

**ANALISIS KOMPARATIF MODEL DEEP LEARNING
PRETRAINED UNTUK DETEKSI DINI AUTISME
PADA ANAK BERBASIS CITRA WAJAH**

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi S1-Informatika



disusun oleh

ARVI RIZQI FADHILA

22.11.4693

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2026

**ANALISIS KOMPARATIF MODEL DEEP LEARNING
PRETRAINED UNTUK DETEKSI DINI AUTISME
PADA ANAK BERBASIS CITRA WAJAH**

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi S1-Informatika



disusun oleh

ARVI RIZQI FADHILA

22.11.4693

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2026**

HALAMAN PERSETUJUAN

SKRIPSI

**ANALISIS KOMPARATIF MODEL DEEP LEARNING PRETRAINED
UNTUK DETEKSI DINI AUTISME PADA ANAK BERBASIS CITRA
WAJAH**

yang disusun dan diajukan oleh

ARVI RIZQI FADHILA

22.11.4693

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 27 Januari 2026

Dosen Pembimbing,

Dr. Andi Sunyoto, S.Kom., M.Kom.

NIK. 190302052

HALAMAN PENGESAHAN
SKRIPSI
ANALISIS KOMPARATIF MODEL DEEP LEARNING PRETRAINED
UNTUK DETEKSI DINI AUTISME PADA ANAK BERBASIS CITRA
WAJAH

yang disusun dan diajukan oleh

ARVI RIZQI FADHILA

22.11.4693

Telah dipertahankan di depan Dewan Penguji
pada tanggal 27 Januari 2026

Susunan Dewan Penguji

Nama Penguji

Majid Rahardi, S.Kom., M.Eng.
NIK. 190302393

Ike Verawati, S.Kom., M.Kom.
NIK. 190302237

Dr. Andi Sunyoto, S.Kom., M.Kom.
NIK. 190302052

Tanda Tangan



Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 27 Januari 2026

DEKAN FAKULTAS ILMU KOMPUTER



Prof. Dr. Kusriani, M.Kom.

NIK. 190302106

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Arvi Rizqi Fadhila
NIM : 22.11.4693

Menyatakan bahwa Skripsi dengan judul berikut:

ANALISIS KOMPARATIF MODEL DEEP LEARNING PRETRAINED UNTUK DETEKSI DINI AUTISME PADA ANAK BERBASIS CITRA WAJAH

Dosen Pembimbing: Dr. Andi Sunyoto, S.Kom., M.Kom

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 27 Januari 2026

Yang Menyatakan,



Arvi Rizqi Fadhila

HALAMAN PERSEMBAHAN

Dengan penuh rasa syukur kehadirat Allah SWT, yang telah melimpahkan rahmat dan karunia-Nya, skripsi ini penulis persembahkan kepada:

1. Kedua orang tua, kakek, nenek, dan adik penulis, terima kasih atas segala doa, dukungan, dan pengorbanan yang tiada henti. Kalian adalah alasan terbesar saya mampu menyelesaikan pendidikan hingga jenjang sarjana.
2. Saudara-saudara yang senantiasa memberikan dukungan, semangat, dan motivasi.
3. Sahabat dan teman-teman seperjuangan yang selalu memberikan semangat, bantuan, dan menemani dalam suka maupun duka selama masa perkuliahan.
4. Almamater tercinta, Universitas Amikom Yogyakarta, Terima kasih atas kesempatan menimba ilmu dan berkembang menjadi pribadi yang lebih baik.
5. Dan kepada diri sendiri, terima kasih untuk semua usaha dan kerja keras sehingga dapat bertahan dan tidak berhenti sehingga dapat menyelesaikan skripsi ini.

KATA PENGANTAR

Puji syukur kehadiran Allah SWT yang telah melimpahkan rahmat, hidayah, dan karunia-Nya sehingga penulis dapat menyelesaikan skripsi yang berjudul **"ANALISIS KOMPARATIF MODEL DEEP LEARNING PRETRAINED UNTUK DETEKSI DINI AUTISME PADA ANAK BERBASIS CITRA WAJAH"**. Shalawat serta salam semoga senantiasa tercurah kepada Nabi Muhammad SAW, keluarga, sahabat, dan para pengikutnya hingga akhir zaman.

