggunakan *confusion matrix* menunjukkan bahwa model mampu mengklasifikasikan sebagian besar data secara benar dengan tingkat kesalahan klasifikasi yang relatif kecil. Berdasarkan hasil pengujian, model memperoleh nilai *precision* sebesar 0,84, *recall* sebesar 0,87, dan *F1-score* sebesar 0,85.

Selain itu, dilakukan perbandingan performa antara MobileNetV3 dan YOLOv8 pada implementasi sistem deteksi fokus secara *real-time*. Hasil pengujian menunjukkan bahwa kedua model memiliki tingkat akurasi yang sama, namun MobileNetV3 memiliki performa *real-time* yang lebih stabil dan ringan pada perangkat CPU.



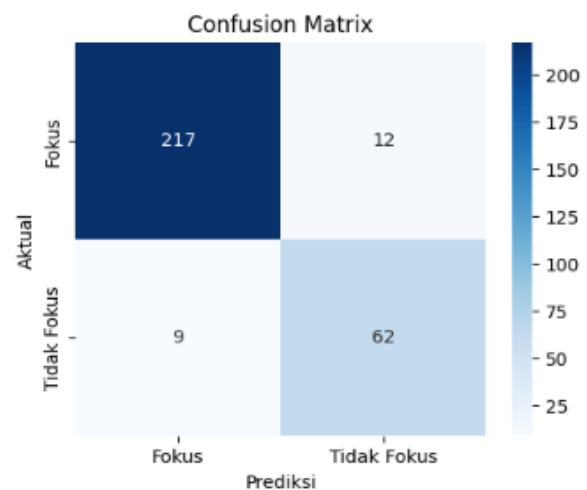
Gambar 1. distribusi data

Pada gambar 1 menunjukkan distribusi data antara kelas Fokus dan Tidak Fokus pada dataset yang digunakan. Berdasarkan hasil visualisasi, jumlah data pada kelas Fokus lebih banyak dibandingkan dengan kelas Tidak Fokus. Ketidakseimbangan jumlah data ini berpotensi menyebabkan model cenderung lebih bias terhadap kelas mayoritas.



Gambar 2. grafik accuracy mobilnetv3

Gambar 2 menunjukkan grafik *accuracy* dan *loss* selama proses pelatihan model MobileNetV3. Pada grafik *accuracy* terlihat bahwa nilai akurasi training dan validation mengalami peningkatan secara bertahap hingga mencapai lebih dari 90%. Hal ini menunjukkan bahwa model mampu mempelajari pola dari data dengan baik.



Gambar 3. confusion matrix hasil testing Mobilenetv3

Gambar 3 menunjukkan confusion matrix hasil testing yang digunakan untuk mengevaluasi performa model dalam mengklasifikasikan data Fokus dan Tidak Fokus. Berdasarkan hasil tersebut, model mampu mengklasifikasikan sebagian besar data dengan benar. Pada kelas Fokus, sebagian besar data berhasil diprediksi dengan benar, dengan hanya sedikit kesalahan klasifikasi ke kelas Tidak Fokus.

Begitu juga pada kelas Tidak Fokus, model mampu mengenali sebagian besar data dengan baik meskipun masih terdapat beberapa kesalahan prediksi. Hasil ini menunjukkan bahwa model memiliki performa yang cukup baik dalam membedakan antara kondisi fokus dan tidak fokus.

```
# =====
# 4. LABELING
# =====
def classify_focus(yaw, pitch):
    if abs(yaw) <= 30 and abs(pitch) <= 20:
        return 0 # Fokus
    elif abs(yaw) > 50 or abs(pitch) > 30:
        return 1 # Tidak Fokus
    else:
        return 0 # anggap masih fokus
```

