

# **IMPLEMENTASI SISTEM PENERJEMAH BISINDO DUA ARAH BERBASIS MEDIAPIPE DAN ARTIFICIAL INTELLEGENCE**

**TUGAS AKHIR**

**Diajukan Oleh:**

**ADEA MARLINDA**

**NIM. 220705021**

**Mahasiswa Fakultas Sains dan Teknologi**

**Program Studi Teknologi Informasi**



**FAKULTAS SAINS DAN TEKNOLOGI  
UNIVERSITAS ISLAM NEGERI AR-RANIRY  
BANDA ACEH  
2026 M / 1447 H**

## LEMBAR PERSETUJUAN

### IMPLEMENTASI SISTEM PENERJEMAH BISINDO DUA ARAH BERBASIS MEDIAPIPE DAN ARTIFICIAL INTELLIGENCE

#### TUGAS AKHIR

Diajukan Kepada Fakultas Sains dan Teknologi  
Universitas Islam Negeri (UIN) Ar-Raniry Banda Aceh  
Sebagai Salah Satu Persyaratan Penulisan Tugas Akhir  
Dalam Prodi Teknologi Informasi

Oleh:  
**ADEA MARLINDA**  
220705021

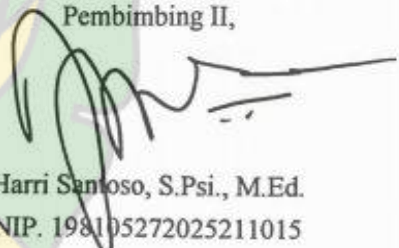
Mahasiswa Fakultas Sains dan Teknologi  
Program Studi Teknologi Informasi

Disetujui Untuk di Munaqasyahkan Oleh:

Pembimbing I,


  
Khairan AR, M.Kom  
NIP. 198607042014031001

Pembimbing II,

  
Harri Santoso, S.Psi., M.Ed.  
NIP. 198105272025211015

Mengetahui,

Ketua Program Studi Teknologi Informasi

  
Malahayati, M.T

NIP. 198301272015032003

## LEMBAR PENGESAHAN

### IMPLEMENTASI SISTEM PENERJEMAH BISINDO DUA ARAH BERBASIS MEDIAPIPE DAN ARTIFICIAL INTELLIGENCE

#### TUGAS AKHIR

Telah Diuji Oleh Panitia Ujian Munaqasyah Tugas Akhir  
Fakultas Sains dan Teknologi UIN Ar-Raniry Banda Aceh dan Dinyatakan Lulus  
Serta Diterima Sebagai Salah Satu Beban Studi Program Sarjana (S1)  
Dalam Prodi Teknologi Informasi


Pada Hari/Tanggal: Selasa, 24 Februari 2026  
06 Ramadan 1447 H  
Di Darussalam, Banda Aceh

Panitia Ujian Munaqasyah Tugas Akhir:


Ketua,

  
Khairan AR, M.Kom  
NIP. 198607042014031001


Sekretaris,

  
Harri Santoso, S.Psi., M.Ed.  
NIP. 198105272025211015

Penguji I,

  
Dr. Hendri Ahmadian, S.Si, M.I.M  
NIP. 198301042014031002

Penguji II,

  
Ridha Ilahi, M.T  
NIP. 197905302014031001

Mengetahui,

Dekan Fakultas Sains dan Teknologi  
UIN Ar-Raniry Banda Aceh



  
Prof. Dr. Ir. Muhammad Dirhamsyah, M.T., IPU  
NIP. 196210021988111001

## SURAT PERNYATAAN KEASLIAN TUGAS AKHIR

Yang bertanda tangan di bawah ini:

Nama : Adea Marlinda

NIM : 220705021

Program Studi : Teknologi Informasi

Fakultas : Sains dan Teknologi

Judul : Implementasi Sistem Penerjemah BISINDO Dua Arah Berbasis  
MediaPipe Dan Artificial Intelligence

Dengan ini menyatakan bahwa dalam penulisan tugas akhir ini, saya:

1. Tidak menggunakan ide orang lain tanpa mampu mengembangkan dan mempertanggungjawabkan;
2. Tidak melakukan plagiasi terhadap naskah karya orang lain;
3. Tidak menggunakan karya orang lain tanpa menyebutkan sumber asli atau tanpa izin pemilik karya;
4. Tidak memanipulasi dan memalsukan data;
5. Mengerjakan sendiri karya ini dan mampu bertanggungjawab atas karya ini.

Bila dikemudian hari ada tuntutan dari pihak lain atas karya saya, dan telah melalui pembuktian yang dapat dipertanggungjawabkan dan ternyata memang ditemukan bukti bahwa saya telah melanggar pernyataan ini, maka saya siap dikenai sanksi berdasarkan aturan yang berlaku di Fakultas Sains dan Teknologi UIN Ar-Raniry Banda Aceh.

Demikian pernyataan ini saya buat dengan sesungguhnya dan tanpa paksaan dari pihak manapun.

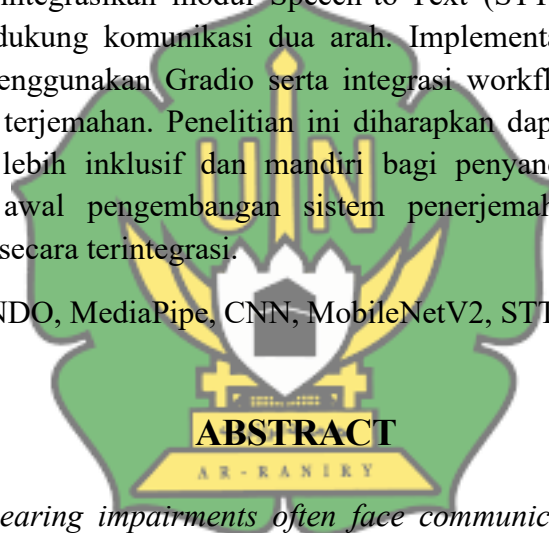
Banda Aceh, 13 Februari 2026

Menyatakan,  
  
Adea marlinda

## ABSTRAK

Komunikasi merupakan kebutuhan dasar manusia, namun penyandang tunarungu sering mengalami hambatan dalam komunikasi verbal sehingga memerlukan media alternatif. Di Indonesia, Bahasa Isyarat Indonesia (BISINDO) digunakan secara luas karena bersifat alami dan fleksibel. Keterbatasan jumlah juru bahasa isyarat serta rendahnya pemahaman masyarakat terhadap BISINDO menyebabkan terjadinya kesenjangan komunikasi. Perkembangan kecerdasan buatan memberikan peluang untuk mengembangkan sistem penerjemah bahasa isyarat yang lebih efektif dan terintegrasi. Penelitian ini bertujuan merancang dan mengembangkan sistem penerjemah BISINDO dua arah berbasis MediaPipe dan Artificial Intelligence. Sistem difokuskan pada pengenalan huruf alfabet BISINDO (A-Z) yang dirangkai menjadi kata sederhana. Proses klasifikasi gestur dilakukan menggunakan Convolutional Neural Network (CNN) dengan pendekatan transfer learning pada arsitektur MobileNetV2 serta fine-tuning untuk menyesuaikan karakteristik dataset. Sistem juga mengintegrasikan modul Speech-to-Text (STT) dan Text-to-Speech (TTS) untuk mendukung komunikasi dua arah. Implementasi dilakukan melalui antarmuka web menggunakan Gradio serta integrasi workflow automation untuk dokumentasi hasil terjemahan. Penelitian ini diharapkan dapat memberikan solusi komunikasi yang lebih inklusif dan mandiri bagi penyandang tunarungu serta menjadi langkah awal pengembangan sistem penerjemah BISINDO berbasis kecerdasan buatan secara terintegrasi.

**Kata kunci:** BISINDO, MediaPipe, CNN, MobileNetV2, STT, TTS.



*Individuals with hearing impairments often face communication barriers due to limited public understanding of Indonesian Sign Language (BISINDO) and the shortage of interpreters. Advances in artificial intelligence offer opportunities to develop more effective and integrated sign language translation systems. This study aims to develop a bidirectional BISINDO translation system based on MediaPipe and Artificial Intelligence. The system focuses on recognizing BISINDO alphabet gestures (A-Z) using a Convolutional Neural Network (CNN) with a transfer learning approach based on MobileNetV2 and fine-tuning. It also integrates Speech-to-Text (STT) and Text-to-Speech (TTS) to support two-way communication. The system is implemented through a web-based interface and integrated with workflow automation for documentation. This research is expected to contribute to more inclusive and independent communication for individuals with hearing impairments through an AI-based translation system.*

**Keywords:** BISINDO, MediaPipe, CNN, MobileNetV2, STT, TTS.

## KATA PENGANTAR

Puji syukur ke hadirat Allah SWT atas segala rahmat, hidayah, dan karunia-Nya sehingga penulis dapat menyelesaikan tugas akhir yang berjudul “Implementasi Sistem Penerjemah BISINDO Dua Arah Berbasis MediaPipe dan Artificial Intelligence.” Tak lupa pula Shalawat beriringkan salam tercurahkan kepada sang baginda Nabi Muhammad SAW. Tugas akhir ini disusun sebagai salah satu persyaratan untuk menyelesaikan pendidikan pada Program Studi Teknologi Informasi, Fakultas Sains dan Teknologi, Universitas Islam Negeri (UIN) Ar-Raniry Banda Aceh.

