P-ISSN : 2580-4316
E-ISSN : 2654-8054
https://doi.org/10.37817/ikraith-informatika.v10i3

Rancang Bangun Aplikasi untuk Deteksi Alfabet Bahasa Isyarat Sibi menggunakan Algoritma Convolutional Neural Network (CNN)

[1]Fadly Maulana Adiyasa, [2]Angries
[1,2]Informatika, Universitas Persada Y.A.I, Jakarta Pusat

E-mail: [1]fadlyadiyasa9@gmail.com, [2]griezst@gmail.com

ABSTRAK

Komunikasi merupakan kebutuhan vital, namun penyandang tunarungu sering mengalami hambatan saat berinteraksi dengan masyarakat awam yang minim pemahaman terhadap bahasa isyarat. Penelitian ini bertujuan merancang bangun aplikasi penerjemah Sistem Isyarat Bahasa Indonesia (SIBI) ke dalam teks secara real-time berbasis Android guna mempermudah komunikasi tersebut. Sistem ini dikembangkan memanfaatkan kecerdasan buatan *deep learning* dengan algoritma *Convolutional Neural Network* (CNN) berarsitektur MobileNetV2 dan diintegrasikan dengan kerangka kerja MediaPipe HandLandmarker untuk melacak koordinat batas tangan. Proses klasifikasi berjalan secara offline di dalam perangkat menggunakan TensorFlow Lite untuk menerjemahkan 24 huruf alfabet SIBI yang bersifat statis. Berdasarkan pengujian langsung di lingkungan dengan kondisi ideal, sistem berhasil mencapai tingkat confidence rata-rata sebesar 95,19% pada smartphone dan 97,50% pada tablet. Secara keseluruhan, aplikasi telah berhasil beroperasi secara responsif tanpa terjadinya kebocoran memori (force close) dan tangguh dalam menangani gangguan dua tangan. Meskipun demikian, performa ekstraksi fitur sistem masih dapat menurun pada kondisi lingkungan ekstrem, seperti paparan cahaya silau (backlight) dan latar belakang visual yang ramai.

Kata Kunci: Android, Convolutional Neural Network (CNN), Deep Learning, MediaPipe, Sistem Isyarat Bahasa Indonesia (SIBI).

ABSTRACT

Communication is a vital necessity, however, the deaf community often faces barriers when interacting with the general public, who typically have limited understanding of sign language. This research aims to develop an Android-based application that translates the Indonesian Sign Language System (SIBI) into text in real-time to facilitate such communication. The system utilizes deep learning—specifically a Convolutional Neural Network (CNN) with a MobileNetV2 architecture—and integrates the MediaPipe HandLandmarker framework to track hand landmark coordinates. Classification occurs offline on the device using TensorFlow Lite to translate the 24 static letters of the SIBI alphabet. Testing in ideal environments demonstrated that the system achieved an average confidence level of 95.19% on smartphones and 97.50% on tablets. Overall, the application operates responsively without memory leaks (crashes) and robustly handles two-handed gestures. Nevertheless, feature extraction performance may decline under extreme environmental conditions, such as strong backlighting or cluttered visual backgrounds.

Keywords: ***Android, Convolutional Neural Network (CNN), Deep Learning, MediaPipe, Indonesian Sign Language System (SIBI).***

https://doi.org/10.37817/ikraith-informatika.v10i3
P-ISSN : 2580-4316
E-ISSN : 2654-8054

1. PENDAHULUAN

Komunikasi memainkan peran yang sangat penting bagi manusia sebagai makhluk sosial dalam menjalani aktivitas sehari-hari. Komunikasi bukan sekadar alat interaksi, melainkan sebuah kebutuhan hidup yang memungkinkan individu untuk membangun ikatan, saling memahami, serta bertukar informasi dan emosi dengan orang lain dalam berbagai situasi sosial. Maka dari itu, setiap orang perlu memiliki keterampilan berkomunikasi yang efektif agar bisa menjalani kehidupan sosialnya dengan baik dan bermakna, termasuk bagi mereka yang memiliki keterbatasan fisik maupun sensorik seperti penyandang tunarungu.

Pentingnya komunikasi untuk penyandang tunarungu tidak hanya terletak pada aspek sosial, tetapi juga dalam peningkatan kualitas hidup, akses pendidikan, dan keterlibatan penuh dalam masyarakat. Fungsi komunikasi bagi mereka menjadi sangat vital sebagai sarana utama untuk mengekspresikan keinginan, menyampaikan perasaan, serta mengakses dan menyebarkan informasi yang sangat diperlukan dalam aktivitas sehari-hari. Namun, seringkali mereka menghadapi tantangan besar, terutama individu yang mengalami kesulitan dalam mendengar dan berbicara, karena keterbatasan tersebut dapat menjadi penghalang yang cukup berarti dalam melakukan komunikasi baik secara lisan maupun nonverbal.

