



PROGRAM STUDI
INFORMATIKA
PROGRAM SARJANA

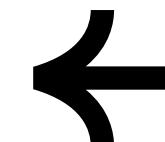
ANALISIS PERBANDINGAN KINERJA ALGORITMA YOLOV8 HINGGA YOLOV12 UNTUK DETEKSI EMOSI PADA EKSPRESI WAJAH

Rizky Prayogi Reksomulyo – 20523047

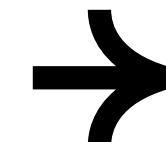
Dosen Pembimbing
Arrie Kurniawardhani, S.Si., M.Kom.

Latar Belakang

Batasan
Masalah



Pendahuluan



Tujuan
Penelitian



Rumusan
Masalah



Latar belakang

- Deteksi emosi menjadi bidang penelitian yang berkembang pesat karena perannya yang penting dalam interaksi manusia sehari-hari.
- Emosi memengaruhi keputusan, perilaku, dan komunikasi, sehingga pengenalannya penting untuk berbagai aplikasi seperti keamanan, interaksi manusia-komputer, dan pemantauan kesehatan mental.
- Studi sebelumnya menunjukkan bahwa model YOLO memiliki akurasi dan kecepatan yang lebih unggul dibandingkan dengan metode konvensional seperti Faster R-CNN.
- Dalam studi sebelumnya ada perbedaan kinerja di antara berbagai versi YOLO yang ditemukan.

Rumusan Masalah

Dari latar belakang, diketahui bahwa deteksi emosi adalah bidang penting dan algoritma YOLO menunjukkan performa yang unggul dalam tugas deteksi objek secara real-time. Penelitian sebelumnya telah membandingkan berbagai versi YOLO untuk tugas spesifik dan menunjukkan adanya perbedaan kinerja dalam hal akurasi, kecepatan dan efisiensi model. Oleh karena itu, muncul pertanyaan tentang bagaimana perbandingan kinerja YOLOv8 sampai YOLOv12 jika diterapkan pada deteksi emosi wajah.