Penulis menyadari bahwa penyusunan tugas akhir ini masih memiliki banyak kekurangan. Oleh karena itu, kritik dan saran yang bersifat membangun sangat diharapkan untuk penyempurnaan di masa mendatang. Penyelesaian tugas akhir ini tidak terlepas dari dukungan dan bantuan berbagai pihak. Dengan penuh rasa terima kasih, penulis menyampaikan apresiasi kepada semua pihak yang telah membantu, baik secara langsung maupun tidak langsung, sejak proses penelitian hingga penyusunan laporan ini.

Oleh karena itu, penulis ingin mengucapkan terima kasih dengan teramat dalam kepada:

1. Penulis menyampaikan rasa terima kasih yang sebesar-besarnya kepada Ayah, Ibu, Kak Helsa, Sisil, dan Arsyah atas segala doa, kasih sayang, dukungan, serta motivasi yang tiada henti diberikan sejak awal perkuliahan hingga penulis dapat menyelesaikan proses penyusunan tugas akhir ini.
2. Ibu Malahayati, M.T, selaku Ketua Program Studi Teknologi Informasi, yang telah banyak memberikan bimbingan serta pembelajaran kepada penulis.
3. Bapak k Khairan AR, M.Kom dan Bapak Harri Santoso, S.Psi., M.Ed, selaku dosen pembimbing 1 dan dosen pembimbing 2, di mana sudah memberikan arahan dan bimbingan selama proses pembuatan dan penyusunan tugas akhir penulis.
4. Ibu Cut Ida Rahmadiana, S.Si, selaku seseorang yang telah banyak membantu selama masa perkuliahan penulis.
5. Ucapan terima kasih teramat dalam dan hangat kepada seseorang yang begitu banyak membantu penulis, terima kasih atas segala rasa, kasih sayang, upaya, ucapan, serta doa baik yang tak pernah lupa di panjatkan kepada penulis, penulis berharap jalinan hubungan ini bisa terus belayar sampai kapan pun itu.
6. Terima kasih kepada teman-teman yang selalu memberikan dukungan, tawa dan semangat kepada penulis, ucapan ini tertuju kepada Puja, Naqqa, Nurul, Ojak, dan teman-teman Angkatan 22.

Banda Aceh, 13 Februari 2026

Yang menyatakan

Adea Marlinda

# DAFTAR ISI

<b>LEMBAR PERSETUJUAN</b> .....	<b>ii</b>
<b>LEMBAR PENGESAHAN</b> .....	<b>iii</b>
<b>SURAT PERNYATAN KEASLIAN TUGAS AKHIR</b> .....	Error! Bookmark not defined.
<b>ABSTRAK</b> .....	<b>v</b>
<b>KATA PENGANTAR</b> .....	<b>vi</b>
<b>DAFTAR ISI</b> .....	<b>vii</b>
<b>DAFTAR GAMBAR</b> .....	<b>x</b>
<b>DAFTAR TABEL</b> .....	<b>xi</b>
<b>BAB I PENDAHULUAN</b> .....	<b>1</b>
I.1 Latar Belakang .....	1
I.2 Rumusan Masalah .....	2
I.3 Tujuan Penelitian .....	2
I.4 Manfaat Penelitian .....	3
I.5 Batasan Masalah .....	4
I.6 Metodologi Penelitian .....	4
I.7 Sistematika Penulisan .....	5
<b>BAB II TINJAUAN PUSTAKA</b> .....	<b>7</b>
II.1 Penelitian Terdahulu .....	7
II.2 Kajian Teoritis .....	10
II.2.1 Bahasa Isyarat Indonesia (BISINDO) .....	10
II.2.2 Computer Vision dalam Pengenalan Gesture .....	11
II.2.3 MediaPipe untuk Ekstraksi Landmark Tangan .....	11
II.2.4 Convolutional Neural Network (CNN) .....	12
II.2.5 Arsitektur MobileNetV2 dan Transfer Learning .....	13
II.2.6 Text-to-Speech (TTS).....	14
II.2.7 Speech-to-Text (STT).....	15
II.2.8 Antarmuka Interaktif Berbasis Gradio .....	15
II.2.9 Google Colab .....	16
II.2.10 Workflow Automation dengan n8n .....	16
II.2.11 Evaluasi Sistem .....	17
<b>BAB III METODOLOGI PENELITIAN</b> .....	<b>20</b>

III.1	Jenis Penelitian .....	20
III.2	Tahapan Penelitian .....	21
III.3	Sistem Penerjemahan Bahasa Isyarat Bisindo .....	23
III.3.1	Arsitektur Sistem .....	24
III.3.2	Perancangan 2 Mode (Gesture dan Suara).....	25
III.3.3	Perancangan Alur Sistem (System Flow).....	26
III.3.4	Pertimbangan Desain Sistem.....	27
III.4	Pengumpulan Data .....	28
III.4.1	Jenis Data.....	28
III.4.2	Proses Akuisisi Data.....	28
III.4.3	Struktur Dataset .....	28
III.4.4	Pembagian Dataset .....	29
III.4.5	Pertimbangan Penggunaan Dataset Mandiri.....	29
III.5	Preprocessing Data Untuk Pelatihan .....	29
III.5.1	Penyesuaian Ukuran Citra.....	30
III.5.2	Normalisasi Nilai Pixel.....	30
III.5.3	Augmentasi Data .....	30
III.5.4	Encoding Label.....	31
III.5.5	Pembagian Data Training dan Validation.....	31
III.6	Metode Pelatihan Model.....	31
III.6.1	Konfigurasi Arsitektur Model.....	32
III.6.2	Tahap 1 Transfer Learning (Initial Training) .....	32
III.6.3	Tahap 2 Fine-Tuning Bertahap.....	33
III.6.4	Pertimbangan Metodologis .....	34
III.7	Metode Evaluasi .....	34
III.7.1	Evaluasi Model Klasifikasi .....	34
III.7.2	Evaluasi Sistem Secara Fungsional.....	36
<b>BAB IV</b>	<b>IMPLEMENTASI DAN HASIL PENELITIAN .....</b>	<b>37</b>
IV.1	Implementasi Dataset dan Preprocessing.....	37
IV.2	Implementasi Pelatihan Model .....	40
IV.2.1	Implementasi Transfer Learning.....	40
IV.2.2	Implementasi Fine-Tuning.....	43

IV.3	Implementasi Sistem Penerjemahan Dua Arah.....	45
IV.3.1	Implementasi Mode Gestur .....	46
IV.3.2	Implementasi Mode Suara .....	48
IV.3.3	Integrasi Webhook n8n .....	49
IV.4	Evaluasi dan Pengujian Sistem .....	51
IV.4.1	Evaluasi Model Klasifikasi .....	51
IV.4.2	Pengujian Fungsional Sistem.....	58
IV.5	Analisis Hasil.....	61
IV.5.1	Analisis Performa Model Klasifikasi.....	62
IV.5.2	Analisis Implementasi Sistem.....	62
IV.5.3	Kesesuaian dengan Tujuan Penelitian .....	63
IV.5.4	Keterbatasan Sistem.....	64
<b>BAB V</b>	<b>KESIMPULAN DAN SARAN .....</b>	<b>65</b>
V.1	Kesimpulan.....	65
V.2	Saran.....	65
<b>DAFTAR PUSTAKA .....</b>		<b>66</b>



## DAFTAR GAMBAR

Gambar II. 1	Huruf Alfabet BISINDO .....	10
Gambar II. 2	Arsitektur Umum CNN .....	13
Gambar II. 3	Landmark Hand.....	12
Gambar II. 4	Orkestrasi Workflow n8n .....	16
Gambar III. 1	Tahapan Penelitian .....	21
Gambar III. 2	Perancangan Sistem Mode Gesture dan Suara.....	26
Gambar III. 3	Struktur Dataset.....	29
Gambar IV. 1	Folder Dataset di Google Drive .....	37
Gambar IV. 2	Banyak Gambar Per Label Huruf BISINDO.....	39
Gambar IV. 3	Output Pembagian Data Training dan Validation .....	39
Gambar IV. 4	Hasil Beberapa Gambar Augmentasi Data.....	40
Gambar IV. 5	Ringkasan Arsitektur MobileNetV2.....	41
Gambar IV. 6	Proses Pelatihan Transfer Learning.....	43
Gambar IV. 7	Proses Pelatihan Fine-Tuning.....	45
Gambar IV. 8	Pilihan Mode BISINDO .....	46
Gambar IV. 9	Antarmuka Upload Citra Gestur .....	47
Gambar IV. 10	Hasil Prediksi Huruf BISINDO .....	47
Gambar IV. 11	Antarmuka Rekam/Upload Suara .....	48
Gambar IV. 12	Hasil Konversi Suara ke Teks.....	49
Gambar IV. 13	Visualisasi Huruf BISINDO .....	49
Gambar IV. 14	Workflow Integrasi Webhook n8n.....	50
Gambar IV. 15	Hasil Penyimpanan ke Google Spreadsheets Gesture.....	50
Gambar IV. 16	Hasil Penyimpanan ke Google Spreadsheets Suara .....	51
Gambar IV. 17	Grafik Akurasi Model.....	52
Gambar IV. 18	Grafik Loss Model.....	53
Gambar IV. 19	Classification Report Accuracy .....	53
Gambar IV. 20	Confusion Matrix Klasifikasi Huruf BISINDO.....	56
Gambar IV. 21	Hasil Uji Sistem Dapat menerima File Gambar .....	58
Gambar IV. 22	Hasil Uji Confidence Perhuruf .....	59
Gambar IV. 23	Hasil Uji Konversi Suara ke Text.....	59
Gambar IV. 24	Hasil Uji STT dan visualisasi huruf.....	60
Gambar IV. 25	Hasil Koneksi n8n.....	60
Gambar IV. 26	Hasil Uji Terjemahan Berhasil Terkirim Mode Gesture .....	61
Gambar IV. 27	Hasil Uji Terjemahan Berhasil Terkirim Mode Suara .....	61

## DAFTAR TABEL

Tabel II. 1 Ringkasan Penelitian Terdahulu .....	8
Tabel II. 2 Perbandingan Penelitian Terdahulu .....	8
Tabel II. 3 Ringkasan Kajian Teoritis .....	18
Tabel IV.1 Clasification Report.....	54



# BAB I

## PENDAHULUAN

### I.1 Latar Belakang

Komunikasi merupakan kebutuhan dasar manusia dalam menyampaikan ide, informasi, dan perasaan. Namun, penyandang tunarungu mengalami hambatan dalam komunikasi verbal sehingga membutuhkan media alternatif. Di Indonesia, Bahasa Isyarat Indonesia (BISINDO) menjadi bahasa yang paling banyak digunakan oleh komunitas Tuli karena bersifat alami dan fleksibel dibandingkan Sistem Isyarat Bahasa Indonesia (SIBI) (Abdillah & Andamisari, 2023).