Bahasa isyarat berperan sebagai alat komunikasi yang sangat penting bagi individu dengan disabilitas pendengaran dalam kehidupan sehari-hari untuk menjaga komunikasi yang efektif dengan orang-orang di sekitarnya. Bahasa isyarat berfungsi sebagai alat penghubung komunikasi yang penting bagi penyandang tunarungu dan individu dengan disabilitas pendengaran lainnya, memberi mereka kesempatan untuk menyampaikan gagasan, emosi, serta kebutuhan dengan lebih jelas dan mudah dimengerti oleh orang lain, sehingga menciptakan interaksi yang nyaman dan saling menghormati. Menurut data yang dipublikasikan oleh Kementerian Sosial Indonesia pada tahun 2023, terungkap bahwa jumlah individu dengan disabilitas tunarungu di Indonesia mencapai sekitar 22,97 juta orang. Dalam konteks nasional, salah satu pedoman komunikasi utama yang dibakukan oleh pemerintah untuk memfasilitasi interaksi pendidikan dan sosial bagi penyandang tunarungu adalah Sistem Isyarat Bahasa Indonesia (SIBI). Namun, minimnya pemahaman masyarakat umum terhadap abjad dan gestur SIBI seringkali memutus rantai komunikasi tersebut. Angka yang signifikan ini dengan jelas menunjukkan perlunya pengembangan metode komunikasi yang inklusif dan teknologi yang bisa mendukung kehidupan sosial serta akses informasi bagi individu dengan disabilitas ini secara lebih efisien.

Untuk mengatasi masalah ini, dirancang sebuah aplikasi untuk menerjemahkan alfabet bahasa isyarat dengan memanfaatkan *machine learning* dan bekerja pada platform Android. Upaya pembuatan aplikasi penerjemah huruf bahasa isyarat ini bertujuan untuk menawarkan solusi teknologi yang modern dan mudah diakses melalui smartphone Android, sehingga memudahkan dalam menerjemahkan ejaan abjad bahasa isyarat yang biasa digunakan oleh penyandang tunarungu ke dalam teks. Diharapkan aplikasi ini bisa berfungsi sebagai sarana yang mendukung kegiatan sosial serta meningkatkan akses komunikasi di berbagai tempat, mulai dari rumah, sekolah, hingga tempat kerja.

Aplikasi ini direncanakan akan menggunakan kamera smartphone Android untuk merekam gerakan tangan pengguna. Gerakan tangan tersebut akan dianalisis secara langsung menggunakan model *deep learning* yang dikembangkan dengan algoritma *Convolutional Neural Network* (CNN) berbasis TensorFlow, sebuah teknologi *machine learning* untuk pengenalan pola dan klasifikasi gambar. Model CNN ini akan mengolah data gambar dari gerakan tangan pengguna dan mengubahnya menjadi teks huruf, menciptakan interaksi yang mulus antara orang dengan disabilitas dan lingkungan mereka, sambil memberikan pengalaman komunikasi yang lebih alami dan bebas dari batasan teknologi yang rumit.

2. LANDASAN TEORI

Bahasa Isyarat

Bahasa isyarat adalah bahasa yang tidak mengeluarkan bunyi melainkan menggunakan gerakan tangan dalam menyampaikan informasi. Bahasa ini digunakan oleh penyandang tunarungu atau Tuli dan tunawicara. Penyandang tunarungu adalah orang yang memiliki kekurangan dalam pendengaran, sementara

P-ISSN : 2580-4316
E-ISSN : 2654-8054
https://doi.org/10.37817/ikraith-informatika.v10i3

tunawicara adalah orang yang memiliki kekurangan dalam berbicara (Azizah et al., 2023)

SIBI

SIBI atau Sistem Isyarat Bahasa Indonesia adalah bahasa yang telah ditetapkan oleh pemerintah sebagai bahasa isyarat resmi di Indonesia dan sebagai standar acuan untuk proses pendidikan di Sekolah Berkebutuhan Khusus (SLB). Sistem bahasa ini awalnya dikembangkan oleh masyarakat sebagai upaya untuk menjembatani kebutuhan komunikasi dan pembelajaran akademis (Fatmawati et al., 2022)

***Convolutional Neural Network* (CNN)**

Convolutional Neural Network (CNN) adalah salah satu algoritma *deep learning* yang diperkenalkan pada tahun 1960 dan banyak digunakan dalam *computer vision*. CNN banyak digunakan untuk menyelesaikan masalah visual yang membutuhkan kemampuan komputasi tinggi sehingga cocok untuk masalah klasifikasi gambar, segmentasi, deteksi objek (Purwono et al., 2022). CNN bekerja dengan mengikuti cara kerja neuron pada otak manusia dengan membentuk rangkaian sel yang kompleks untuk disimulasikan oleh CNN. CNN memiliki fitur yang membagi bobot untuk mengurangi parameter sehingga lebih terhindar dari overfitting. Kemampuan lain dari CNN adalah mempelajari atau memproses klasifikasi secara bersamaan sehingga model yang keluar lebih terorganisir dan bergantung pada fitur yang diekstrak (Ramadhani et al., 2023)