Gambar 4. Pelebelan arah pandangan

Gambar 4 menunjukkan proses pelabelan arah pandangan data yang dilakukan berdasarkan nilai yaw dan pitch. Data diklasifikasikan menjadi dua kategori, yaitu Fokus dan Tidak Fokus. Jika nilai yaw berada pada rentang -30° hingga $+30^{\circ}$ dan pitch berada pada rentang -20° hingga $+20^{\circ}$, maka data dikategorikan sebagai Fokus. Sedangkan jika nilai yaw lebih dari 50° atau pitch lebih dari 30° , maka data dikategorikan sebagai Tidak Fokus. Pendekatan ini digunakan untuk merepresentasikan arah pandang kepala sebagai indikator tingkat fokus.

```
=== CLASSIFICATION REPORT ===
              precision    recall  f1-score   support

     0       0.96       0.95       0.95        229
     1       0.84       0.87       0.86         71

 accuracy         0.90
 macro avg       0.90       0.91       0.90        300
 weighted avg    0.93       0.93       0.93        300
```

Gambar 5. Matrix hasil testing mobilnetv3

```
Classification Report:
              precision    recall  f1-score   support

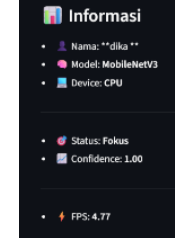
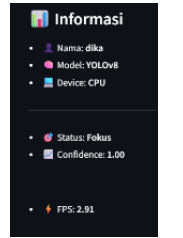
     0       0.94       0.96       0.95        229
     1       0.87       0.82       0.84         71

 accuracy         0.90
 macro avg       0.90       0.89       0.90        300
 weighted avg    0.93       0.93       0.93        300
```

Gambar 6. matrix hasil testing yolov8

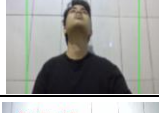
Pada penelitian ini dilakukan perbandingan performa antara model MobileNetV3 dan YOLOv8 dalam mengklasifikasikan kondisi fokus dan tidak fokus mahasiswa. Pengujian dilakukan menggunakan dataset test yang sama untuk memastikan hasil yang objektif dan adil. Berdasarkan hasil pengujian, kedua model memiliki nilai akurasi yang sama yaitu sebesar 93%. Namun, terdapat perbedaan pada metrik evaluasi lainnya seperti precision, recall, dan F1-score. Model YOLOv8 memiliki nilai precision yang lebih tinggi, yaitu sebesar 0.87, yang menunjukkan bahwa model lebih selektif dalam melakukan klasifikasi. Sementara itu, MobileNetV3 memiliki nilai recall sebesar 0.87 dan F1-score sebesar 0.85, yang menunjukkan bahwa model lebih seimbang dalam mendeteksi kedua kelas.

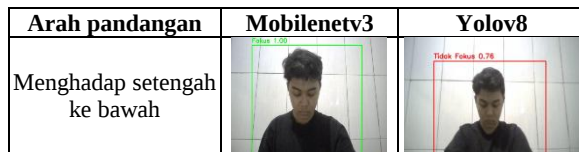
Tabel 2. perbandingan FPS

MobileNetV3	Yolov8	FPS
		Mobilnetv3 lebih unggul untuk FPS, yang mendapatkan 4.77
		sedangkan Yolov8 hanya mendapatkan FPS 2.91

Selain akurasi, dilakukan juga perbandingan dari sisi kecepatan pemrosesan (FPS) dalam implementasi real-time menggunakan kamera. Berdasarkan hasil pengujian, model MobileNetV3 memiliki performa yang lebih ringan dibandingkan YOLOv8, sehingga menghasilkan nilai FPS yang lebih tinggi.

Tabel 3. perbandingan tes

Arah pandangan	MobileNetv3	Yolov8
Menghadap lurus ke depan		
Menghadap ke samping kiri		
Menghadap ke samping kanan		
Menghadap ke atas		
Menghadap ke bawah		
Menghadap setengah ke samping kiri		
Menghadap setengah ke samping kanan		
Menghadap setengah ke atas		



Berdasarkan hasil pengujian yang ditampilkan pada Tabel 4.2, dilakukan perbandingan performa antara model MobileNetV3 dan YOLOv8 dalam mendeteksi tingkat fokus mahasiswa berdasarkan variasi arah pandangan kepala.