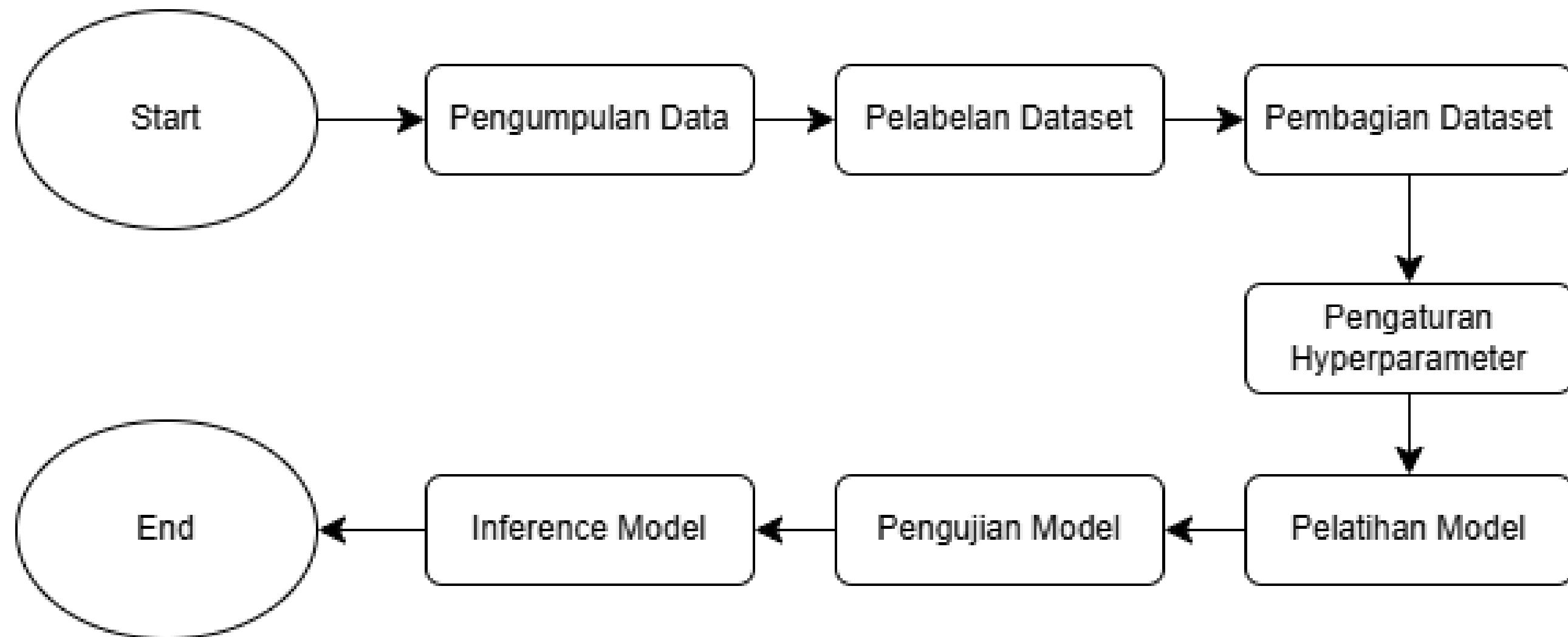
Tujuan Penelitian

Mengimplementasikan serangkaian model deteksi objek menggunakan algoritma YOLOv8, YOLOv9, YOLOv10, YOLOv11, dan YOLOv12 untuk mengenali berbagai ekspresi emosi pada wajah manusia.

Batasan Penelitian

- Menggunakan algoritma YOLO varian nano.
- Data yang diolah berupa gambar.
- Dataset yang digunakan diambil dari website dengan link “www.kdef.se”.
- Hanya menggunakan dataset dengan angle straight, left half, dan right half.
- Hyperparameter yang diatur hanya batch, epoch dan image sized.
- Inference menggunakan 7 gambar tiap emosi dari video dataset RAVDESS dan image hasil foto dari handphone.

Tahapan Pengerjaan



Pengumpulan Data

- Menggunakan dataset The Karolinska Directed Emotional Faces (KDEF)
- Berisi 4.900 gambar yang dikategorikan ke dalam tujuh kelas emosi yaitu angry, disgusted, fearful, happy, neutral, sad, dan surprised
- Terdapat 5 sudut yang berbeda yaitu profil kiri penuh, profil setengah kiri, lurus, profil setengah kanan, dan profil kanan penuh
- Ada 2 sesi pemotretan
- Data yang menggunakan hanya 1.470 gambar

Pelabelan Dataset

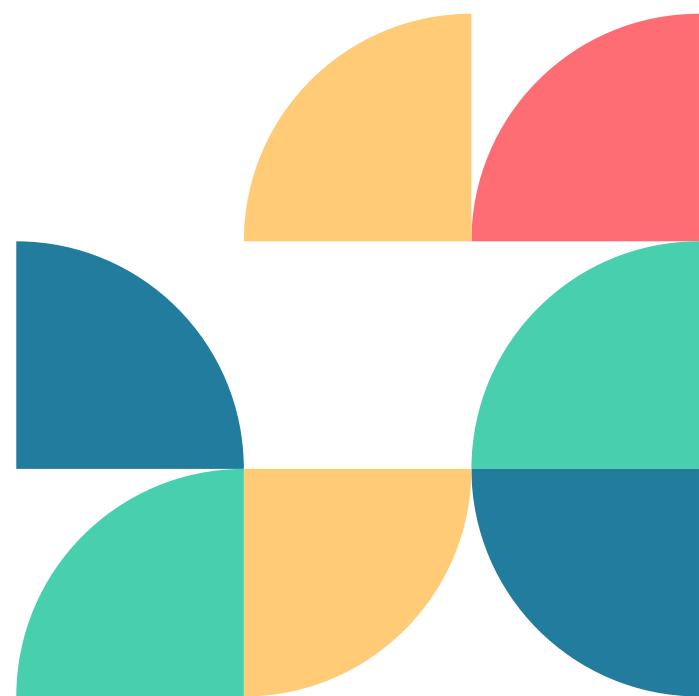
- Tujuan utamanya adalah untuk memberikan label kelas pada objek dan menentukan posisi kotak pada objek.
- Pelabelan dilakukan sendiri dengan platform roboflow
- Mengkategorikan ketujuh kelas emosi tersebut ke dalam label numerik

Emosi	fearful	angry	disgusted	happy	neutral	sad	surprised
Label	0	1	2	3	4	5	6

Pembagian Data

- pembagian menggunakan platform Roboflow.
- membagi dataset menjadi dua bagian utama: data pelatihan dan data validasi, dengan rasio 80:20. Sebanyak 1.175 gambar (80%) dialokasikan untuk pelatihan dan 293 gambar sisanya (20%) untuk validasi.