MediaPipe

MediaPipe adalah sebuah framework yang bisa digunakan di berbagai platform, yang memungkinkan para pengembang membuat alur kerja *machine learning* (ML) yang bisa menangani berbagai jenis data. Ini termasuk pemrosesan video, audio, dan juga data deret waktu lainnya. Framework ini memiliki banyak model untuk mendeteksi dan melacak tubuh manusia. Model-model ini telah dilatih dengan menggunakan kumpulan data yang sangat besar dan bervariasi dari Google. Dengan menggunakan struktur kerangka yang terdiri dari simpul, tepi, atau landmark, MediaPipe dapat mengikuti titik-titik penting di berbagai bagian tubuh. Semua titik koordinatnya sudah dinormalisasi dalam bentuk tiga dimensi. Model yang dibuat dengan TensorFlow Lite ini memungkinkan aliran informasi yang fleksibel, sehingga mudah untuk disesuaikan dan diubah melalui tampilan grafik. Secara teknis, pipeline MediaPipe terdiri dari beberapa simpul yang membentuk sebuah grafik, yang biasanya didefinisikan dalam berkas pbtxt. Simpul-simpul tersebut kemudian dihubungkan dengan berkas berbasis C++, di mana pengembangannya bertumpu pada kelas kalkulator dasar (base calculator class) yang terdapat di dalam MediaPipe (Halder & Tayade, 2021).

TensorFlow

TensorFlow merupakan sebuah framework open-source yang dirancang dan dikembangkan secara khusus oleh Google untuk memfasilitasi proses pengembangan dan pelatihan model *machine learning*. Sistem ini menyediakan serangkaian perangkat dan library yang lengkap untuk membangun berbagai arsitektur model, termasuk neural network. Keunggulan utama TensorFlow terletak pada kemampuannya mengeksekusi komputasi numerik berkinerja tinggi melalui representasi data berbasis tensor, serta keluwesannya dalam mendistribusikan beban kerja pelatihan model ke berbagai platform komputasi, seperti CPU, GPU, maupun TPU (Alfiansyah & Litanianda, 2024).

Deep Learning

Deep Learning adalah cabang dari *machine learning* yang menggunakan neural network, yaitu sebuah model komputasi yang terinspirasi oleh cara kerja dan struktur dari otak manusia untuk menyelesaikan sejumlah permasalahan yang kompleks. Berbeda dengan algoritma *machine learning* tradisional yang cenderung lebih efisien dalam menganalisis data terstruktur seperti yang terdapat dalam pemodelan finansial atau pemasaran, *deep learning* menawarkan keunggulan yang jauh lebih signifikan dalam mengolah data yang tidak terstruktur dan tanpa label, contohnya teks, gambar, dan suara. Kemampuan untuk memproses data yang tidak terstruktur ini menjadikan *deep learning* sangat andal dan efisien untuk diterapkan pada tugas-tugas komputasi yang lebih rumit, termasuk dalam pengenalan objek dan pengenalan suara, bahkan juga dalam pengembangan sistem kendaraan tanpa pengemudi (Sharifani & Amini, 2023).

MobileNetV2

MobileNetV2 adalah salah satu arsitektur *Convolutional Neural Network* (CNN)

https://doi.org/10.37817/ikraith-informatika.v10i3
P-ISSN : 2580-4316
E-ISSN : 2654-8054

yang sangat populer dalam bidang pemrosesan dan klasifikasi citra. Pada arsitektur dasarnya, MobileNetV2 awalnya memiliki sekitar 1.000 titik pada lapisan klasifikasinya. Arsitektur ini memiliki struktur inverse residual dan struktur linear bottleneck yang berfungsi untuk membantu mengurangi kalkulasi pada proses konvolusi. Arsitektur ini lebih disukai dibandingkan model lain karena memiliki arsitektur yang sederhana dan karakteristik yang sangat efisien dalam penggunaan memori. Model ini secara khusus dirancang dan dikembangkan untuk aplikasi pada perangkat seluler serta lingkungan yang memiliki keterbatasan sumber daya komputasi (Gulzar, 2023).

Computer Vision

Computer Vision merupakan cabang dari kecerdasan buatan yang memberikan kemampuan kepada komputer untuk memahami dan menganalisis dunia visual melalui beragam langkah seperti pengambilan gambar, pemrosesan data, pengenalan gambar, serta pengambilan keputusan. *Computer Vision* berusaha meniru cara sistem penglihatan manusia yang rumit terutama bagaimana manusia mengamati objek melalui penglihatan dan mengirimkan gambar yang diterima ke otak untuk dianalisis, sehingga mereka bisa mengenali apa yang mereka lihat. Penafsiran ini kemudian menjadi dasar untuk proses pengambilan keputusan. Dalam *computer vision*, mesin mampu mengidentifikasi dan mengklasifikasikan objek dengan tepat memanfaatkan gambar digital dan video yang diperoleh dari kamera, yang diintegrasikan dengan model *deep learning* sehingga sistem bisa beraksi berdasarkan apa yang dilihat (Fachrisyam et al., 2025).