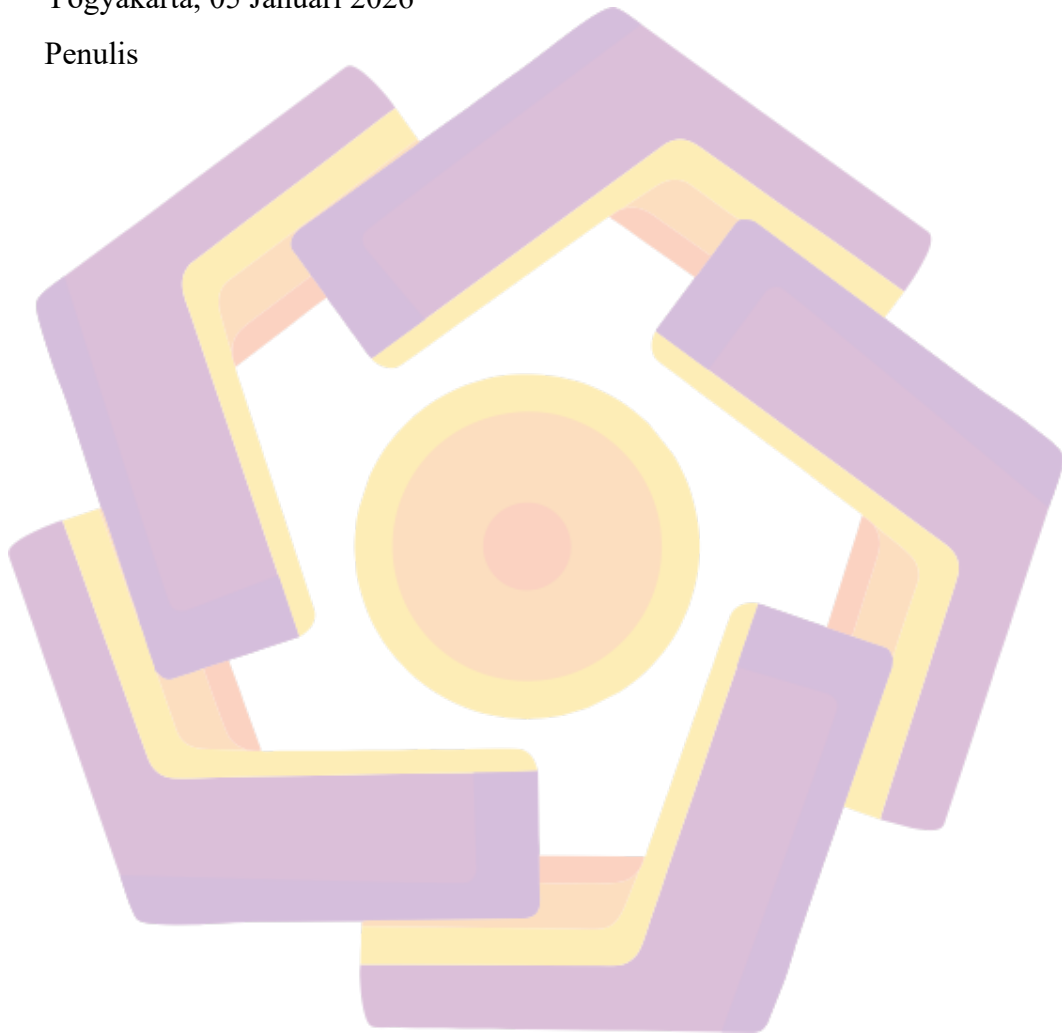
Skripsi ini disusun sebagai salah satu syarat untuk memperoleh gelar Sarjana Informatika pada Universitas Amikom Yogyakarta. Penulis menyadari bahwa dalam penyusunan skripsi ini tidak lepas dari bantuan, bimbingan, dan dukungan dari berbagai pihak. Oleh karena itu, pada kesempatan ini penulis ingin mengucapkan terima kasih yang sebesar-besarnya kepada:

1. Bapak Prof. Dr. Mohammad Suyanto, M.M., selaku Rektor Universitas Amikom Yogyakarta.
2. Ibu Prof. Dr. Kusrini, M. Kom., selaku Dekan Fakultas Ilmu Komputer Universitas Amikom Yogyakarta.
3. Ibu Eli Pujiastuti, M. Kom., selaku Ketua Program Studi S1 Informatika Universitas Amikom Yogyakarta.
4. Bapak Dr. Andi Sunyoto, S.Kom., M.Kom. selaku Dosen Pembimbing yang telah meluangkan waktu, tenaga, dan pikiran dalam memberikan bimbingan, arahan, serta masukan selama proses penyusunan skripsi ini.
5. Seluruh Bapak/Ibu dosen dan staf program studi S1 Informatika yang telah memberikan ilmu dan bantuan selama masa perkuliahan
6. Orang tua penulis yang telah memberikan dukungan moral dan material serta doa yang tidak pernah terputus selama penulis menempuh pendidikan.
7. Seluruh pihak yang telah membantu dan mendukung penyelesaian skripsi ini yang tidak dapat disebutkan satu per satu.

Penulis menyadari bahwa skripsi ini masih jauh dari sempurna. Oleh karena itu, penulis mengharapkan kritik dan saran yang membangun untuk perbaikan di masa mendatang. Semoga skripsi ini dapat bermanfaat bagi pembaca dan perkembangan ilmu pengetahuan.

Yogyakarta, 05 Januari 2026

Penulis



DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERSETUJUAN.....	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI	iv
HALAMAN PERSEMBAHAN	v
KATA PENGANTAR	vi
DAFTAR ISI.....	viii
DAFTAR TABEL.....	xi
DAFTAR GAMBAR.....	xii
DAFTAR RUMUS	xiv
DAFTAR LAMBANG DAN SINGKATAN	xv
DAFTAR ISTILAH	xvi
INTISARI	xvii
<i>ABSTRACT</i>	xviii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah.....	2
1.3 Batasan Masalah	2
1.4 Tujuan Penelitian	4
1.5 Manfaat Penelitian	4
1.6 Sistematika Penulisan	5
BAB II TINJAUAN PUSTAKA	7

2.1	Studi Literatur	7
2.2	Dasar Teori.....	12
BAB III METODE PENELITIAN		19
3.1	Objek Penelitian.....	19
3.2	Alur Penelitian	20
3.2.1	Dataset dan Akuisisi Data.....	21
3.2.2	Preprocessing Dataset	22
3.2.3	Augmentasi Data.....	23
3.2.4	Pembagian Data dan Strategi Sampling.....	25
3.2.5	Data Loader Configuration	25
3.2.6	Arsitektur Model.....	26
3.2.7	Pelatihan Model	29
3.2.8	Evaluasi.....	40
3.3	Alat dan Bahan.....	42
BAB IV HASIL DAN PEMBAHASAN		47
4.1	Gambaran Umum Eksperimen.....	47
4.1.1	Konfigurasi Eksperimen	48
4.1.2	Dataset dan Preprocessing	49
4.1.3	Pembagian Dataset dan Penyeimbangan Data	50
4.1.4	Augmentasi Data.....	50
4.1.5	Lingkungan dan Konfigurasi Eksperimen	51
4.1.6	Metode Evaluasi dan Strategi Pelatihan	51
4.2	Dinamika Pelatihan Model.....	52
4.2.1	Evaluasi Komparatif Tiga Arsitektur pada Phase 1	53
4.2.2	Evaluasi Komparatif Tiga Arsitektur pada Phase 2	53