Pengujian dilakukan pada beberapa skenario posisi kepala, seperti menghadap lurus, menoleh ke samping, mendongak, serta menunduk. Setiap skenario diuji menggunakan kedua model untuk melihat konsistensi dan akurasi prediksi.

Analisis skenario:

a. Posisi Menghadap Lurus

Pada kondisi menghadap lurus ke depan, kedua model yaitu MobileNetV3 dan YOLOv8 mampu mengklasifikasikan kondisi sebagai fokus dengan tingkat confidence yang tinggi. Hal ini menunjukkan bahwa kedua model memiliki performa yang baik pada kondisi wajah frontal.

b. Posisi Setengah ke Samping (Kiri & Kanan)

Pada kondisi wajah sedikit menoleh ke samping, MobileNetV3 masih mampu mempertahankan prediksi fokus, sedangkan YOLOv8 mulai mengklasifikasikan sebagai tidak fokus pada beberapa kasus. Hal ini menunjukkan bahwa YOLOv8 lebih sensitif terhadap perubahan sudut wajah dibandingkan MobileNetV3.

c. Posisi Menghadap Samping (Ekstrem)

Pada posisi wajah menghadap ke samping secara signifikan, kedua model secara konsisten mengklasifikasikan sebagai tidak fokus. Hal ini sesuai dengan definisi fokus yang digunakan dalam penelitian, dimana orientasi kepala yang jauh dari arah depan dianggap sebagai tidak fokus.

d. Posisi Mendongak (Atas)

Pada kondisi mendongak, terdapat perbedaan hasil antara kedua model. MobileNetV3 cenderung masih mengklasifikasikan sebagai fokus, sedangkan YOLOv8 lebih sering mengklasifikasikan sebagai tidak fokus. Hal ini menunjukkan bahwa YOLOv8 lebih ketat dalam mendeteksi arah pandangan vertikal.

e. Posisi Menunduk (Bawah)

Pada kondisi menunduk, kedua model menunjukkan variasi hasil. MobileNetV3 dalam beberapa kasus masih mengklasifikasikan sebagai fokus, sedangkan YOLOv8 lebih konsisten dalam mendeteksi sebagai tidak fokus. Hal ini menunjukkan bahwa YOLOv8 lebih sensitif terhadap perubahan pitch (atas-bawah).

f. Posisi Kombinasi (Miring + Atas/Bawah)

Pada kondisi kombinasi seperti menoleh sambil mendongak atau menunduk, kedua model secara umum mengklasifikasikan sebagai tidak fokus,

namun YOLOv8 menunjukkan tingkat konsistensi yang lebih tinggi dibandingkan MobileNetV3.

Kesimpulan Perbandingan:

Berdasarkan hasil pengujian tersebut, dapat disimpulkan bahwa:

- MobileNetV3 memiliki keunggulan dalam:
 - Stabil pada kondisi wajah frontal dan semi frontal & Tidak terlalu sensitif terhadap perubahan kecil
- YOLOv8 memiliki keunggulan dalam:
 - Lebih sensitif terhadap perubahan arah pandangan (yaw dan pitch) & Lebih konsisten dalam mendeteksi kondisi tidak fokus

Secara umum:

YOLOv8 lebih cocok untuk deteksi yang ketat (strict detection) & MobileNetV3 lebih cocok untuk kondisi real-time yang lebih fleksibel.