Secara linguistik, BISINDO merupakan bahasa alami (natural language) yang berkembang dalam komunitas Tuli dengan struktur visual-spasial yang mandiri serta memanfaatkan ekspresi wajah dan gerakan tubuh sebagai bagian dari makna. Sebaliknya, SIBI adalah sistem isyarat resmi yang mengikuti tata bahasa Bahasa Indonesia secara ketat dan lebih banyak digunakan dalam konteks pendidikan formal. Karena lebih ekspresif dan sesuai dengan kebutuhan komunikasi sehari-hari, BISINDO lebih diterima dalam praktik sosial komunitas Tuli.

Meskipun demikian, sebagian besar masyarakat umum belum memahami bahasa isyarat. Data Kementerian Kesehatan RI tahun 2021 mencatat sekitar 2,3 juta individu dengan gangguan pendengaran di Indonesia, sementara jumlah juru bahasa isyarat masih terbatas. Kondisi ini menimbulkan ketergantungan pada pendamping serta potensi miskomunikasi dalam pelayanan publik dan interaksi sosial (Isnaniah et al., 2023).

Perkembangan Artificial Intelligence memberikan peluang untuk menjembatani kesenjangan tersebut melalui sistem pengenalan gestur berbasis computer vision. MediaPipe digunakan untuk mendeteksi landmark tangan pada citra, sedangkan Convolutional Neural Network (CNN) efektif dalam mengklasifikasikan pola visual dengan akurasi tinggi dibandingkan metode konvensional seperti K-Nearest Neighbor (KNN) atau Support Vector Machine (SVM) (Eriyadi et al., 2024; Saleh, 2025). Untuk meningkatkan performa dan efisiensi pelatihan, pendekatan transfer learning menggunakan MobileNetV2 dimanfaatkan karena ringan dan memiliki stabilitas yang baik dalam tugas klasifikasi citra.

Namun, penelitian terdahulu lebih banyak berfokus pada American Sign Language (ASL) serta umumnya bersifat satu arah. Hal ini menunjukkan adanya research gap dalam

pengembangan sistem penerjemah berbasis BISINDO yang mendukung komunikasi dua arah secara terintegrasi (Sanjaya & Ilone, 2023).

Berdasarkan permasalahan tersebut, penelitian ini mengembangkan sistem penerjemah BISINDO dua arah berbasis MediaPipe dan Artificial Intelligence yang mengintegrasikan klasifikasi huruf menggunakan CNN dengan pendekatan transfer learning, serta fitur Speech-to-Text dan Text-to-Speech. Implementasi dilakukan melalui antarmuka web berbasis Gradio dengan dukungan workflow automation n8n untuk dokumentasi hasil terjemahan (Kelana et al., 2024). Tahap awal difokuskan pada pengenalan huruf alfabet BISINDO (A-Z) yang dirangkai menjadi kata sederhana guna menjaga stabilitas dan akurasi model.

Penelitian ini diharapkan dapat berkontribusi dalam pengembangan teknologi inklusif yang lebih relevan dengan kebutuhan komunitas Tuli di Indonesia serta mendukung komunikasi yang lebih mandiri dan efektif.

## **I.2 Rumusan Masalah**

Berdasarkan latar belakang yang telah diuraikan, penelitian ini dirumuskan dalam dua pertanyaan utama sebagai berikut:

1. Bagaimana merancang dan mengimplementasikan sistem klasifikasi gestur huruf Bahasa Isyarat Indonesia (BISINDO) berbasis MediaPipe dan Convolutional Neural Network (CNN) dengan pendekatan transfer learning untuk menghasilkan teks yang dapat disusun menjadi kata sederhana?
2. Bagaimana mengintegrasikan sistem klasifikasi tersebut dengan fitur komunikasi dua arah berbasis web (Gradio) melalui Speech-to-Text (STT), Text-to-Speech (TTS), serta workflow automation sehingga terbentuk sistem penerjemah BISINDO yang interaktif dan terintegrasi?

## **I.3 Tujuan Penelitian**

Penelitian ini bertujuan untuk merancang dan mengembangkan sistem penerjemah Bahasa Isyarat Indonesia (BISINDO) dua arah berbasis Artificial Intelligence guna mendukung komunikasi yang lebih inklusif antara penyandang tunarungu dan masyarakat umum. Secara khusus, tujuan penelitian ini adalah:

1. Mengembangkan sistem klasifikasi gestur huruf BISINDO berbasis MediaPipe dan Convolutional Neural Network (CNN) dengan pendekatan transfer learning untuk

menghasilkan teks yang dapat disusun menjadi kata sederhana secara akurat dan stabil.

2. Mengintegrasikan sistem klasifikasi tersebut dengan fitur komunikasi dua arah berbasis web (Gradio) melalui Speech-to-Text (STT), Text-to-Speech (TTS), serta workflow automation sehingga terbentuk sistem penerjemah BISINDO yang interaktif dan terintegrasi.

#### **I.4 Manfaat Penelitian**

Penelitian ini diharapkan dapat memberikan kontribusi dalam pengembangan teknologi asistif berbasis Artificial Intelligence, khususnya dalam mendukung komunikasi bagi penyandang disabilitas pendengaran. Adapun manfaat penelitian ini adalah sebagai berikut:

1. Manfaat Teoritis

Penelitian ini menambah kajian ilmiah mengenai penerapan MediaPipe dan Convolutional Neural Network (CNN) dengan pendekatan transfer learning dalam sistem penerjemah Bahasa Isyarat Indonesia (BISINDO). Hasil penelitian ini diharapkan dapat memperkaya referensi mengenai pengembangan sistem komunikasi dua arah berbasis AI serta menjadi dasar bagi penelitian lanjutan di bidang pengolahan citra dan pengenalan suara.

2. Manfaat Praktis

Sistem yang dikembangkan dapat membantu penyandang tunarungu dalam menyampaikan dan menerima informasi melalui penerjemahan gestur huruf menjadi teks, serta konversi teks dan suara melalui fitur Speech-to-Text (STT) dan Text-to-Speech (TTS). Implementasi berbasis web memungkinkan sistem digunakan secara interaktif, sementara integrasi workflow automation mendukung proses dokumentasi dan distribusi hasil terjemahan secara lebih efisien.

3. Manfaat bagi Pengembang dan Institusi

Penelitian ini dapat menjadi referensi dalam pengembangan sistem penerjemah bahasa isyarat yang lebih kompleks di masa mendatang, baik dalam cakupan kosakata maupun peningkatan performa model. Bagi institusi pendidikan, penelitian ini dapat menjadi contoh penerapan Artificial Intelligence yang inklusif, aplikatif, dan berorientasi pada penyelesaian permasalahan sosial.

## **I.5 Batasan Masalah**

Untuk menjaga fokus penelitian agar terarah dan dapat diselesaikan secara efektif, penelitian ini dibatasi pada hal-hal berikut:

1. Sistem hanya menerjemahkan gestur huruf alfabet BISINDO (A-Z) yang dirangkai menjadi kata sederhana, tanpa mencakup pengenalan kalimat, ekspresi wajah, maupun gesture kompleks lainnya.
2. Proses klasifikasi gestur menggunakan MediaPipe untuk ekstraksi landmark tangan dan Convolutional Neural Network (CNN) dengan pendekatan transfer learning sebagai model utama, tanpa melakukan perbandingan dengan arsitektur atau metode lain.
3. Fitur komunikasi dua arah dibatasi pada konversi dasar melalui Speech-to-Text (STT) dan Text-to-Speech (TTS) tanpa analisis semantik lanjutan, koreksi otomatis, maupun pengaturan intonasi atau dialek khusus.
4. Integrasi workflow automation menggunakan n8n dibatasi pada dokumentasi dan distribusi hasil terjemahan tertentu, tanpa pengembangan integrasi ke berbagai platform secara luas.
5. Pengujian sistem dilakukan pada dataset huruf BISINDO dalam lingkungan terkontrol, sehingga performa pada kondisi pencahayaan ekstrem, pergerakan cepat, atau variasi latar belakang kompleks tidak menjadi fokus utama.
6. Penelitian berfokus pada aspek teknis dan kinerja sistem, tanpa melakukan evaluasi mendalam terkait aspek psikologis, sosial, tingkat adopsi pengguna, maupun pengembangan aplikasi mandiri di luar lingkungan demonstrasi berbasis web.