3. METODOLOGI

Metode penelitian yang digunakan dalam pengembangan sistem ini menerapkan pendekatan *prototyping*. Pendekatan ini dipilih karena memungkinkan siklus pengembangan yang berkelanjutan, sehingga evaluasi dan penyempurnaan dapat dilakukan secara bertahap. Tahapan penelitian ini meliputi:

Studi literatur

Mengkaji berbagai literatur akademik dan dokumentasi teknis terkait visi komputer (*computer vision*), *Deep Learning*, dan arsitektur *Convolutional Neural Network* (CNN). Kajian ini difokuskan pada metode prapemrosesan citra, optimalisasi model klasifikasi, konversi model jaringan saraf tiruan ke format TensorFlow Lite, serta teknik integrasi model *machine learning* ke dalam lingkungan *native* Android Studio. Pengumpulan dan Persiapan dataset Pengumpulan dataset meliputi pengumpulan gambar pose dalam bahasa isyarat berdasarkan pedoman SIBI dan persiapan dataset meliputi pengategorian dan pelabelan.

Pengumpulan dan Persiapan *dataset*

Tahap ini mengumpulkan data primer berupa 7.680 citra digital yang merepresentasikan gestur 24 abjad statis Sistem Isyarat Bahasa Indonesia (SIBI). Tahap persiapan (preprocessing) diawali dengan mengisolasi 20% dari total dataset, yaitu sebanyak 1.536 gambar (64 gambar per kelas abjad), murni sebagai data pengujian (testing data) yang tidak dilibatkan dalam proses pelatihan.

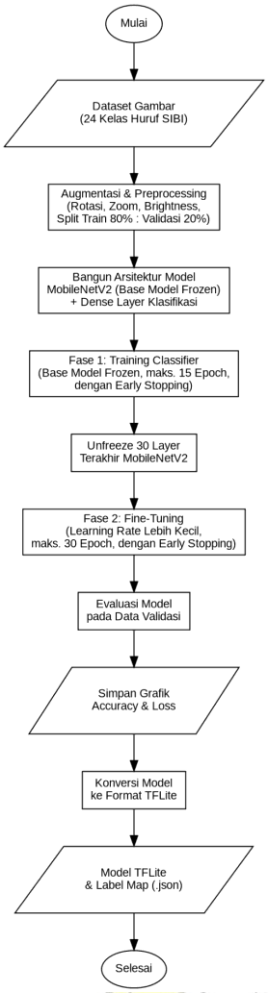
Sisa dataset sejumlah 6.144 gambar kemudian dialokasikan untuk proses pemodelan dengan pembagian rasio 80:20, menghasilkan 4.913 data latih (training data) dan 1.231 data validasi (validation data). Langkah prapemrosesan selanjutnya mencakup normalisasi piksel, pelabelan kelas, serta penerapan teknik augmentasi (seperti rotasi dan penyesuaian kontras) secara eksklusif pada data latih. Pendekatan ini bertujuan memperkaya variasi visual sehingga model memiliki ketahanan (robustness) yang tinggi terhadap kondisi pencahayaan dan perbedaan latar belakang di lingkungan nyata.

Perancangan Sistem

Sistem klasifikasi dirancang menggunakan arsitektur *Convolutional Neural Network* (CNN). Proses diawali dengan prapemrosesan dan augmentasi dataset gambar (24 kelas abjad SIBI statis) yang dibagi menjadi 80% data latih dan 20% data validasi. Pelatihan model menerapkan metode *transfer learning* menggunakan *base model* MobileNetV2 dan dieksekusi dalam dua fase. Fase pertama (*Training Classifier*) melatih lapisan klasifikasi tambahan selama maksimal 15 *epoch* dengan membekukan lapisan dasar. Fase kedua (*Fine-Tuning*) membuka 30 lapisan terakhir MobileNetV2 untuk dilatih ulang menggunakan *learning rate* yang lebih kecil selama maksimal 30 *epoch*. Kedua fase menggunakan *early*