Tabel 4. Perbandingan Model

Aspek	MobileNetV3	YOLOv8
Akurasi	93%	93%
Precision	0.84	0.87
Recall	0.87	0.82
F1-score	0.85	0.84
FPS	± 2.5 – 4.7 FPS	± 2.2 – 4.4 FPS
Stabilitas	Lebih stabil pada berbagai arah wajah	Lebih sensitif terhadap perubahan arah
Kesesuaian realtime	Sangat sesuai (ringan & konsisten)	Cukup sesuai (lebih berat & fluktuatif)

Berdasarkan hasil perbandingan pada Tabel 4.3, kedua model yaitu MobileNetV3 dan YOLOv8 memiliki tingkat akurasi yang sama sebesar 93%. Namun, terdapat perbedaan pada metrik lainnya. YOLOv8 memiliki nilai precision yang lebih tinggi yaitu 0.87, yang menunjukkan kemampuan model dalam meminimalkan kesalahan prediksi positif.

Sementara itu, MobileNetV3 memiliki nilai recall yang lebih tinggi sebesar 0.87 dibandingkan YOLOv8 sebesar 0.82, yang menunjukkan bahwa MobileNetV3 lebih baik dalam mendeteksi seluruh kondisi fokus secara keseluruhan.

Dari sisi performa real-time, MobileNetV3 menunjukkan performa yang lebih stabil dengan nilai FPS yang relatif konsisten. Hal ini menunjukkan bahwa MobileNetV3 lebih ringan dan lebih cocok untuk implementasi sistem deteksi fokus secara real-time.

Sebaliknya, YOLOv8 memiliki sensitivitas yang lebih tinggi terhadap perubahan arah wajah, namun hal ini menyebabkan model menjadi kurang stabil dalam beberapa kondisi tertentu. Oleh karena itu, berdasarkan hasil evaluasi, MobileNetV3 dinilai lebih sesuai untuk digunakan dalam sistem deteksi fokus mahasiswa secara real-time.

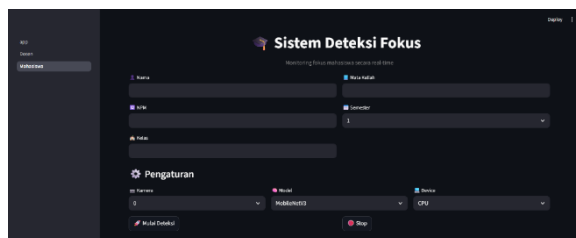
Tabel 5. MobileNetV3 vs YOLOv8

Model	Accuracy	Precision	Recall	F1-score	FPS
MobileNetV3	93%	0,84	0,87	0,85	4,7
YOLOv8	93%	0,87	0,82	0,84	2,5

Model kemudian diimplementasikan menggunakan framework Streamlit untuk melakukan monitoring fokus peserta didik secara *real-time* melalui webcam. Sistem mampu menampilkan hasil klasifikasi fokus, persentase fokus, serta dashboard monitoring dosen secara langsung.

Berdasarkan hasil penelitian, MobileNetV3 mampu mengklasifikasikan tingkat fokus peserta didik berdasarkan orientasi kepala menggunakan parameter *yaw* dan *pitch* dengan tingkat akurasi sebesar 93%. Selain memiliki performa klasifikasi yang baik, model juga menunjukkan performa *real-time* yang lebih stabil dibandingkan YOLOv8 sehingga lebih optimal diterapkan pada perangkat dengan keterbatasan sumber daya. Sistem berhasil diimplementasikan menggunakan Streamlit untuk monitoring fokus peserta didik secara otomatis melalui webcam. Penelitian selanjutnya dapat dikembangkan dengan menambahkan fitur *eye tracking*, *multi-face detection*, serta optimasi GPU agar sistem mampu digunakan pada lingkungan pembelajaran dengan jumlah peserta didik yang lebih banyak.