## **I.6 Metodologi Penelitian**

Penelitian ini menggunakan pendekatan rekayasa sistem (system engineering) yang meliputi tahapan analisis kebutuhan, perancangan, pengembangan, implementasi, dan evaluasi kinerja sistem penerjemah Bahasa Isyarat Indonesia (BISINDO). Pada tahap awal, dilakukan analisis kebutuhan sistem serta pengumpulan data gestur huruf alfabet BISINDO melalui proses perekaman dan penyusunan dataset. Data citra kemudian diproses menggunakan MediaPipe untuk mengekstraksi landmark tangan sebagai fitur masukan bagi model klasifikasi.

Model klasifikasi dikembangkan menggunakan Convolutional Neural Network (CNN) dengan pendekatan transfer learning berbasis arsitektur MobileNetV2. Tahapan pengembangan meliputi preprocessing data, pembagian dataset, proses pelatihan (training), fine-tuning, serta

pengujian dalam lingkungan terkontrol. Hasil prediksi huruf selanjutnya dirangkai menjadi kata sederhana sebagai bentuk keluaran sistem.

Untuk mendukung komunikasi dua arah, sistem diintegrasikan dengan modul Speech-to-Text (STT) dan Text-to-Speech (TTS) dalam antarmuka berbasis web. Selain itu, sistem juga memanfaatkan workflow automation untuk mendukung proses dokumentasi dan distribusi hasil terjemahan secara terintegrasi.

Evaluasi kinerja model dilakukan menggunakan metrik kuantitatif seperti akurasi, precision, recall, dan F1-Score guna menilai performa klasifikasi gestur huruf BISINDO. Seluruh tahapan pengembangan dan hasil pengujian dianalisis serta didokumentasikan sebagai bagian dari laporan penelitian.

## **I.7 Sistematika Penulisan**

Adapun sistematika penulisan laporan penelitian ini disusun secara terstruktur dalam lima bab utama sebagai berikut:

### **BAB I PENDAHULUAN**

Bab ini berisi latar belakang penelitian, rumusan masalah, tujuan penelitian, manfaat penelitian, batasan masalah, metodologi penelitian, serta sistematika penulisan.

### **BAB II TINJAUAN PUSTAKA**

Bab ini membahas teori dan penelitian terdahulu yang relevan sebagai landasan dalam pengembangan sistem, meliputi:

- ✓ Bahasa Isyarat Indonesia (BISINDO).
- ✓ Konsep computer vision dan pengenalan gestur tangan.
- ✓ Framework MediaPipe untuk ekstraksi landmark tangan.
- ✓ Convolutional Neural Network (CNN), termasuk konsep transfer learning dan arsitektur MobileNetV2.
- ✓ Speech-to-Text (STT) dan Text-to-Speech (TTS).
- ✓ Visualisasi gesture dalam sistem komunikasi berbasis suara.
- ✓ Antarmuka interaktif berbasis Gradio.
- ✓ Konsep workflow automation menggunakan n8n.
- ✓ Studi terdahulu terkait sistem penerjemah bahasa isyarat.

### **BAB III METODOLOGI PENELITIAN**

Bab ini menjelaskan tahapan analisis kebutuhan, perancangan, pengembangan, dan implementasi sistem. Uraian meliputi proses pengumpulan dan preprocessing data, ekstraksi landmark menggunakan MediaPipe, perancangan dan pelatihan model CNN berbasis transfer learning, integrasi modul STT dan TTS, serta perancangan workflow automation. Diagram alur sistem dan arsitektur keseluruhan dipaparkan pada bab ini.

### **BAB IV IMPLEMENTASI DAN HASIL PENELITIAN**

Bab ini menyajikan hasil implementasi setiap komponen sistem, tampilan antarmuka berbasis web, konfigurasi lingkungan pengujian, serta hasil evaluasi kinerja model menggunakan metrik akurasi, precision, recall, dan F1-score. Selain itu, bab ini membahas analisis performa sistem komunikasi dua arah secara terintegrasi.

### **BAB V KESIMPULAN DAN SARAN**

Bab ini berisi kesimpulan yang diperoleh berdasarkan hasil penelitian serta saran untuk pengembangan lebih lanjut, seperti perluasan kosakata gestur, pengenalan ekspresi wajah, peningkatan performa model, atau pengujian pada kondisi lingkungan yang lebih beragam.



## BAB II

### TINJAUAN PUSTAKA

#### II.1 Penelitian Terdahulu

Penelitian mengenai penerjemah bahasa isyarat telah banyak dilakukan dalam beberapa tahun terakhir, khususnya dalam konteks Bahasa Isyarat Indonesia (BISINDO). Beberapa penelitian yang relevan adalah sebagai berikut:

- a. Saputra & Rakun (2024)  
Recognizing Indonesian Sign Language (Bisindo) Gesture in Complex Backgrounds. Penelitian ini menggunakan Faster R-CNN dan YOLOv5 untuk mendeteksi tangan dan wajah, MobileNetV2 untuk ekstraksi fitur, serta LSTM untuk klasifikasi. Fokus penelitian adalah pengenalan kalimat BISINDO pada latar belakang kompleks. Hasil pengujian menunjukkan sistem mampu menangani kondisi lingkungan nyata, meskipun akurasi pada tingkat kalimat masih terbatas.
- b. Damatraseta, Novariany, & Ridhani (2023)  
BISINDO Hand Gesture Detection and Recognition with Deep Learning CNN. Penelitian ini menggunakan CNN (LeNet-5 dan AlexNet) untuk mengenali huruf alfabet BISINDO. Output yang dihasilkan berupa teks alfabet, namun belum dilengkapi dengan fitur konversi suara maupun integrasi sistem komunikasi dua arah.
- c. Norick Eriyadi et al. (2024)  
Object Detection of BISINDO Sign Language Letters Using Residual Network. Penelitian ini menerapkan Residual Network (ResNet) untuk klasifikasi huruf BISINDO berbasis citra tunggal. Hasil penelitian menunjukkan performa cukup baik pada dataset terbatas, tetapi belum mencakup integrasi sistem komunikasi dua arah maupun modul konversi suara.
- d. Agustin, Maulana, & Mandyartha (2024)  
Detection of Actions BISINDO into Text-to-Speech Using Long Short-Term Memory with MediaPipe Holistics. Penelitian ini menggabungkan MediaPipe Holistics dengan LSTM untuk mendeteksi gestur BISINDO, kemudian hasilnya dikonversi menjadi suara menggunakan Text-to-Speech. Penelitian ini sudah mengarah pada sistem terintegrasi, namun masih terbatas pada lingkup gestur tertentu dan belum mencakup integrasi komunikasi dua arah maupun orkestrasi workflow otomatis.

Tabel II. 1 Ringkasan Penelitian Terdahulu

No	Judul	Penulis dan Tahun	Metode dan Teknologi Utama	Fokus dan Cakupan
1	<i>Recognizing Indonesian Sign Language (Bisindo) Gesture in Complex Backgrounds</i>	Saputra & Rakun (2024)	Faster R-CNN, YOLOv5, MobileNetV2, LSTM	Kalimat BISINDO pada latar belakang kompleks
2	<i>BISINDO Hand Gesture Detection and Recognition with Deep Learning CNN</i>	Damatraseta, Novariany, & Ridhani (2023)	CNN (LeNet-5, AlexNet)	Huruf alfabet BISINDO berbasis CNN saja
3	<i>Object Detection of BISINDO Sign Language Letters Using Residual Network</i>	Norick Eriyadi et al. (2024)	ResNet	Deteksi huruf BISINDO berbasis citra tunggal
4	<i>Detection of Actions BISINDO into Text-to-Speech Using LSTM with MediaPipe Holistics</i>	Agustin, Maulana, & Mandyartha (2024)	MediaPipe Holistics, LSTM, TTS	Gestur BISINDO ke suara (TTS)

Tabel II. 2 Perbandingan Penelitian Terdahulu

No	Penelitian	Bahasa Isyarat	Unit yang dikenali	TTS	Media Pipe	Workflow	Dua Arah (STT)
1	Saputra & Rakun (2024)	BISINDO	Kalimat	X	X	X	X
2	Damatraseta et al. (2023)	BISINDO	Huruf alfabet	X	X	X	X

3	Norick Eriyadi et al. (2024)	BISINDO	Huruf alfabet	X	X	X	X
4	Agustin et al. (2024)	BISINDO	Gestur/ aksi	✓ (TTS)	✓	X	X
5	Penelitian ini	BISINDO	Huruf → Kata	✓	✓	✓	✓

Berdasarkan kajian terhadap beberapa penelitian terdahulu, dapat disimpulkan bahwa:

- Sebagian penelitian masih berfokus pada pengenalan huruf atau gestur tunggal tanpa integrasi sistem komunikasi dua arah.
- Pendekatan deep learning seperti CNN dan ResNet telah digunakan, namun belum banyak yang mengoptimalkan transfer learning dengan arsitektur ringan untuk efisiensi implementasi berbasis web.
- Beberapa penelitian telah mengintegrasikan MediaPipe dan Text-to-Speech, tetapi penerapannya masih parsial dan belum membentuk sistem komunikasi yang terintegrasi secara menyeluruh.
- Belum banyak penelitian yang menggabungkan klasifikasi gestur, konversi suara dua arah (STT dan TTS), serta workflow automation dalam satu sistem terpadu berbasis BISINDO.