4.2.3	Evaluasi Komparatif Tiga Arsitektur pada Phase 3	57
4.2.4	Ringkasan Pipeline Tiga Fase	59
4.3	Analisis Kinerja pada Data Pengujian	60
4.4	Analisis Pengaruh Preprocessing Terhadap Performa Model.....	64
4.4.1	Analisis Kinerja Pada Pelatihan	64
4.4.2	Analisis Kinerja Pada Data Tes	67
4.4.3	Analisis Komparasi Dengan Default Dataset.....	70
4.5	Analisis Mendalam Model MobileViTv2	73
4.5.1	Karakteristik Training MobileViTv2	73
4.5.2	Analisis Pola Kesalahan dan Implikasi Klinis	74
4.5.3	Perbandingan dengan Standar Klinis	76
4.6	Analisis Efisiensi Komputasi	77
BAB V PENUTUP		79
5.1	Kesimpulan	79
5.2	Saran	80
REFERENSI		82

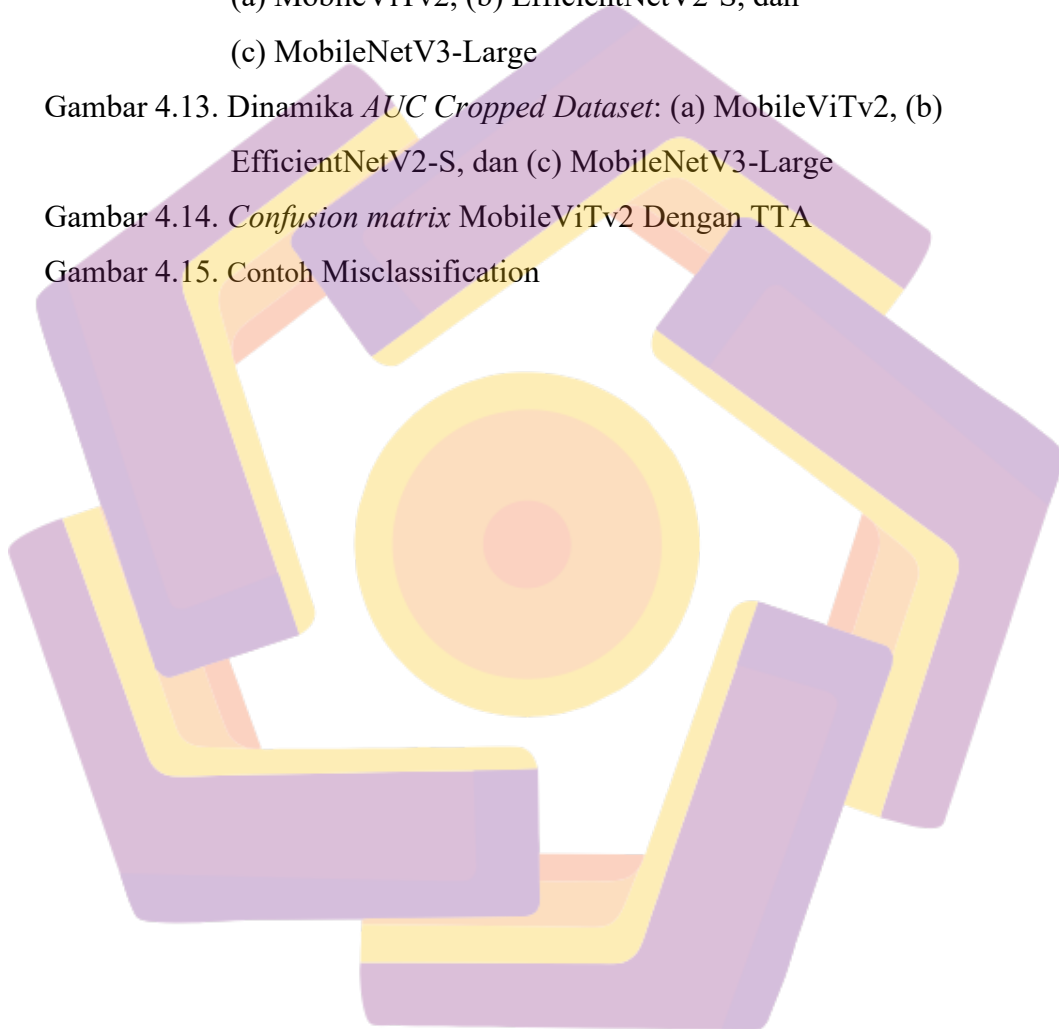
DAFTAR TABEL

Tabel 2.1. Keaslian Penelitian	9
Tabel 3.1. Karakteristik Dataset	22
Tabel 3.2. Contoh Hasil Hyperparameter Terbaik Optuna (MobileViTv2)	36
Tabel 4.1. Konfigurasi Eksperimen	48
Tabel 4.2. Perbandingan Baseline Ketiga Model	53
Tabel 4.3. Best Hyperparameters Hasil Optuna Ketiga Model	54
Tabel 4.4. Perbandingan Phase 2 Ketiga Model	56
Tabel 4.5. Konfigurasi Test-Time Augmentation	57
Tabel 4.6. Perbandingan Phase 3 (TTA) Ketiga Model	58
Tabel 4.7. Perbandingan Phase 2 dan Phase 3 (TTA)	59
Tabel 4.8. Perbandingan Confusion Matrix Tanpa TTA dan Dengan TTA	62
Tabel 4.9. Perbandingan Hasil Tanpa TTA dan Dengan TTA	63
Tabel 4.10. Perbandingan Hasil Phase 1 dan Phase 2 Cropped Dataset	66
Tabel 4.11. Perbandingan Confusion Matrix Tanpa TTA dan Dengan TTA Crop Dataset	67
Tabel 4.12. Perbandingan Hasil Tanpa TTA dan Dengan TTA Crop Dataset	68
Tabel 4.13. Perbandingan Hasil <i>Cropped Dataset</i> dan <i>Default Dataset</i>	70
Tabel 4.14. Perbandingan Kompleksitas Arsitektur dan Efisiensi Performa	77