Pada tahap selanjutnya dilakukan implementasi sistem deteksi fokus mahasiswa berbasis web menggunakan framework Streamlit. Sistem ini dirancang untuk melakukan deteksi fokus mahasiswa secara *real-time* menggunakan kamera.\



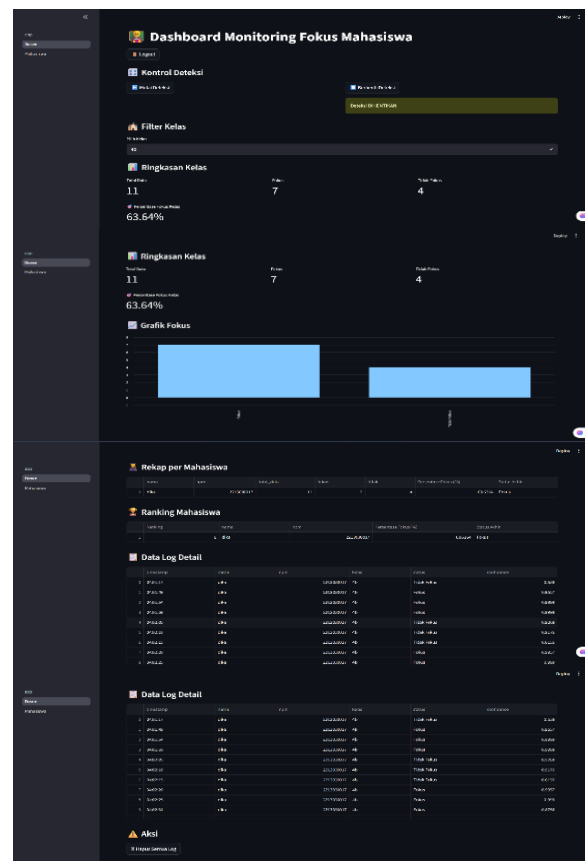
Gambar 7. halaman siswa

Pada gambar 7 untuk Halaman mahasiswa merupakan antarmuka utama dalam sistem yang digunakan untuk melakukan deteksi fokus secara *real-time*. Halaman ini dibangun menggunakan framework Streamlit dengan tampilan berbasis web yang interaktif dan mudah digunakan proses deteksi, yaitu:

- Input Data Mahasiswa, meliputi:
 - Nama
 - NPM
 - Kelas
 - Mata kuliah
 - Semester
- Pengaturan Sistem, yang terdiri dari:
 - Pemilihan kamera (index 0, 1, atau 2)
- Pemilihan model (MobileNetV3 atau YOLOv8)
- Pemilihan perangkat komputasi (CPU atau GPU)

Setelah seluruh data diisi, pengguna dapat menekan tombol “Mulai Deteksi” untuk memulai proses deteksi secara *real-time* menggunakan kamera. Sistem kemudian akan membaca input kamera dan melakukan proses deteksi wajah serta klasifikasi fokus menggunakan model yang telah dipilih.

Selain itu, tersedia tombol “Stop” yang berfungsi untuk menghentikan proses deteksi sehingga sistem tidak terus menerus melakukan pembacaan kamera.



Gambar 8. Dashboard Dosen

Pada Gambar 8 Dashboard dosen merupakan bagian dari sistem yang digunakan untuk memantau tingkat fokus mahasiswa secara keseluruhan berdasarkan data yang diperoleh dari proses deteksi pada sisi mahasiswa. Tampilan dashboard ini dirancang secara interaktif dan informatif sehingga memudahkan dosen dalam melakukan evaluasi.

- Fitur Kontrol Deteksi

Pada bagian atas dashboard terdapat fitur kontrol deteksi yang terdiri dari:

 - Tombol “Mulai Deteksi” → digunakan untuk mulai menerima data dari mahasiswa
 - Tombol “Berhenti Deteksi” → menghentikan proses pencatatan data.
- Filter Kelas

Fitur ini digunakan untuk memilih kelas tertentu (misalnya: 4B). Memungkinkan dosen untuk memantau fokus mahasiswa berdasarkan kelas tertentu sehingga analisis lebih terarah.