Kategori	Emosi							Total
	<i>Affraid</i>	<i>Angry</i>	<i>Disgust</i>	<i>Happy</i>	<i>Neutral</i>	<i>Sad</i>	<i>Surprise</i>	
Data latih	168	168	167	168	168	168	168	1175
Data validasi	42	42	42	42	41	42	42	293



Pengaturan Hyperparameter

- Mengatur tiga hyperparameter utama: ukuran batch, jumlah epoch, dan resolusi gambar.
- Ukuran batch sebesar 32 dipilih untuk menyeimbangkan efisiensi komputasi dengan lamanya waktu pelatihan.
- Jumlah epoch ditetapkan pada 100 untuk memberikan waktu yang cukup bagi model untuk belajar
- Resolusi gambar diatur pada 640x640 piksel, karena terbukti bahwa gambar beresolusi lebih tinggi memungkinkan ekstraksi fitur yang lebih kaya, yang secara langsung berkontribusi pada peningkatan nilai mAP.

Pelatihan Model

- Tahap ini model belajar dari gambar dan anotasi untuk mengidentifikasi pola dan fitur unik yang mendefinisikan setiap kelas objek.
- model yang digunakan pada pelatihan adalah model paling ringan dari setiap versi
- Proses pelatihan menghasilkan berbagai metrik evaluasi, termasuk loss, recall, precision, mAP, dan *inference time*.

Pengujian Model

- Tahap pengujian model dilakukan dengan menggunakan data validasi.
- Hasil dari tahap pengujian adalah confusion matrix untuk menilai kemampuannya dalam mendekripsi ekspresi wajah secara akurat berdasarkan kategori yang telah ditentukan.

Inference Model

- Tahap Inference model dilakukan dengan menggunakan gambar tujuh video emosi berbeda dari dataset Ravidess dan gambar hasil foto menggunakan handphone.
- Tujuannya untuk melihat kinerja pada data yang belum pernah dipelajari oleh model.