DAFTAR GAMBAR

Gambar 2.1.	Ilustrasi cara kerja <i>deep learning</i>	13
Gambar 2.2.	Cara kerja MobileViT	14
Gambar 2.3.	Perbandingan arsitektur MobileViTv1 dan MobileViTv2	14
Gambar 3.1.	Alur Penelitian Sistematis Tiga Fase	20
Gambar 3.2.	Rincian Augmentasi Data	23
Gambar 3.3.	Rincian Konfigurasi Arsitektur	27
Gambar 3.4.	Rincian Pipeline Arsitektur	29
Gambar 3.5.	Sampel Gambar Kelas Non-Autistic Default Dataset	44
Gambar 3.6.	Sampel Gambar Kelas Autistic Default Dataset	45
Gambar 3.7.	Sampel Gambar Kelas Non-Autistic Cropped Dataset	45
Gambar 3.8.	Sampel Gambar Kelas Autistic Cropped Dataset	46
Gambar 4.1.	Cuplikan Kode face cropping MTCNN	48
Gambar 4.2.	Cuplikan Pipeline Augmentasi Data Validation dan Test	50
Gambar 4.3.	Cuplikan Pipeline Augmentasi Data Training	51
Gambar 4.4.	Cuplikan kode WeightedRandomSampler	51
Gambar 4.5.	Dinamika <i>training Accuracy</i> dan <i>validation Accuracy</i> untuk tiga model <i>lightweight</i> : (a) MobileViTv2, (b) EfficientNetV2-S, dan (c) MobileNetV3-Large	55
Gambar 4.6.	Dinamika <i>training Loss</i> dan <i>validation Loss</i> untuk tiga model <i>lightweight</i> : (a) MobileViTv2, (b) EfficientNetV2-S, dan (c) MobileNetV3-Large	55
Gambar 4.7.	Dinamika <i>AUC Progression</i> untuk tiga model setelah TTA: (a) MobileViTv2, (b) EfficientNetV2-S, dan (c) MobileNetV3-Large	55
Gambar 4.8.	Dinamika <i>AUC Progression</i> untuk tiga model dengan TTA: (a) MobileViTv2, (b) EfficientNetV2-S, dan (c) MobileNetV3-Large	58
Gambar 4.9.	Confusion Matrix Tanpa TTA: (a) MobileViTv2, (b) EfficientNetV2-S, dan (c) MobileNetV3-Large	60
Gambar 4.10.	Confusion Matrix Dengan TTA: (a) MobileViTv2,	