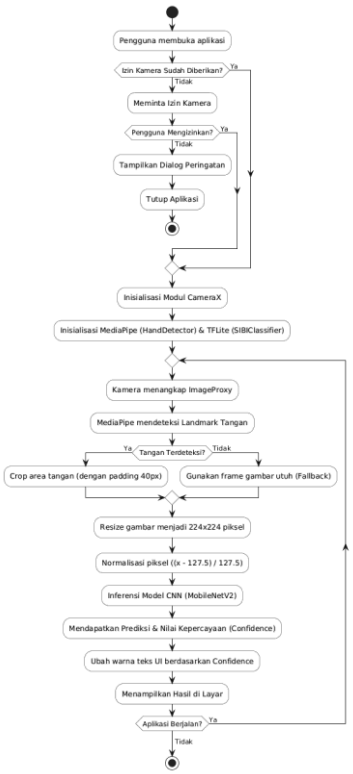
P-ISSN : 2580-4316
E-ISSN : 2654-8054
https://doi.org/10.37817/ikraith-informatika.v10i3

stopping untuk mencegah *overfitting*. Model terbaik kemudian dievaluasi dan dikonversi ke format TensorFlow Lite (.tflite) beserta *label map* (.json) agar dapat dijalankan pada perangkat Android.



Gambar 1. Training CNN

Alur aplikasi dimulai saat aplikasi dijalankan, sistem secara otomatis memverifikasi izin akses kamera. Apabila ditolak, aplikasi akan ditutup untuk mencegah kegagalan sistem. Jika diizinkan, modul CameraX, MediaPipe (*HandDetector*), dan model klasifikasi TFLite diinisialisasi ke dalam memori. Dalam siklus kerjanya, kamera menangkap *frame* secara berkelanjutan untuk dianalisis oleh MediaPipe. Jika *landmark* tangan terdeteksi, sistem melakukan pemotongan gambar (*cropping*) di area tangan dengan *padding* 40 piksel. Jika gagal, sistem menggunakan *frame* utuh sebagai langkah *fallback*. Gambar tersebut kemudian direduksi ukurannya menjadi 224x224 piksel, dinormalisasi, dan diinferensi oleh model CNN. Hasil prediksi dengan nilai keyakinan (*confidence*) tertinggi langsung ditampilkan untuk memperbarui antarmuka secara *real-time*.



Gambar 2. Alur Aplikasi

Implementasi

Menerjemahkan rancangan konseptual ke dalam baris kode program. Implementasi ini terbagi menjadi dua proses utama yaitu, pembangunan dan pelatihan model CNN menggunakan bahasa pemrograman Python dengan *library* TensorFlow dan Keras. Lalu pengembangan antarmuka pengguna aplikasi Android dan integrasi model deteksi melalui fungsi kamera menggunakan bahasa pemrograman Kotlin. Sistem yang telah rampung kemudian dievaluasi melalui tahap pengujian terhadap program yang sudah dibuat dan melakukan perbaikan bila ada kesalahan.

Dokumentasi

Menyusun laporan dari keseluruhan rangkaian pengembangan. Tahapan ini mencakup pencatatan landasan metode, rekaman performa pengujian model, serta penyusunan pedoman teknis aplikasi sebagai acuan referensi untuk pengembangan atau penelitian lanjutan.

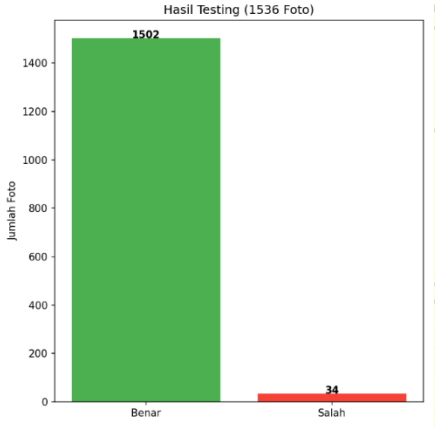
https://doi.org/10.37817/ikraith-informatika.v10i3
P-ISSN : 2580-4316
E-ISSN : 2654-8054

4. HASIL DAN PEMBAHASAN

Pengujian Model

Dataset yang dimiliki berjumlah 7.680 citra. Pada tahap pertama, dataset ini dipecah untuk menyisihkan data pengujian (testing data) sebanyak 1.536 gambar (20%) yang akan digunakan murni untuk evaluasi performa model.

Dengan 1.536 data uji yang dipisahkan dari data latih. Model berhasil memprediksi 1.502 gambar secara tepat dan 34 gambar keliru, menghasilkan tingkat akurasi akhir 97,79% dengan rata-rata confidence 98,31%.