Hasil dan Pembahasan

Hasil Loss Grafik

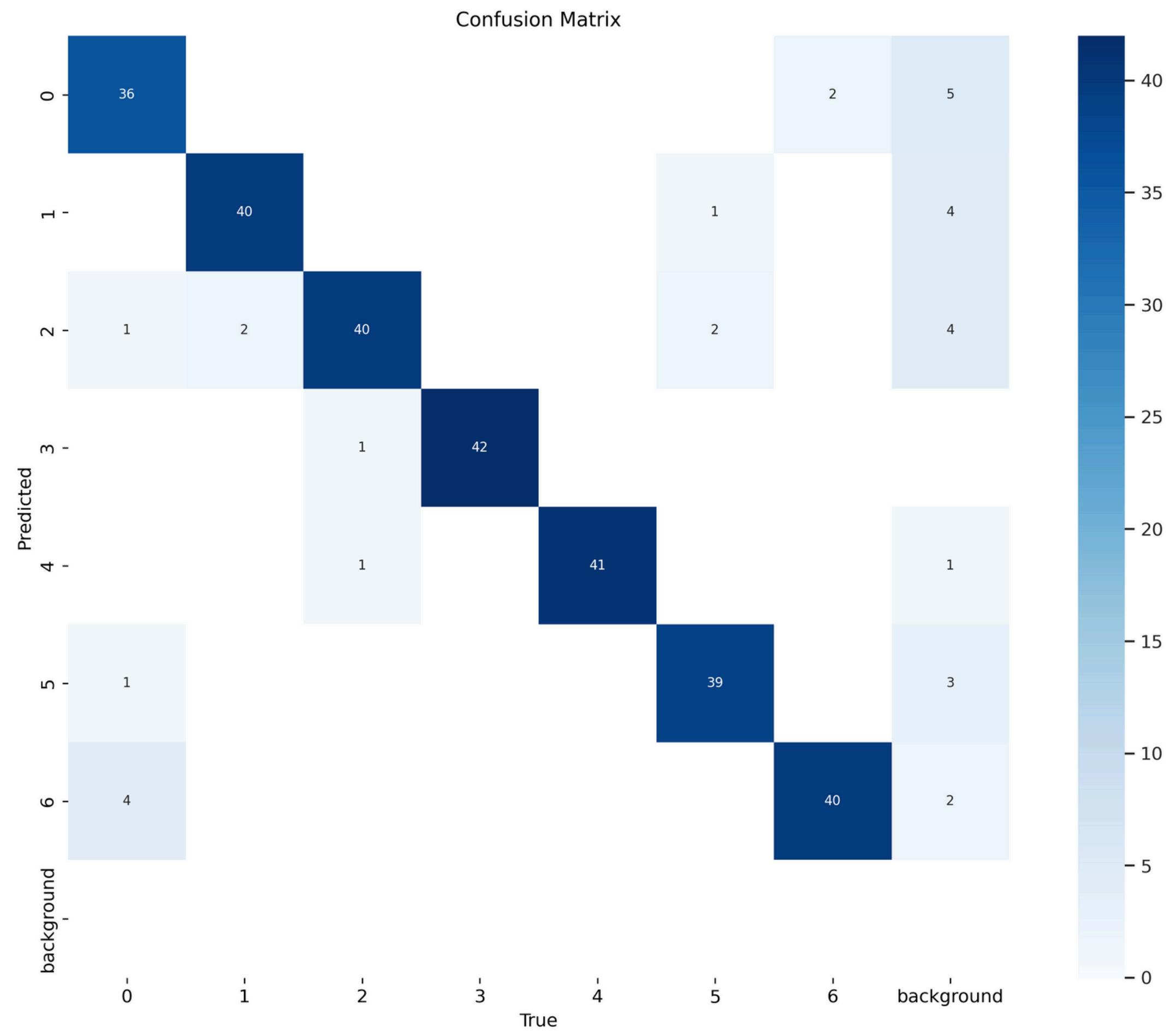
Model		Box_loss	Cls_loss	Dfl_loss
YOLOv12n	Train	0.396	0.197	0.921
	Val	0.563	0.261	1.049
YOLOv11n	Train	0.396	0.193	0.925
	Val	0.574	0.299	1.051
YOLOv10n	Train	0.45	0.241	0.853
	Val	0.565	0.301	0.901
YOLOv9t	Train	0.403	0.177	0.934
	Val	0.583	0.281	1.047
YOLOv8n	Train	0.365	0.187	0.91
	Val	0.578	0.302	1.05

Hasil Evaluasi

model	layers	parameters	precision	recall	mAP50	mAP50:90	Inference time
YOLOv12n	159	2,558,093	0.977	0.986	0.993	0.871	2.9ms
YOLOv11n	100	2,583,517	0.936	0.967	0.989	0.868	2.2ms
YOLOv10n	125	2,697,146	0.982	0.932	0.988	0.866	2.3ms
YOLOv9t	197	1,972,149	0.953	0.957	0.987	0.871	3.3ms
YOLOv8n	72	3,007,013	0.967	0.946	0.988	0.863	1.7ms