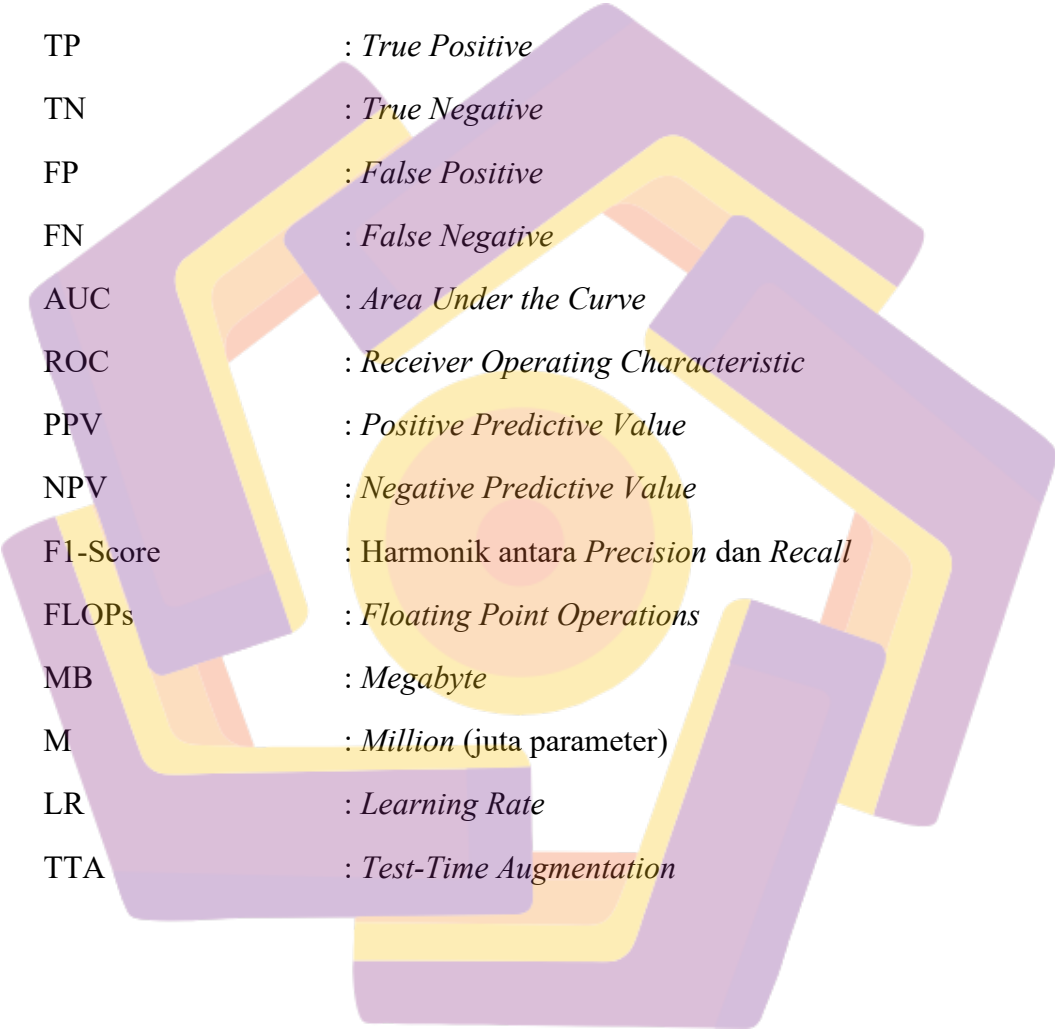
(b) EfficientNetV2-S, dan (c) MobileNetV3-Large	60
Gambar 4.11. Dinamika <i>Training Accuracy</i> dan <i>Validation Accuracy</i> <i>Cropped Dataset</i> : (a) MobileViTv2, (b) EfficientNetV2-S, dan (c) MobileNetV3-Large	65
Gambar 4.12. Dinamika <i>Training Loss</i> dan <i>Validation Loss</i> <i>Cropped Dataset</i> : (a) MobileViTv2, (b) EfficientNetV2-S, dan (c) MobileNetV3-Large	65
Gambar 4.13. Dinamika <i>AUC</i> <i>Cropped Dataset</i> : (a) MobileViTv2, (b) EfficientNetV2-S, dan (c) MobileNetV3-Large	65
Gambar 4.14. <i>Confusion matrix</i> MobileViTv2 Dengan TTA	73
Gambar 4.15. Contoh Misclassification	75