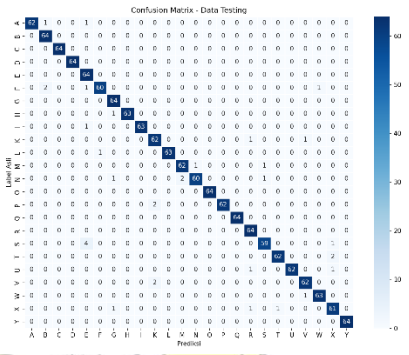
Gambar 2. Diagram batang hasil pengujian

Berdasarkan evaluasi metrik dan Confusion Matrix, Lima kelas huruf (C, D, O, Q, dan Y) mencapai skor absolut (1,000) pada seluruh metrik Precision, Recall, dan F1-Score. Kelas A, H, I, L, P, dan U juga mencatatkan precision sempurna (1,000). Kesalahan prediksi mayoritas disebabkan oleh kemiripan gestur jari. Kasus tertinggi adalah 4 gambar huruf S yang keliru diprediksi sebagai huruf E. Hal ini berimbas pada huruf S yang mencatat *recall* terendah (0,922) dan huruf E dengan *precision* terendah (0,901). Terdapat pola kesalahan saling tertukar dalam skala kecil (1-2 gambar) pada pasangan abjad yang secara visual identik, yakni antara huruf M dengan N, serta huruf T dengan X. Sisa kesalahan klasifikasi hanya bersifat sporadis (maksimal 1-2 gambar) dan tersebar merata, menunjukkan bahwa model tidak memiliki kelemahan sistematis terhadap fitur abjad tertentu.

Tabel 1. Hasil Pengujian Model

Hasil Testing Model SIBI	
Akurasi	97,79%
Confidence	98.31%

Total Foto Uji	1536			
Kelas	*Precision*	*Recall*	*F1-Score*	*Support*
A	1.000	0.969	0.984	64
B	0.955	1.000	0.977	64
C	1.000	1.000	1.000	64
D	1.000	1.000	1.000	64
E	0.901	1.000	0.948	64
F	0.984	0.938	0.960	64
G	0.955	1.000	0.977	64
H	1.000	0.984	0.992	64
I	1.000	0.984	0.992	64
K	0.939	0.969	0.954	64
L	1.000	0.984	0.992	64
M	0.969	0.969	0.969	64
N	0.984	0.938	0.960	64
O	1.000	1.000	1.000	64
P	1.000	0.969	0.984	64
Q	1.000	1.000	1.000	64
R	0.955	1.000	0.977	64
S	0.967	0.922	0.944	64
T	0.984	0.969	0.976	64
U	1.000	0.969	0.984	64
V	0.969	0.969	0.969	64
W	0.984	0.984	0.984	64
X	0.938	0.953	0.946	64
Y	1.000	1.000	1.000	64



Gambar 3. Confusion Matrix

Pengujian Kondisi Nyata

Pengujian real-time dilakukan langsung pada smartphone dan tablet untuk mengukur ketahanan model TensorFlow Lite di lingkungan fisik. Pengujian dibagi menjadi kondisi ideal yang mencakup keseluruhan 24 huruf, serta variasi lingkungan ekstrem yang difokuskan pada 5 huruf dengan tingkat kesalahan tertinggi (E, N, O, S, dan X).

Tabel 2. Pengujian Kondisi Ideal

Label Asli	Prediksi S*mart phone*	Prediksi Tablet	Keterangan
A	A 99%	A 99%	Model dapat mengenali tangan dengan sangat baik pada kedua perangkat.

P-ISSN : 2580-4316
E-ISSN : 2654-8054
https://doi.org/10.37817/ikraith-informatika.v10i3

Label Asli	Prediksi S*mart phone*	Prediksi Tablet	Keterangan
B	B 99%	B 99%	Model dapat mengenali tangan dengan sangat baik pada kedua perangkat.
C	C 99%	C 99%	Model dapat mengenali tangan dengan sangat baik pada kedua perangkat.
D	D 99%	D 99%	Model dapat mengenali tangan dengan sangat baik pada kedua perangkat.
E	E 99%	E 99%	Model dapat mengenali tangan dengan sangat baik pada kedua perangkat.
F	F 78% - 98%	F 71% - 98%	Terjadi fluktuasi *confidence* sebesar 20% pada *smarthphone* dan 27% pada tablet, namun label yang ditampilkan tetap konsisten sebagai huruf F.
G	G 99%	G 99%	Model dapat mengenali tangan dengan sangat baik pada kedua perangkat.
H	H 98% - 99%	H 99%	Model dapat mengenali tangan dengan sangat baik pada kedua perangkat.
I	I 99%	I 99%	Model dapat mengenali tangan dengan sangat baik pada kedua perangkat.
K	K 99%	K 99%	Model dapat mengenali tangan dengan sangat baik pada kedua perangkat.
L	L 99%	L 99%	Model dapat mengenali tangan dengan sangat baik pada kedua perangkat.
M	M 99%	M 99%	Model dapat mengenali tangan dengan sangat baik pada kedua perangkat.
N	N 41% - 92%	N 85% - 98%	Fluktuasi *confidence* sebesar 51% terjadi pada *smartphone*, sementara tablet relatif lebih stabil di 13%. Label yang ditampilkan tetap benar sebagai huruf N pada kedua perangkat.
O	O 99%	O 99%	Model dapat mengenali tangan dengan sangat baik pada kedua perangkat.
P	P 99%	P 90% - 99%	Model dapat mengenali tangan dengan sangat baik pada kedua perangkat, dengan sedikit fluktuasi pada tablet sebesar 9%.
Q	Q 99%	Q 99%	Model dapat mengenali tangan dengan sangat baik pada kedua perangkat.
R	R 99%	R 99%	Model dapat mengenali tangan dengan sangat baik pada kedua perangkat.
S	S 68% - 95%	S 99%	Fluktuasi *confidence* sebesar 27% terjadi pada *smartphone*, sementara tablet stabil pada 99%.
T	T 99%	T 99%	Model dapat mengenali tangan dengan sangat baik pada kedua perangkat.
U	U 72% - 94%	U 97% - 99%	Fluktuasi *confidence* sebesar 22% terjadi pada *smartphone*, sementara tablet relatif stabil.
V	V 76% - 97%	V 99%	Fluktuasi *confidence* terjadi pada *smartphone*, sementara tablet stabil pada 99%.
W	W 96% - 99%	W 87% - 98%	Fluktuasi *confidence* sebesar 21% pada *smartphone* dan 11% pada tablet, namun label tetap konsisten.
X	X 99%	X 96% - 98%	Model dapat mengenali tangan dengan sangat baik pada kedua perangkat.
Y	Y 99%	Y 99%	Model dapat mengenali tangan dengan sangat baik pada kedua perangkat.
Nilai *Confidence* Rata-rata	95,19%	97,50%	