Hasil Confusion Matrix

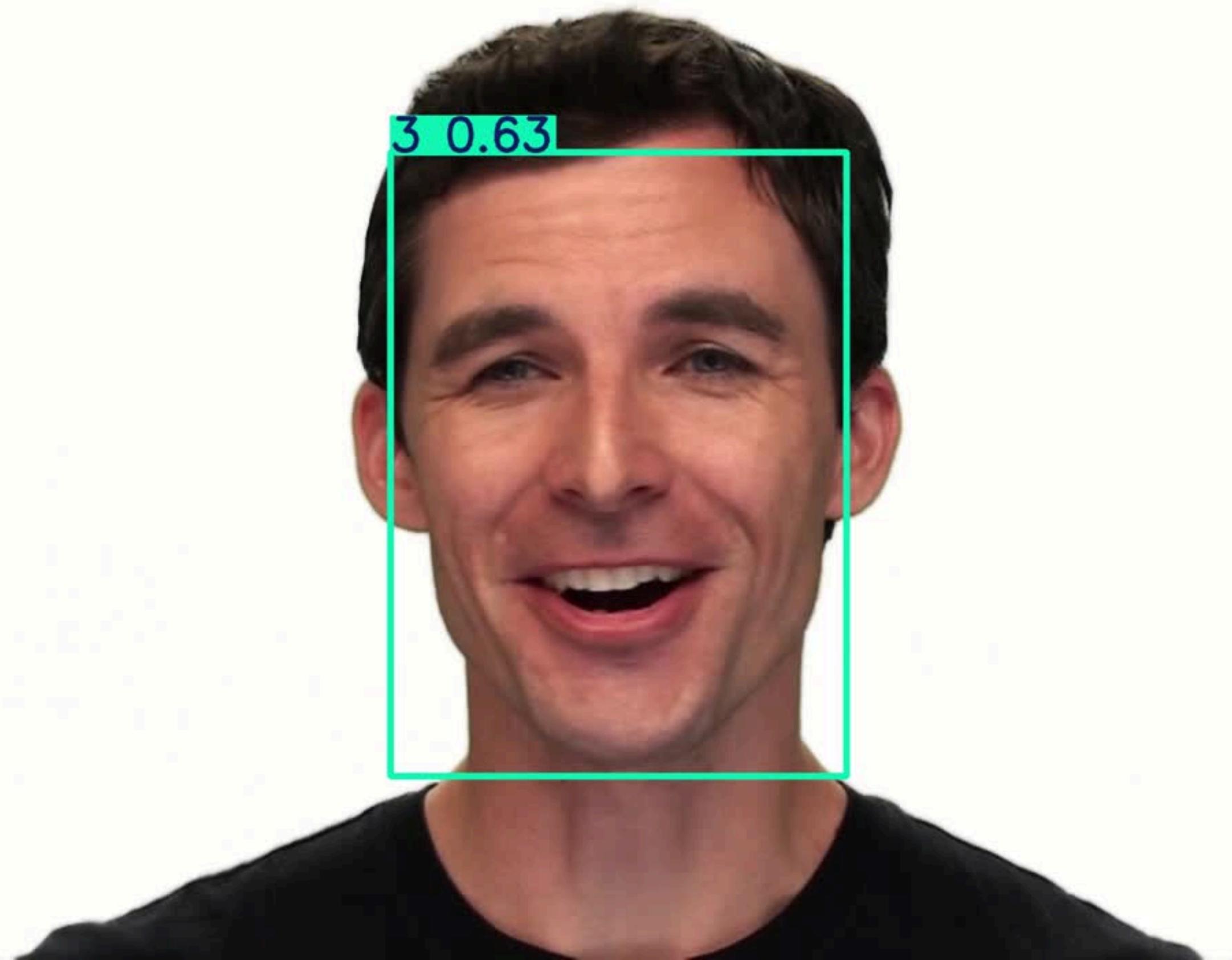




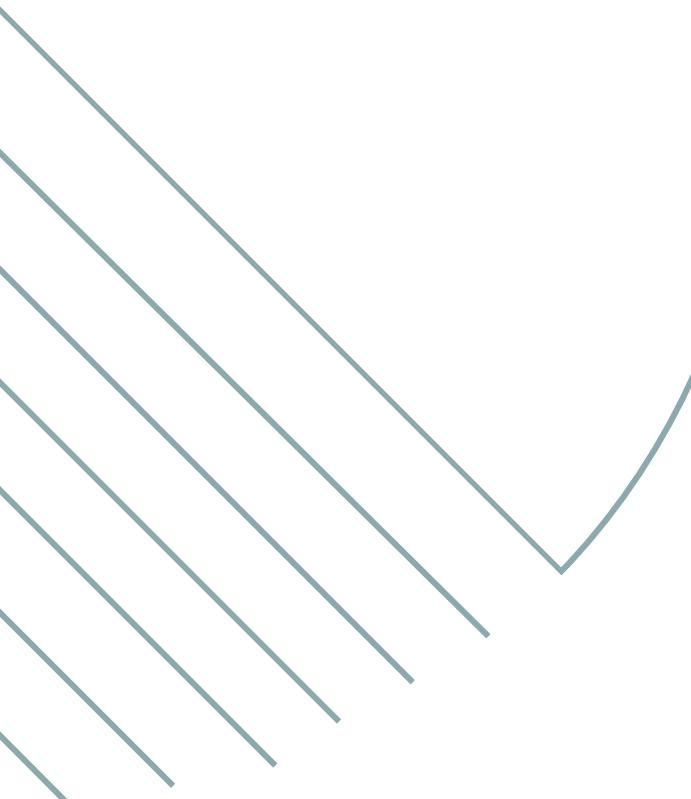


Hasil Inference dari Dataset Ravdess

Label	Emosi
4	neutral
3	happy
5	sad
1	angry
0	fearful
2	disgusted
6	surprised



Model		Emotion						
		0 (Affraid)	1 (Angry)	2 (Disgusted)	3 (happy)	4 (neutral)	5 (sad)	6 (surprised)
YOLOv12n	Confident score	0.65	0.88	0.93	0.94	0.93	0.93	0.94
	Hasil Klasifikasi	0 (Affraid)	1 (Angry)	2 (Disgusted)	3 (happy)	4 (neutral)	5 (sad)	6 (surprised)
YOLOv11n	Confident score	0.56	0.94	0.94	0.93	0.93	0.8	0.92
	Hasil Klasifikasi	0 (Affraid)	1 (Angry)	2 (Disgusted)	3 (happy)	4 (neutral)	5 (sad)	6 (surprised)
YOLOv10n	Confident score	0.71	0.97	0.94	0.91	0.85	0.92	0.7
	Hasil Klasifikasi	5 (sad)	1 (Angry)	2 (Disgusted)	3 (happy)	4 (neutral)	5 (sad)	6 (surprised)
YOLOv9t	Confident score	0.78	0.92	0.93	0.96	0.69	0.92	0.7
	Hasil Klasifikasi	1 (Angry)	1 (Angry)	2 (Disgusted)	3 (happy)	4 (neutral)	5 (sad)	6 (surprised)
YOLOv8n	Confident score	0.85	0.94	0.94	0.91	0.94	0.82	0.93
	Hasil Klasifikasi	1 (Angry)	1 (Angry)	2 (Disgusted)	3 (happy)	4 (neutral)	0 (Affraid)	6 (surprised)