Berdasarkan celah penelitian tersebut, penelitian ini mengusulkan pengembangan sistem penerjemah BISINDO dengan kontribusi utama sebagai berikut:

- Pemanfaatan MediaPipe untuk ekstraksi landmark tangan sebagai fitur input model.
- Penggunaan Convolutional Neural Network (CNN) dengan pendekatan transfer learning berbasis MobileNetV2 untuk klasifikasi huruf alfabet BISINDO.
- Penyusunan hasil prediksi huruf menjadi kata sederhana sebagai keluaran sistem.
- Integrasi Text-to-Speech (TTS) dan Speech-to-Text (STT) untuk mendukung komunikasi dua arah.
- Pemanfaatan workflow automation untuk mendukung dokumentasi dan distribusi hasil terjemahan secara terintegrasi.

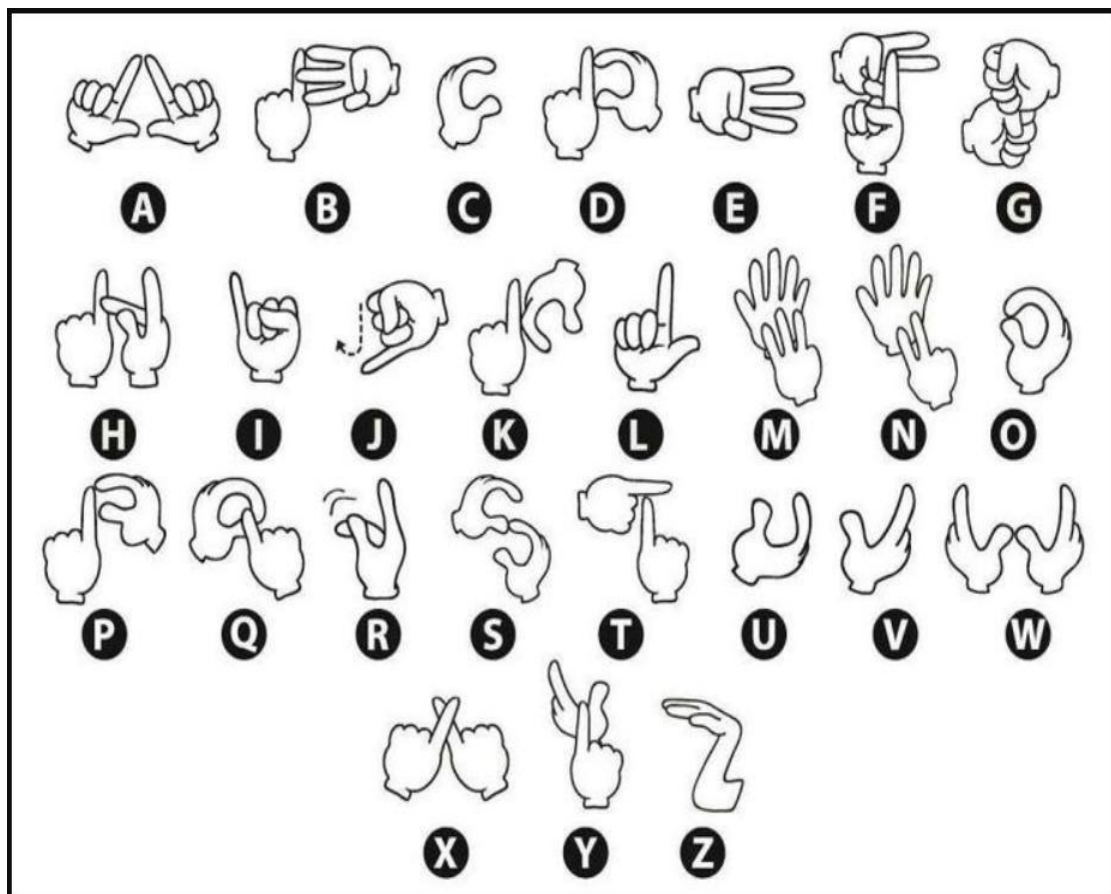
Dengan demikian, penelitian ini menawarkan pendekatan sistem terintegrasi yang mencakup proses ekstraksi fitur, klasifikasi berbasis deep learning, konversi suara dua arah,

serta integrasi workflow dalam satu platform berbasis web. Pendekatan ini menjadi pembeda dibandingkan penelitian terdahulu yang umumnya masih bersifat parsial atau berfokus pada satu komponen tertentu saja.

## II.2 Kajian Teoritis

### II.2.1 Bahasa Isyarat Indonesia (BISINDO)

Bahasa Isyarat Indonesia (BISINDO) adalah bahasa visual alami yang digunakan komunitas tunarungu di Indonesia. BISINDO memiliki struktur linguistik sendiri, berbeda dengan Sistem Isyarat Bahasa Indonesia (SIBI) yang lebih cenderung transliteratif terhadap bahasa lisan. Dalam BISINDO, komunikasi dilakukan dengan kombinasi gerakan tangan, ekspresi wajah, dan posisi tubuh (Munandar et al. 2024).



Gambar II.1 Huruf Alfabet BISINDO (Sumber:Pinterest)

Keunggulan BISINDO adalah sifatnya yang lebih mudah dipahami oleh komunitas tunarungu karena berkembang secara alami. Namun, keterbatasan masyarakat umum dalam memahami BISINDO menyebabkan hambatan komunikasi. Oleh sebab itu, sistem

penerjemah BISINDO berbasis teknologi menjadi solusi untuk menjembatani komunikasi antara penyandang tunarungu dan masyarakat luas.

## **II.2.2 Computer Vision dalam Pengenalan Gesture**

Computer Vision merupakan cabang kecerdasan buatan yang berfokus pada kemampuan sistem komputer untuk memahami dan menginterpretasikan informasi visual dari citra digital. Dalam konteks pengenalan gesture, computer vision digunakan untuk mengekstraksi pola atau fitur dari gambar tangan yang kemudian dianalisis untuk menentukan kelas tertentu (Widyagama 2023).

Secara umum, proses computer vision melibatkan beberapa tahapan, yaitu akuisisi citra, preprocessing, ekstraksi fitur, serta klasifikasi. Pada penelitian ini, citra yang digunakan berasal dari gambar gesture tangan yang diunggah oleh pengguna. Gambar tersebut kemudian diproses untuk memperoleh representasi fitur yang relevan sebelum dilakukan proses klasifikasi menggunakan model pembelajaran mendalam.

Pendekatan berbasis citra statis dipilih karena lebih stabil dalam pengujian dan memudahkan proses pelatihan model. Dengan memanfaatkan teknik computer vision, sistem dapat mengenali pola visual pada gesture huruf BISINDO dan mengubahnya menjadi representasi teks sesuai kelas yang telah dipelajari model.

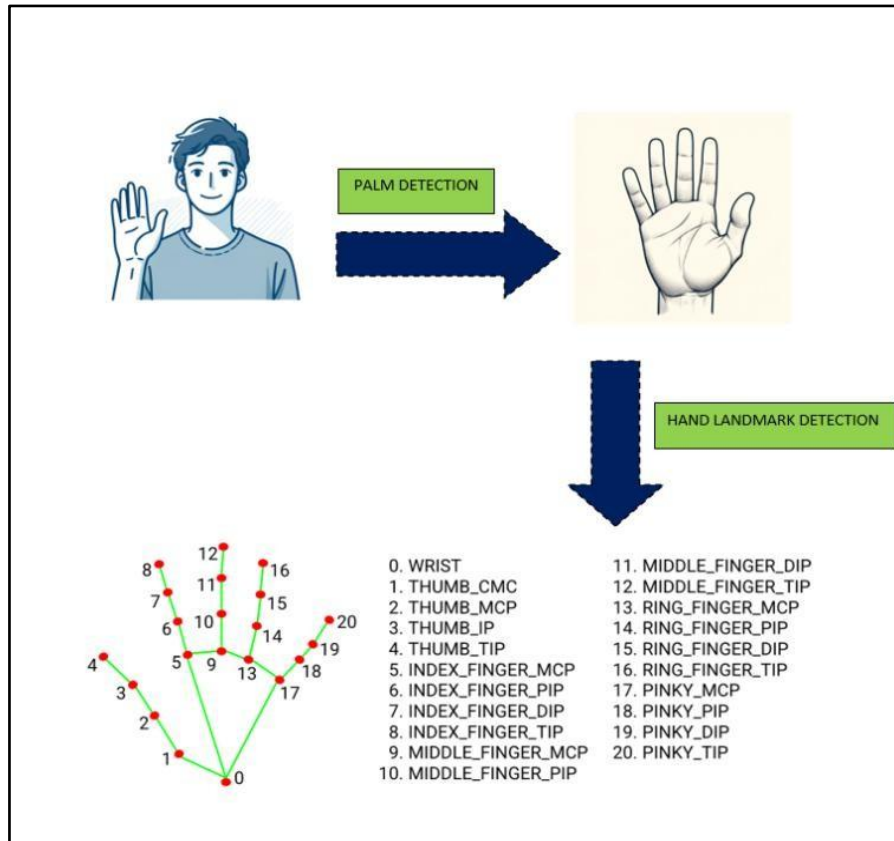
## **II.2.3 MediaPipe untuk Ekstraksi Landmark Tangan**

MediaPipe merupakan framework yang dikembangkan oleh Google untuk pemrosesan data multimodal, termasuk citra, video, dan sinyal sensor. Salah satu modul yang banyak digunakan adalah MediaPipe Hands, yang dirancang untuk mendeteksi serta melacak posisi tangan dan menghasilkan titik-titik landmark secara terstruktur.

MediaPipe Hands mampu mengidentifikasi 21 titik landmark pada tangan, yang merepresentasikan posisi setiap ruas jari dan telapak tangan dalam bentuk koordinat numerik. Landmark tersebut menjadi representasi fitur yang lebih terstruktur dibandingkan citra mentah, sehingga memudahkan proses analisis dan klasifikasi menggunakan model pembelajaran mendalam.