- c. Ringkasan Kelas
Menampilkan statistik utama:
 - Total data
 - Jumlah Fokus
 - Jumlah Tidak Fokus
 - Persentase fokus
- d. Grafik Fokus
Ditampilkan dalam bentuk grafik batang yang membandingkan:
 - Jumlah mahasiswa fokus
 - Jumlah mahasiswa tidak fokus
- e. Rekap per Mahasiswa
Menampilkan data agregat tiap mahasiswa:
 - Nama
 - NPM
 - Kelas
 - Jumlah fokus
 - Jumlah tidak fokus
 - Persentase fokus
 - Status akhir (Fokus / Tidak Fokus)
- f. Ranking Mahasiswa
 - Menampilkan urutan mahasiswa berdasarkan: Persentase fokus (dari tertinggi ke terendah)
- g. Berisi data mentah hasil deteksi:
 - Timestamp
 - Nama
 - NPM
 - Kelas
 - Status (Fokus / Tidak Fokus)
 - Confidence
- h. Fitur Aksi (Hapus Data)
Tombol:
 - Hapus Semua Log

Berdasarkan implementasi yang telah dilakukan, dashboard dosen mampu menampilkan informasi monitoring fokus mahasiswa secara real-time dan terstruktur. Fitur yang disediakan memungkinkan dosen untuk melakukan analisis baik secara keseluruhan maupun per individu mahasiswa. Selain itu, adanya fitur kontrol deteksi dan filter kelas memberikan fleksibilitas dalam penggunaan sistem sesuai kebutuhan proses pembelajaran.

5. KESIMPULAN DAN SARAN

Berdasarkan hasil penelitian yang telah dilakukan, sistem deteksi fokus peserta didik berbasis computer vision menggunakan arsitektur MobileNetV3 berhasil dikembangkan dan diimplementasikan secara real-time. Penelitian ini memanfaatkan orientasi kepala berupa yaw dan pitch sebagai indikator untuk mengklasifikasikan kondisi fokus dan tidak fokus peserta didik. Proses penelitian dilakukan menggunakan metode CRISP-DM yang meliputi tahap business understanding, data understanding, data preparation, modeling, evaluation, dan deployment. Dataset yang digunakan adalah AFLW2000-3D yang memiliki variasi pose wajah dalam kondisi nyata (in-the-wild), sehingga model

mampu beradaptasi pada berbagai kondisi arah pandangan wajah.

Hasil pengujian menunjukkan bahwa model MobileNetV3 memperoleh performa yang baik dengan nilai accuracy sebesar 93%, precision sebesar 0,84, recall sebesar 0,87, dan F1-score sebesar 0,85. Selain itu, model juga mampu berjalan secara real-time melalui implementasi sistem berbasis Streamlit menggunakan webcam sebagai input utama. Penelitian ini juga melakukan perbandingan performa dengan model YOLOv8. Berdasarkan hasil pengujian, YOLOv8 memiliki sensitivitas yang lebih tinggi terhadap perubahan arah wajah, namun MobileNetV3 menunjukkan performa yang lebih stabil, ringan, dan sesuai untuk implementasi real-time pada perangkat dengan spesifikasi menengah.

Implementasi sistem yang terdiri dari dashboard mahasiswa dan dashboard dosen juga berhasil menampilkan hasil monitoring fokus secara otomatis dan interaktif. Dashboard dosen mampu menampilkan rekapitulasi fokus mahasiswa, grafik monitoring, serta ranking tingkat fokus mahasiswa secara real-time. Dengan demikian, penelitian ini berhasil memenuhi tujuan penelitian, yaitu membangun sistem monitoring fokus peserta didik yang objektif, efisien, dan dapat diterapkan pada lingkungan pembelajaran digital.

Berdasarkan hasil penelitian yang telah dilakukan, sistem deteksi fokus peserta didik ini masih memiliki beberapa keterbatasan yang dapat dikembangkan pada penelitian selanjutnya. Pengembangan sistem dapat dilakukan dengan menambahkan fitur lain seperti deteksi mata, ekspresi wajah, dan eye tracking untuk meningkatkan akurasi dalam menentukan tingkat fokus peserta didik. Selain itu, pengembangan antarmuka pengguna (UI/UX) dan integrasi dengan database berbasis cloud juga dapat dilakukan agar sistem lebih optimal dan mendukung penggunaan multi-user secara bersamaan.