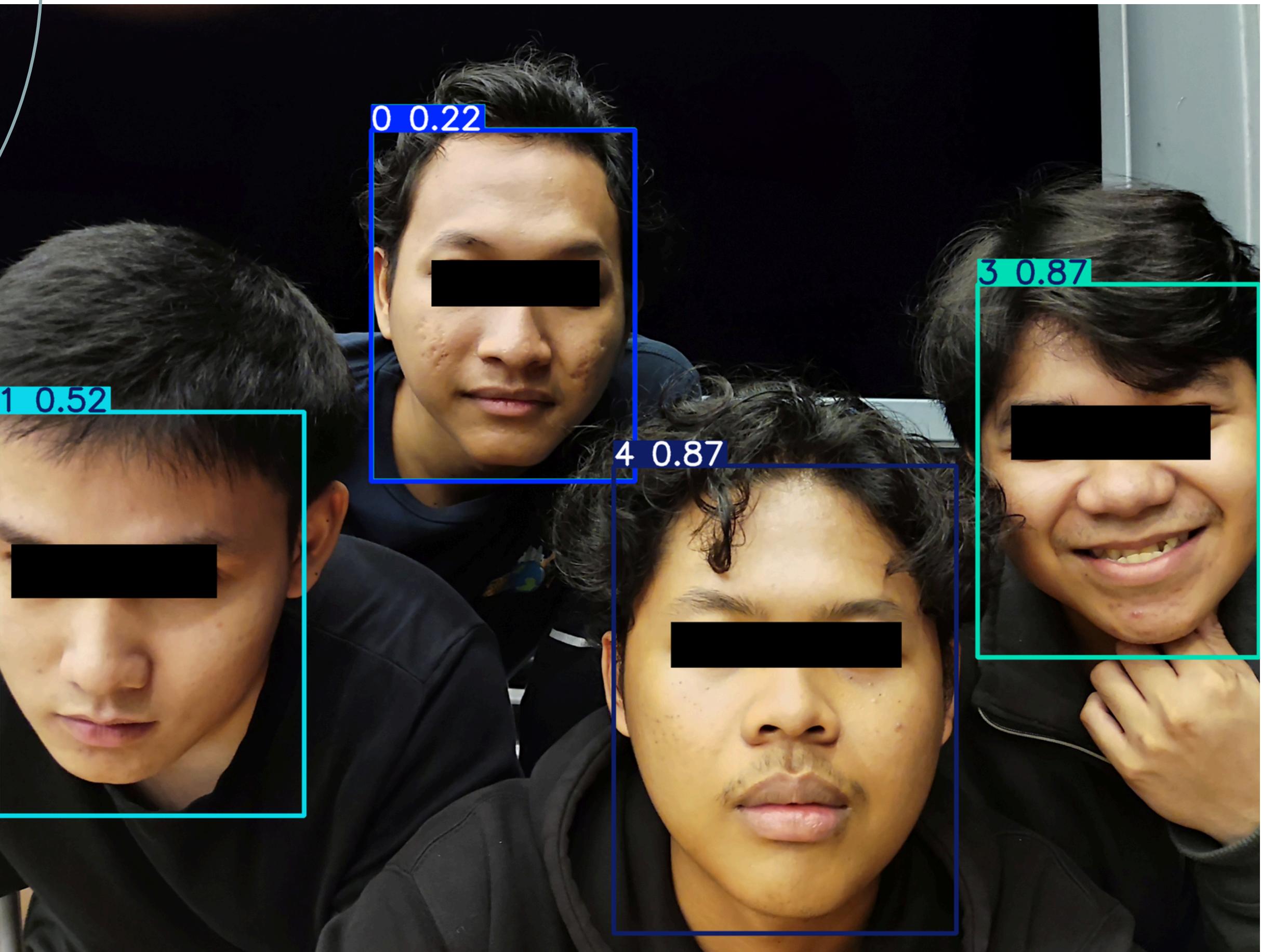
Hasil Inference dari Foto Handphone





Gambar yang digunakan adalah empat orang dengan ekspresi emosional yang telah diatur sebelumnya, dari kiri ke kanan: marah, sedih, netral, dan bahagia.

Label	Emosi
0	fearful
1	angry
2	disgusted
3	happy
4	neutral
5	sad
6	surprised



YOLOv8n

Model		Emotion			
		1 (Angry)	5 (sad)	4 (neutral)	3 (happy)
YOLOv12n	Confident score	0.77	0.57	0.73	0.75
	Hasil Klasifikasi	1 (Angry)	5 (sad)	4 (neutral)	2 (disgusted)
YOLOv11n	Confident score	0.91	0.08	0.76	0.49
	Hasil Klasifikasi	1 (Angry)	2 (disgusted)	1 (Angry)	2 (disgusted)
YOLOv10n	Confident score	0.43	0.49	0.92	0.08
	Hasil Klasifikasi	3 (happy)	5 (sad)	1 (Angry)	3 (happy)
YOLOv9t	Confident score	0.56	0.53	0.59	0.61
	Hasil Klasifikasi	3 (happy)	5 (sad)	1 (Angry)	2 (disgusted)
YOLOv8n	Confident score	0.52	0.22	0.87	0.87
	Hasil Klasifikasi	1 (Angry)	0 (affraid)	4 (neutral)	3 (happy)

Kesimpulan

Berdasarkan hasil penelitian perbandingan kinerja algoritma YOLOv8 hingga YOLOv12 untuk deteksi emosi pada ekspresi wajah menunjukkan adanya perbedaan dalam hal akurasi, kecepatan, dan efisiensi model. Hasil evaluasi dan analisis juga menunjukkan bahwa model YOLOv12n secara konsisten mendapatkan performa terbaik dalam tugas deteksi dan klasifikasi ekspresi wajah dari gambar. YOLOv12n mencatat nilai loss yang relatif seimbang antara data pelatihan dan validasi, yang mengindikasikan kemampuan generalisasi model yang baik dan minim overfitting. Selain itu, model ini juga meraih nilai recall dan mAP50 tertinggi di antara model lainnya, serta kinerja inferensi yang masih dapat diterima untuk pendekripsi *real-time* walaupun bukan yang terbaik. Hasil pengujian pada video RAVDESS dan foto dari *handphone* menunjukkan bahwa YOLOv12n mampu mendekripsi ekspresi emosional secara akurat, meskipun masih terdapat satu kesalahan klasifikasi. Dibandingkan model lain, YOLOv12n memperlihatkan kombinasi terbaik antara akurasi deteksi, *confident score*, dan ketahanan terhadap variasi kondisi seperti pencahayaan dan banyak wajah.