Dalam penelitian ini, MediaPipe digunakan untuk mengekstraksi landmark tangan dari gambar gesture yang diunggah pengguna. Proses ini bertujuan untuk memperoleh fitur spasial yang merepresentasikan bentuk dan posisi tangan pada setiap huruf alfabet BISINDO.

Hasil ekstraksi landmark kemudian digunakan sebagai input pada model klasifikasi berbasis Convolutional Neural Network (CNN).

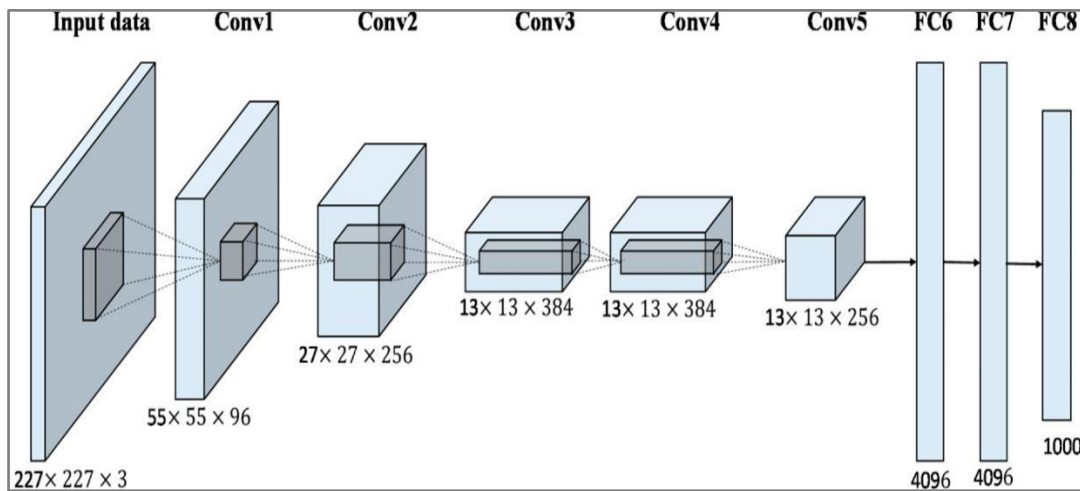


Gambar II.2 Landmark Hand (Sumber:Google)

Penggunaan MediaPipe dipilih karena kemampuannya dalam mendeteksi struktur tangan secara stabil serta menghasilkan data numerik yang efisien untuk diproses lebih lanjut. Dengan demikian, MediaPipe berperan sebagai tahap awal dalam pipeline sistem sebelum dilakukan proses klasifikasi huruf (Lemmuela et al. 2025).

## II.2.4 Convolutional Neural Network (CNN)

Convolutional Neural Network (CNN) merupakan salah satu metode deep learning yang banyak digunakan dalam pengolahan citra dan pengenalan pola visual. CNN dirancang untuk secara otomatis mengekstraksi fitur dari data gambar melalui proses konvolusi, sehingga mampu mengenali pola seperti bentuk, tepi, dan struktur objek (Fadillah et al. 2022).



*Gambar II.3 Arsitektur Umum CNN (Sumber: Google)*

Secara umum, arsitektur CNN terdiri dari beberapa lapisan utama, yaitu convolution layer, activation function, pooling layer, dan fully connected layer. Convolution layer berfungsi untuk mengekstraksi fitur dari citra menggunakan filter atau kernel. Activation function, seperti ReLU, digunakan untuk menambahkan sifat non-linear pada model. Pooling layer bertujuan untuk mengurangi dimensi fitur sekaligus mempertahankan informasi penting. Selanjutnya, fully connected layer digunakan untuk melakukan proses klasifikasi berdasarkan fitur yang telah diekstraksi.

Dalam penelitian ini, CNN digunakan sebagai model klasifikasi untuk mengenali huruf alfabet BISINDO berdasarkan fitur visual yang telah diproses sebelumnya. Kemampuan CNN dalam mempelajari pola spasial dari data citra menjadikannya metode yang sesuai untuk sistem pengenalan gesture berbasis gambar.

Penggunaan CNN memungkinkan proses klasifikasi dilakukan secara otomatis tanpa perlu perancangan fitur manual, sehingga model dapat menyesuaikan diri dengan variasi bentuk gesture pada setiap huruf.

## **II.2.5 Arsitektur MobileNetV2 dan Transfer Learning**

MobileNetV2 merupakan salah satu arsitektur Convolutional Neural Network (CNN) yang dirancang untuk efisiensi komputasi dengan tetap mempertahankan performa klasifikasi yang baik. Arsitektur ini menggunakan konsep depthwise separable convolution, yaitu pemisahan proses konvolusi menjadi dua tahap: depthwise convolution dan pointwise convolution. Pendekatan ini mampu mengurangi jumlah parameter dan kompleksitas perhitungan dibandingkan CNN konvensional, sehingga lebih ringan dan efisien (Lemmuella et al. 2025).

Selain itu, MobileNetV2 menerapkan konsep inverted residual block dan linear bottleneck, yang bertujuan untuk mempertahankan informasi penting selama proses ekstraksi fitur sekaligus mengurangi dimensi fitur secara optimal. Struktur ini memungkinkan model tetap akurat meskipun digunakan pada perangkat dengan sumber daya terbatas.

Dalam penelitian ini, MobileNetV2 digunakan sebagai model dasar (pre-trained model) yang telah dilatih sebelumnya pada dataset skala besar. Untuk menyesuaikan model dengan dataset gesture huruf BISINDO, diterapkan metode transfer learning. Transfer learning merupakan teknik pemanfaatan pengetahuan dari model yang telah dilatih sebelumnya untuk menyelesaikan tugas baru yang serupa. Dengan pendekatan ini, proses pelatihan menjadi lebih efisien dan tidak memerlukan dataset dalam jumlah sangat besar.

Selanjutnya, dilakukan proses fine-tuning dengan menyesuaikan beberapa lapisan akhir model agar lebih adaptif terhadap karakteristik gesture huruf BISINDO. Kombinasi MobileNetV2 dan transfer learning dipilih karena mampu meningkatkan akurasi klasifikasi sekaligus menjaga efisiensi proses pelatihan (Kelana et al. 2024).

## II.2.6 Text-to-Speech (TTS)

Text-to-Speech (TTS) merupakan teknologi yang digunakan untuk mengubah teks menjadi suara secara otomatis. Sistem TTS bekerja dengan memproses input berupa teks, kemudian menghasilkan keluaran dalam bentuk sinyal audio yang dapat dipahami sebagai ujaran manusia. Teknologi ini banyak dimanfaatkan dalam sistem asistif untuk membantu komunikasi, termasuk bagi penyandang disabilitas (Rachmawati et al. 2023).

Secara umum, proses TTS melibatkan beberapa tahapan, yaitu analisis teks, pemodelan fonetik, serta sintesis suara. Pada tahap analisis, sistem mengidentifikasi struktur kata dan karakter. Selanjutnya, teks diubah menjadi representasi fonem, yang kemudian disintesis menjadi gelombang suara melalui model suara tertentu.

Dalam penelitian ini, TTS digunakan sebagai media keluaran suara dari hasil klasifikasi gesture huruf BISINDO yang telah disusun menjadi teks. Implementasi TTS memungkinkan hasil terjemahan tidak hanya ditampilkan dalam bentuk tulisan, tetapi juga dapat didengar oleh pengguna lainnya, sehingga mendukung proses komunikasi dua arah.

Penggunaan teknologi TTS pada sistem ini bertujuan untuk meningkatkan aksesibilitas dan memperluas fungsi sistem penerjemah, sehingga tidak terbatas pada visualisasi teks saja.

### **II.2.7 Speech-to-Text (STT)**

Speech-to-Text (STT) merupakan teknologi yang berfungsi untuk mengubah sinyal suara menjadi teks secara otomatis. Sistem STT bekerja dengan mengenali pola gelombang suara, kemudian mengonversinya menjadi representasi fonetik yang selanjutnya diterjemahkan menjadi kata atau kalimat dalam bentuk teks (Yuniarno 2025).

Secara umum, proses STT melibatkan tahapan perekaman suara, ekstraksi fitur audio, serta pemodelan bahasa untuk menentukan hasil transkripsi yang paling sesuai. Teknologi ini memanfaatkan model pembelajaran mesin untuk mengenali variasi pengucapan, intonasi, dan karakteristik suara pengguna.

Dalam penelitian ini, STT digunakan sebagai bagian dari komunikasi dua arah, yaitu untuk mengubah suara pengguna menjadi teks yang kemudian dapat ditampilkan atau diproses lebih lanjut dalam sistem. Integrasi STT memungkinkan pengguna yang tidak menggunakan bahasa isyarat tetap dapat berinteraksi melalui media suara, sehingga sistem menjadi lebih inklusif (Sanjaya & Ilone 2023).

Dengan adanya kombinasi TTS dan STT, sistem tidak hanya berfungsi sebagai penerjemah gesture ke teks dan suara, tetapi juga mendukung alur komunikasi timbal balik antara pengguna bahasa isyarat dan pengguna bahasa lisan.

### **II.2.8 Antarmuka Interaktif Berbasis Gradio**

Gradio merupakan framework berbasis Python yang digunakan untuk membangun antarmuka interaktif bagi aplikasi machine learning secara sederhana dan cepat. Framework ini memungkinkan pengembang untuk menghubungkan model yang telah dilatih dengan komponen input dan output berbasis web tanpa perlu membangun sistem front-end secara kompleks (Jiang 2024).