DAFTAR RUMUS

2.1. Rumus <i>Feature Map</i>	13
3.1. Rumus <i>Focal Loss</i>	32
3.2. Rumus <i>Mixup Augmentation</i>	32
3.3. Rumus <i>Weighted Average</i>	39
3.4. Rumus <i>True Positive Rate (TPR)</i>	40
3.5. Rumus <i>False Positive Rate (FPR)</i>	40
3.6. Rumus <i>Accuracy</i>	40
3.7. Rumus <i>Precision</i>	41
3.8. Rumus <i>Recall</i>	41
3.9. Rumus <i>Specificity</i>	41
3.10. Rumus <i>F1-Score</i>	41
3.11. Rumus <i>Binary Cross-Entropy (BCE) Loss</i>	42

DAFTAR LAMBANG DAN SINGKATAN



ASD	: <i>Autism Spectrum Disorder</i>
DL	: <i>Deep learning</i>
CNN	: <i>Convolutional Neural Network</i>
ViT	: <i>Vision Transformer</i>
TP	: <i>True Positive</i>
TN	: <i>True Negative</i>
FP	: <i>False Positive</i>
FN	: <i>False Negative</i>
AUC	: <i>Area Under the Curve</i>
ROC	: <i>Receiver Operating Characteristic</i>
PPV	: <i>Positive Predictive Value</i>
NPV	: <i>Negative Predictive Value</i>
F1-Score	: <i>Harmonik antara Precision dan Recall</i>
FLOPs	: <i>Floating Point Operations</i>
MB	: <i>Megabyte</i>
M	: <i>Million (juta parameter)</i>
LR	: <i>Learning Rate</i>
TTA	: <i>Test-Time Augmentation</i>

DAFTAR ISTILAH

ASD	: Gangguan perkembangan saraf.
<i>Deep learning</i>	: Pembelajaran mesin berbasis jaringan saraf.
CNN	: Model ekstraksi fitur citra.
Transformer	: Arsitektur dengan mekanisme self-attention.
MobileViTv2	: Model <i>lightweight</i> hybrid CNN–Transformer.
EfficientNetV2	: Model visi besar berperforma tinggi.
MobileNetV3	: Model CNN ringan untuk perangkat mobile.
Default Face	: Dataset utuh tanpa <i>preprocessing</i> .
<i>Cropped Dataset</i>	: <i>Preprocessing</i> wajah dengan <i>potongan</i> ketat.
Feature Adaptation	: Pelatihan awal <i>backbone</i> beku.
Fine-Tuning	: Pelatihan lanjut membuka sebagian <i>layer</i> .
<i>Confusion matrix</i>	: Tabel TP, TN, FP, FN.
TP	: Prediksi benar untuk kelas positif.
FN	: Positif yang diprediksi negatif.
FP	: Negatif yang diprediksi positif.
TN	: Prediksi benar untuk kelas negatif.
AUC	: Ukuran kemampuan pemisahan kelas.
Sensitivity / <i>Recall</i>	: Deteksi kasus positif.
Specificity	: Deteksi kasus negatif.
<i>Precision</i> (PPV)	: Akurasi prediksi positif.
NPV	: Akurasi prediksi negatif.
F1-Score	: Rata-rata harmonik <i>Precision</i> dan <i>Recall</i> .
FLOPs	: Beban komputasi model.
Model Size	: Ukuran file model.
AUC/Parameter Ratio	: Efisiensi performa per parameter.
Epoch	: Satu siklus penuh pelatihan dataset
Batch Size	: Jumlah sampel per iterasi pelatihan
Fine-Tuning	: Pelatihan lanjutan dengan membuka sebagian <i>layer</i>
Backbone	: Arsitektur inti ekstraksi fitur