Pada jarak ±50 cm dengan pencahayaan dan latar belakang optimal, model berhasil memprediksi ke-24 abjad SIBI dengan tingkat keberhasilan 100%. Rata-rata *confidence* pada tablet mencapai 97,50%, sedikit lebih tinggi dan stabil dibandingkan *smartphone* 95,19%. Meskipun sempat terjadi fluktuasi *confidence* yang cukup lebar pada gestur huruf N, S, U, dan

https://doi.org/10.37817/ikraith-informatika.v10i3
P-ISSN : 2580-4316
E-ISSN : 2654-8054

V, label akhir yang ditebak oleh sistem tetap konsisten dan akurat pada kedua perangkat.

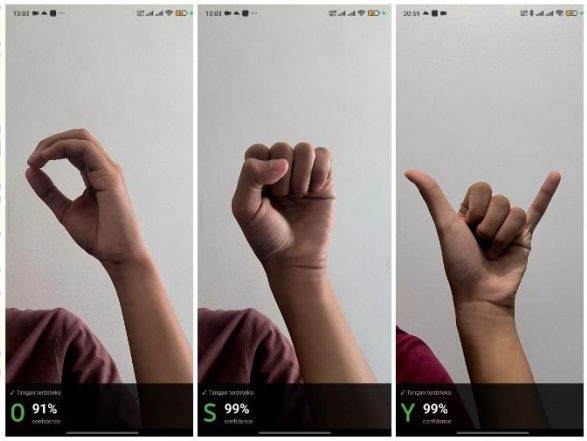
Tabel 3. Pengujian Variasi Lingkungan

Skenario Pengujian	Variabel Uji	Hasil & Analisis Prediksi (Smartphone & Tablet)
Variasi Jarak	Dekat (±30 cm)	Akurasi menurun karena pembingkaian (*cropping*) kurang optimal. Huruf S kerap keliru terbaca sebagai E (45-86%) pada *smartphone*.
	Jauh (±100 cm)	Prediksi kembali akurat, namun resolusi kamera tablet yang lebih rendah memicu fluktuasi *confidence* pada gestur rapat seperti S dan X.
Intensitas Cahaya	*Low Light* (Redup)	Toleransi Sangat Baik. Selama kontur tangan masih dapat diekstraksi, model berhasil menebak seluruh label dengan akurat.
	Backlight (Silau)	Kelemahan Utama. Efek siluet menghilangkan informasi kedalaman dan lipatan jari. Huruf S konsisten salah terbaca sebagai E, dan X gagal dikenali secara stabil oleh tablet.
Gangguan Visual	Latar Ramai (*Noise*)	Ekstraksi fitur terganggu oleh objek di belakang. Akurasi huruf X mengalami degradasi parah (dominan salah diprediksi sebagai Q pada *smartphone*).
	Objek Terpotong	Gagal Total. Huruf X tidak dikenali sama sekali jika ujung jari terpotong batas layar. Huruf Y salah terbaca sebagai L pada tablet akibat hilangnya rentangan ibu jari.
	Multi-objek (Dua Tangan)	Berhasil 100%. Algoritma sukses memprioritaskan tangan di latar depan (area terbesar) dan secara cerdas mengabaikan gestur pengecoh di latar belakang.