Gradio menyediakan berbagai komponen input seperti unggah gambar, teks, dan audio, serta komponen output seperti tampilan teks, gambar, dan suara. Kemudahan integrasi ini menjadikan Gradio banyak digunakan dalam tahap prototyping maupun demonstrasi sistem berbasis kecerdasan buatan.

Dalam penelitian ini, Gradio digunakan sebagai media antarmuka untuk menerima input berupa gambar gesture yang diunggah pengguna, menampilkan hasil klasifikasi dalam bentuk teks, serta memutar keluaran suara hasil konversi Text-to-Speech (TTS). Dengan

menggunakan Gradio, sistem dapat dioperasikan melalui browser secara interaktif tanpa memerlukan pengembangan aplikasi web secara mandiri.

Penggunaan Gradio mendukung kemudahan pengujian dan demonstrasi sistem, sehingga pengguna dapat berinteraksi langsung dengan model penerjemah gesture yang telah dikembangkan.

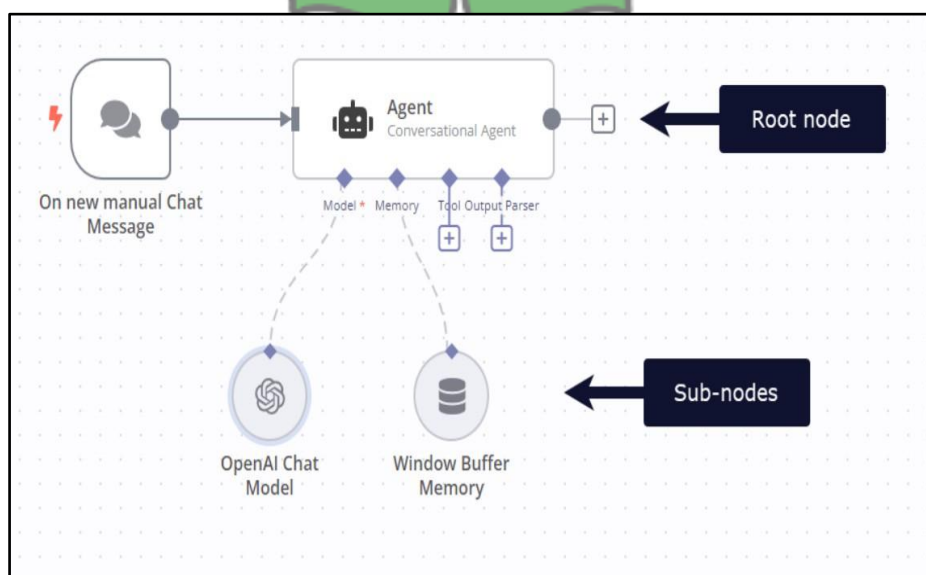
### II.2.9 Google Colab

Google Colab adalah platform berbasis cloud yang memungkinkan peneliti menjalankan kode Python tanpa konfigurasi rumit di komputer lokal. Colab mendukung penggunaan GPU, yang sangat penting untuk mempercepat proses pelatihan model deep learning (Nugroho et al. 2023).

Dalam penelitian ini, Google Colab digunakan sebagai lingkungan pelatihan model CNN, mulai dari preprocessing dataset hingga evaluasi model. Keunggulannya adalah gratis, mudah digunakan, dan dapat diakses dari mana saja.

### II.2.10 Workflow Automation dengan n8n

n8n merupakan platform workflow automation berbasis open-source yang digunakan untuk menghubungkan berbagai layanan dan aplikasi secara otomatis melalui mekanisme alur kerja (workflow). n8n memungkinkan integrasi antar sistem menggunakan webhook dan berbagai konektor API tanpa memerlukan pengembangan backend yang kompleks (Eriyadi et al. 2024).



Gambar II.4 Orkestrasi Workflow n8n (Sumber:Google)

Secara umum, workflow pada n8n terdiri dari trigger, proses pengolahan data, dan aksi keluaran. Trigger dapat berupa permintaan webhook atau event tertentu, kemudian data diproses sesuai kebutuhan, dan hasilnya dikirim ke layanan lain seperti database, spreadsheet, atau aplikasi pesan.

Dalam penelitian ini, n8n digunakan sebagai orkestrator untuk mengelola hasil penerjemahan sistem. Teks hasil klasifikasi gesture maupun hasil konversi suara dapat dikirim secara otomatis ke Google Spreadsheet untuk keperluan pencatatan, serta diteruskan ke platform komunikasi seperti WhatsApp. Integrasi ini memungkinkan sistem tidak hanya berfungsi sebagai alat penerjemah, tetapi juga mendukung dokumentasi dan distribusi informasi secara terstruktur.

Penggunaan n8n memberikan nilai tambah pada penelitian ini karena menghadirkan integrasi end-to-end, mulai dari proses pengenalan gesture hingga pengiriman hasil ke platform eksternal secara otomatis.

### **II.2.11 Evaluasi Sistem**

Evaluasi sistem dilakukan untuk menilai kinerja dan efektivitas model serta integrasi komponen yang digunakan dalam penelitian ini. Proses evaluasi difokuskan pada model klasifikasi gesture huruf BISINDO, komponen komunikasi dua arah, serta integrasi workflow otomatis.

Pada model klasifikasi berbasis Convolutional Neural Network (CNN), evaluasi dilakukan menggunakan metrik akurasi, precision, recall, F1-score, dan confusion matrix. Akurasi mengukur persentase prediksi yang benar terhadap seluruh data uji. Precision menunjukkan tingkat ketepatan model dalam mengklasifikasikan suatu kelas huruf tertentu, sedangkan recall mengukur kemampuan model dalam mengenali seluruh data yang termasuk dalam kelas tersebut. F1-score merupakan rata-rata harmonis antara precision dan recall yang memberikan gambaran keseimbangan performa model. Confusion matrix digunakan untuk melihat distribusi prediksi benar dan salah pada setiap kelas huruf alfabet BISINDO.

Untuk komponen Speech-to-Text (STT), evaluasi dilakukan dengan membandingkan hasil transkripsi sistem terhadap teks referensi. Pengukuran dapat dilakukan menggunakan Word Error Rate (WER) atau Confidence untuk mengetahui tingkat kesalahan pengenalan kata, serta tingkat akurasi kalimat untuk melihat kesesuaian hasil transkripsi secara keseluruhan.

Evaluasi Text-to-Speech (TTS) dilakukan secara fungsional dengan memastikan bahwa teks hasil klasifikasi dapat dikonversi menjadi suara secara utuh dan dapat dipahami dengan jelas. Aspek yang diperhatikan meliputi kesesuaian antara teks dan suara yang dihasilkan serta keberhasilan proses konversi.

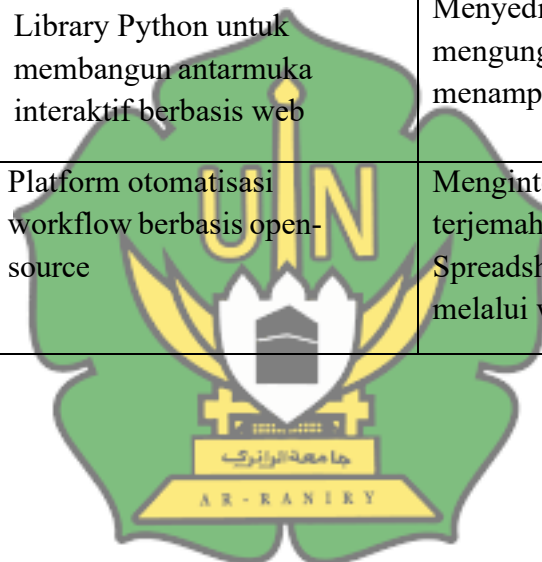
Pada integrasi workflow menggunakan n8n, evaluasi difokuskan pada reliabilitas pengiriman data. Sistem diuji untuk memastikan bahwa hasil terjemahan dapat terkirim dengan benar ke layanan eksternal seperti Google Spreadsheet melalui mekanisme webhook yang telah dikonfigurasi. Tingkat keberhasilan pengiriman menjadi indikator utama dalam penilaian integrasi sistem.

Dengan pendekatan evaluasi tersebut, kinerja sistem dapat dianalisis secara menyeluruh, baik dari sisi akurasi model klasifikasi maupun keberhasilan integrasi antar komponen.