INTISARI

Autism Spectrum Disorder (ASD) adalah gangguan perkembangan saraf yang memerlukan deteksi dini untuk intervensi yang efektif. Penelitian ini mengevaluasi kinerja dan efektivitas arsitektur ringan MobileViTv2 dalam klasifikasi ASD berbasis cintra wajah. Selain itu, dengan membandingkan dataset default (dataset asli) dan dataset yang dipotong (fokus pada ROI wajah), penelitian ini menganalisis dampak strategi preprocessing. Metode penelitian menggunakan kerangka eksperimental yang terdiri dari tiga fase. Pertama, pelatihan dasar; kedua, optimasi hyperparameter menggunakan framework Optuna dengan penyempurnaan parsial; dan ketiga, evaluasi akhir menggunakan Peningkatan Waktu Tes (TTA). 2.940 gambar wajah anak dengan pembagian stratified 70:15:15 digunakan. Akurasi, sensitivitas, spesifisitas, dan analisis efisiensi komputasi mendukung evaluasi dengan metrik AUC sebagai indikator utama. Hasil penelitian menunjukkan bahwa MobileViTv2 adalah model terbaik dengan nilai AUC tertinggi sebesar 0,9575 dan akurasi tertinggi sebesar 0,9002. Dengan parameter $\pm 5,8$ juta dan ukuran model ± 41 MB, MobileViTv2 mengungguli EfficientNetV2-S dan MobileNetV3-Large. Untuk menjaga keseimbangan antara sensitivitas dan spesifisitas selama proses preprocessing, penggunaan sampel default secara konsisten lebih baik daripada sampel cropped. Hasilnya menunjukkan bahwa konteks wajah utuh memberi model informasi visual yang lebih kaya. MobileViTv2 sangat cocok untuk digunakan sebagai sistem skrining awal ASD pada perangkat mobile karena efisiensi dan akurasi yang tinggi.

Kata kunci: *Autism Spectrum Disorder, MobileViTv2, deep learning, lightweight model, preprocessing, AUC, Test-Time Augmentation*

ABSTRACT

Autism Spectrum Disorder (ASD) is a neurodevelopmental disorder that requires early detection for effective intervention. This study evaluates the performance and effectiveness of the lightweight MobileViTv2 architecture in ASD classification based on facial images. Additionally, by comparing the default dataset (original dataset) and the cropped dataset (focusing on facial ROI), this study analyzes the impact of preprocessing strategies. The research method uses an experimental framework consisting of three phases. First, baseline training; second, hyperparameter optimization using the Optuna framework with partial refinement; and third, final evaluation using Test Time Augmentation (TTA). 2,940 children's facial images with a stratified split of 70:15:15 were used. Accuracy, sensitivity, specificity, and computational efficiency analysis support evaluation using the AUC metric as the primary indicator. The results show that MobileViTv2 is the best model with the highest AUC value of 0,9575 and the highest accuracy of 0,9002. With parameters of ± 5.8 million and a model size of ± 41 MB, MobileViTv2 outperforms EfficientNetV2-S and MobileNetV3-Large. To maintain a balance between sensitivity and specificity during the preprocessing process, the use of default samples is consistently better than cropped samples. The results show that the fullface context provides the model with richer visual information. MobileViTv2 is well suited for use as an initial ASD screening system on mobile devices due to its high efficiency and accuracy.

Keyword: Autism Spectrum Disorder, MobileViTv2, deep learning, lightweight model, preprocessing, AUC, Test-Time Augmentation