Penelitian selanjutnya juga disarankan menggunakan dataset yang lebih besar dengan variasi pengguna, kondisi pencahayaan, dan lingkungan yang lebih beragam agar kemampuan generalisasi model menjadi lebih baik. Selain itu, penelitian berikutnya dapat mengombinasikan metode lain seperti EfficientNet, ResNet, atau pendekatan hybrid deep learning untuk meningkatkan performa klasifikasi fokus secara real-time. Sistem yang dikembangkan juga berpotensi diterapkan secara langsung pada lingkungan pembelajaran daring, laboratorium komputer, maupun konsep smart classroom untuk membantu dosen dalam memonitor tingkat perhatian peserta didik secara lebih efektif dan objektif.

DAFTAR PUSTAKA

- [1] A. Rohmah and S. Pradikto, "Jurnal Cakrawala Pendidikan dan Biologi Volume. 2, Nomor. 4, Tahun," vol. 2025, pp. 42–53, doi: 10.61132/jucapenbi.v2i1.140.
- [2] M. Rahma Setyani, "ANALISIS TINGKAT KONSENTRASI BELAJAR SISWA DALAM

- PROSES PEMBELAJARAN MATEMATIKA DITINJAU DARI HASIL BELAJAR,” *Pendidikan Matematika*, vol. 01, 2018.
- [3] E. Simbolon and D. Amir Soleh, “Analisis Dampak Lingkungan Kelas terhadap Konsentrasi Belajar Siswa,” *Jurnal Studi Kemahasiswaan*, vol. 5, no. 1, pp. 116–128, 2025, doi: 10.54437/irsyaduna.
- [4] J. Wang, Y. Sun, and S. Tian, “Deep Learning for Student Behavior Detection in Smart Classroom Environments,” *Information (Switzerland)*, vol. 16, no. 11, Nov. 2025, doi: 10.3390/info16110949.
- [5] G. Fanelli, T. Weise, J. Gall, and L. Van Gool, “Real Time Head Pose Estimation from Consumer Depth Cameras.”
- [6] A. Howard *et al.*, “Searching for MobileNetV3.”
- [7] N. Zhou, R. Liang, and W. Shi, “A Lightweight Convolutional Neural Network for Real-Time Facial Expression Detection,” *IEEE Access*, vol. 9, pp. 5573–5584, 2021, doi: 10.1109/ACCESS.2020.3046715.
- [8] F. Marpaung, F. Aulia, N. Suryani SKom, and R. Cyra Nabila SKom, *COMPUTER VISION DAN PENGOLAHAN CITRA DIGITAL*. [Online]. Available: www.pustakaaksara.co.id
- [9] A Nugroho and SYLLA AYU KUSUMAHATI, “PENERAPAN CONVOLUTIONAL NEURAL NETWORK DENGAN TRANSFER LEARNING MOBILENETV2 PADA KLASIFIKASI PENYAKIT KULIT WAJAH,” 2022.
- [10] A. M. Iqbal, J. Jumadi, and E. Nurlatifah, “Implementasi YOLOv8 Sebagai Pendeteksi Nominal Uang Rupiah Kertas Berbasis Android,” *SMATIKA JURNAL*, vol. 15, no. 02, pp. 268–277, Dec. 2025, doi: 10.32664/smatika.v15i02.1545.
- [11] buulolo, “MOBILENETV3 DENGAN CBAM UNTUK KLASIFIKASI PENYAKIT DAUN KENTANG,” 2025.
- [12] Fitri p.t, “MODEL ATENSI CITRA WAJAH AUDIENS BERBASIS EDGE MACHINE LEARNING,” 2024.
- [13] Y. Imelda, L. Karmila Daulay, D. Paramitha Purba, and H. Syahputra, “JETBUS Journal of Education Transportation and Business Deteksi Gerakan Kepala Secara Real-Time Menggunakan OpenCV dan Python,” vol. 2, no. 1, p. 426