Skenario ini secara spesifik menguji huruf E, N, O, S, dan X (serta gestur lebar seperti Y pada skenario terpotong) untuk menilai batas kinerja integrasi sensor kamera, MediaPipe, dan CNN saat dihadapkan pada kondisi fisik yang tidak standar. Secara keseluruhan, model *computer vision* ini sangat andal menangani distraksi multi-objek dan minim cahaya, namun terbukti rentan terhadap hilangnya detail anatomi spasial yang diakibatkan oleh bayangan *backlight* maupun batas tepi kamera (*out-of-bound*).

Implementasi Antarmuka Pengguna

Antarmuka aplikasi dibangun pada platform Android terintegrasi sensor kamera untuk memfasilitasi penerjemahan SIBI secara *real-time*. Selama proses terjemahan berlangsung, pengguna mendapatkan umpan balik visual *real-time* di layar. Karakter alfabet yang diprediksi oleh model muncul secara langsung, disertai dengan persentase tingkat kepercayaannya. Untuk meningkatkan kejelasan, warna teks akan berubah secara otomatis. Hijau menandakan tingkat kepercayaan yang tinggi (≥ 70%), oranye menunjukkan tingkat yang sedang, dan merah mencerminkan tingkat kepercayaan yang rendah. Selain itu, sistem terus memberikan informasi pelacakan—menampilkan "✓ Tangan terdeteksi" untuk mengonfirmasi posisi yang akurat atau memberikan peringatan jika MediaPipe gagal mengidentifikasi tangan.



Gambar 4. Implementasi Antarmuka Pengguna

5. KESIMPULAN

Penelitian ini berhasil mengintegrasikan model *Convolutional Neural Network* (CNN) berarsitektur MobileNetV2 dan kerangka kerja MediaPipe ke dalam aplikasi penerjemah SIBI berbasis Android yang responsif dan stabil tanpa indikasi kebocoran memori. Secara teoretis, evaluasi model menunjukkan performa yang sangat baik dengan tingkat akurasi mencapai

97,79%, meskipun masih terdapat margin kesalahan minor pada pengklasifikasian gestur yang memiliki kemiripan visual tinggi, seperti pasangan huruf M-N, T-X, dan S-E.

Pada pengujian kondisi nyata (*on-device*), sistem mampu memprediksi ke-24 abjad statis secara sempurna pada lingkungan ideal dan terbukti sangat cerdas dalam memprioritaskan objek utama saat dihadapkan pada distraksi dua tangan sekaligus. Namun, kinerja ekstraksi fitur model memiliki kelemahan yang sangat bergantung pada kualitas tangkapan sensor kamera perangkat. Sistem rentan mengalami degradasi akurasi pada kondisi *backlight* yang mengubah tekstur jari menjadi siluet gelap, latar belakang visual yang bising (*noise*), serta gagal berfungsi apabila sebagian anatomi tangan terpotong oleh batas layar. Selain itu, perbedaan spesifikasi perangkat keras seperti resolusi lensa kamera juga secara langsung memengaruhi stabilitas nilai prediksi, terutama pada pengujian dengan jarak pandang yang jauh.

DAFTAR PUSTAKA

Alfiansyah, N. S., & Litanianda, Y. (2024). Identifikasi Lumpy Skin Disease Menggunakan Tensorflow Dengan Metode Convolutional Neuron Network. *JATI (Jurnal Mahasiswa Teknik Informatika)*, *8*(4), 7330–7336.

Azizah, E. N., Resmi, M. G., & Alam, S. (2023). Penerapan Metode Design Thinking Pada Perancangan User Interface Aplikasi Mobile Pengenalan Bahasa Isyarat Indonesia (Bisindo). *Jurnal Mnemonic*, *6*(1), 71–76.

Fachrisyam, M., Indra, D., & Hasnawi, M. (2025). Implementasi metode YOLO dalam mendeteksi jenis sampah berbasis computer vision. *LINIER: Literatur Informatika Dan Komputer*, *2*(1), 68–76.

Fatmawati, R., Asmara, R., Prayogi, Y. R., & Hakkun, R. Y. (2022). Aplikasi Pembelajaran Sistem Isyarat Bahasa Indonesia (SIBI) Berbasis Voice Menggunakan OpenSIBI. *Technomedia J*, *7*(1), 22–39.

Gulzar, Y. (2023). Fruit image classification model based on MobileNetV2 with deep transfer learning technique. *Sustainability*, *15*(3), 1906.

Halder, A., & Tayade, A. (2021). Real-time vernacular sign language recognition using mediapipe and machine learning. *Journal Homepage: Www. Ijrpr. Com ISSN*, *2582*(7421), 2.

Ramadhani, F., Satria, A., & Salamah, S. (2023). Implementasi Algoritma Convolutional Neural Network dalam Mengidentifikasi Dini Penyakit pada Mata Katarak. *Sudo Jurnal Teknik Informatika*, *2*(4), 167–175.

Sharifani, K., & Amini, M. (2023). Machine learning and deep learning: A review of methods and applications. *World Information Technology and Engineering Journal*, *10*(07), 3897–3904.