*Tabel II. 3 Ringkasan Kajian Teoritis*

Komponen	Definisi	Peran dalam Penelitian
<b>BISINDO</b>	Bahasa visual alami komunitas tunarungu di Indonesia yang menggunakan gerakan tangan dan ekspresi non-verbal	Menjadi objek utama sistem, dengan fokus pada pengenalan huruf alfabet A-Z yang dapat disusun menjadi kata
<b>Computer Vision</b>	Bidang kecerdasan buatan yang memungkinkan sistem memahami dan menganalisis citra digital	Digunakan untuk memproses gambar gesture yang diunggah pengguna sebelum dilakukan klasifikasi
<b>MediaPipe</b>	Framework dari Google untuk mendeteksi dan mengekstraksi landmark tangan dari citra	Mengekstraksi 21 titik landmark tangan sebagai representasi fitur gesture
<b>Convolutional Neural Network (CNN)</b>	Arsitektur deep learning untuk pengenalan pola visual pada citra	Mengklasifikasikan gesture tangan menjadi huruf alfabet BISINDO
<b>MobileNetV2</b>	Arsitektur CNN ringan berbasis depthwise separable convolution	Digunakan sebagai model dasar dalam proses klasifikasi gesture

<b>Transfer Learning</b>	Teknik pemanfaatan model yang telah dilatih sebelumnya untuk tugas baru yang serupa	Mempercepat proses pelatihan model dengan menyesuaikan MobileNetV2 pada dataset BISINDO
<b>Google Colab</b>	Platform komputasi berbasis cloud untuk menjalankan kode Python	Digunakan untuk pelatihan dan evaluasi model
<b>Text-to-Speech (TTS)</b>	Teknologi yang mengubah teks menjadi suara sintetis	Mengonversi hasil teks klasifikasi gesture menjadi audio
<b>Speech-to-Text (STT)</b>	Teknologi yang mengubah suara menjadi teks	Menerjemahkan ucapan pengguna menjadi teks sebagai bagian komunikasi dua arah
<b>Gradio</b>	Library Python untuk membangun antarmuka interaktif berbasis web	Menyediakan antarmuka untuk mengunggah gambar gesture dan menampilkan hasil klasifikasi
<b>n8n</b>	Platform otomatisasi workflow berbasis open-source	Mengintegrasikan hasil terjemahan ke Google Spreadsheet dan WhatsApp melalui webhook



## BAB III

### METODOLOGI PENELITIAN

#### III.1 Jenis Penelitian

Penelitian ini merupakan penelitian terapan (applied research) yang berfokus pada implementasi sistem penerjemahan Bahasa Isyarat Indonesia (BISINDO) dua arah berbasis MediaPipe dan Artificial Intelligence. Penelitian bertujuan membangun sistem yang mampu menerjemahkan gestur tangan menjadi teks dan suara, serta menerjemahkan suara menjadi teks dan visual huruf BISINDO secara real-time.

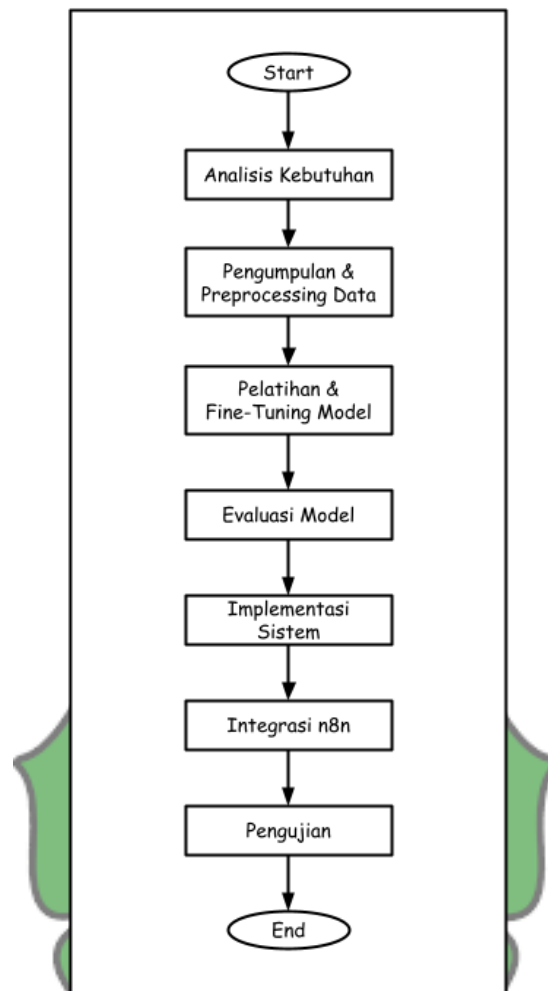
Pendekatan yang digunakan adalah rekayasa sistem (system engineering approach) yang meliputi tahapan perancangan arsitektur, pengumpulan dataset, pelatihan model deep learning menggunakan metode transfer learning dan fine-tuning bertahap, implementasi sistem real-time berbasis web, serta evaluasi performa sistem.

Dataset citra gestur BISINDO dikumpulkan secara mandiri dan disimpan pada Google Drive, kemudian diakses melalui proses mounting pada Google Colab untuk proses pelatihan model. Pelatihan dilakukan dalam dua tahap. Tahap pertama menghasilkan model awal yang disimpan dalam format *model\_bisindo.h5*. Selanjutnya dilakukan proses fine-tuning untuk meningkatkan kemampuan adaptasi model terhadap karakteristik dataset BISINDO, sehingga diperoleh model akhir yang disimpan dalam format *model\_finetune.h5*.

Model *model\_finetune.h5* kemudian diintegrasikan ke dalam sistem antarmuka berbasis Gradio yang memanfaatkan MediaPipe untuk deteksi tangan serta modul *Speech-to-Text* dan *Text-to-Speech* untuk mendukung komunikasi dua arah.

Metode penelitian bersifat kuantitatif, karena kinerja sistem diukur menggunakan metrik evaluasi klasifikasi dan pengujian performa sistem secara langsung.

## III.2 Tahapan Penelitian



*Gambar III. 1 Tahapan Penelitian*

Tahapan penelitian dalam pengembangan sistem penerjemahan BISINDO dua arah dilakukan secara sistematis sebagai berikut:

- a. Analisis Kebutuhan Sistem
  - Kebutuhan penerjemahan dua arah antara gestur dan suara
  - Kebutuhan sistem deteksi area tangan untuk meningkatkan fokus klasifikasi
  - Kebutuhan model klasifikasi berbasis deep learning
  - Kebutuhan modul konversi suara dan teks
  - Kebutuhan pencatatan hasil terjemahan untuk dokumentasi dan monitoring sistemHasil tahap ini menjadi dasar perancangan arsitektur sistem.
- b. Pengumpulan dan Penyusunan Dataset

Dataset citra gestur huruf BISINDO(A-Z) dikumpulkan secara mandiri menggunakan webcam. Dataset disusun dalam struktur folder berdasarkan kelas huruf dan disimpan

pada Google Drive. Dataset kemudian diakses melalui Google Colab dengan proses mounting untuk keperluan pelatihan model klasifikasi.

Dataset yang digunakan dalam penelitian ini terdiri dari 4458 citra gestur huruf BISINDO yang terbagi ke dalam 26 kelas (A-Z). Rata-rata jumlah citra per kelas adalah  $\pm 171$  gambar. Dataset kemudian dibagi menjadi data training dan validation dengan rasio 80:20 (training 3566 dan validation 892 gambar).

### c. Preprocessing Data Awal

Sebelum proses pelatihan model, citra gestur BISINDO diproses melalui beberapa tahapan preprocessing untuk menyesuaikan dengan kebutuhan arsitektur MobileNetV2 dan meningkatkan kemampuan generalisasi model. Tahapan tersebut meliputi:

- Resize citra ke ukuran  $128 \times 128$  piksel agar sesuai dengan dimensi input model.
- Normalisasi nilai piksel dengan skala 0–1 menggunakan teknik rescaling ( $1./255$ ).
- Augmentasi data menggunakan ImageDataGenerator untuk menghasilkan variasi citra melalui rotasi, pergeseran, zoom, dan flipping.
- Pembagian dataset menjadi data training dan validation dengan rasio tertentu untuk keperluan evaluasi model.

Tahapan preprocessing ini bertujuan untuk meningkatkan stabilitas pelatihan dan mengurangi risiko overfitting pada model

### d. Pelatihan Model Tahap Pertama

Pada tahap ini, model dikembangkan menggunakan pendekatan transfer learning dengan memanfaatkan arsitektur MobileNetV2 yang telah memiliki bobot awal (pretrained weights) dari dataset ImageNet. Seluruh layer pada base model MobileNetV2 dibekukan (freeze) sehingga berfungsi sebagai feature extractor dan tidak mengalami pembaruan bobot selama proses pelatihan awal.

Selanjutnya, ditambahkan layer klasifikasi berupa Global Average Pooling dan Dense layer dengan fungsi aktivasi softmax sesuai jumlah kelas huruf BISINDO (26 kelas).

Model kemudian dikompilasi menggunakan optimizer Adam dengan fungsi loss categorical crossentropy serta metrik evaluasi accuracy. Model dilatih menggunakan dataset BISINDO yang telah melalui proses preprocessing. Hasil pelatihan tahap pertama disimpan dalam format file bisindo.h5.

e. Fine-Tuning Model

Setelah tahap pelatihan awal selesai, dilakukan proses fine-tuning untuk meningkatkan kemampuan adaptasi model terhadap karakteristik spesifik dataset huruf BISINDO. Pada tahap ini, sebagian layer akhir dari MobileNetV2 dibuka kembali (unfreeze) sehingga bobotnya dapat diperbarui selama proses pelatihan.

Proses fine-tuning dilakukan dengan menggunakan learning rate yang lebih kecil dibandingkan tahap pelatihan awal guna menjaga stabilitas pembaruan bobot serta menghindari perubahan parameter yang terlalu besar pada model pretrained.

Model kemudian dilatih kembali selama beberapa epoch tambahan hingga menunjukkan performa yang stabil. Hasil fine-tuning ini akan disimpan sebagai model\_finetune.h5 dan digunakan sebagai model final dalam sistem penerjemahan.

Proses fine-tuning direncanakan dilakukan selama beberapa epoch tambahan dengan learning rate yang lebih kecil untuk meningkatkan adaptasi model terhadap dataset BISINDO.

f. Implementasi Sistem berbasis web

Model model\_finetune.h5 diintegrasikan dengan:

- MediaPipe digunakan untuk deteksi dan ekstraksi area tangan pada gambar input sebelum proses klasifikasi
- OpenCV untuk preprocessing citra
- Gradio sebagai antarmuka web
- Modul STT dan TTS untuk komunikasi dua arah

Sistem dirancang untuk memproses input secara otomatis setelah pengguna memberikan masukan melalui unggah gambar atau audio.

g. Integrasi Webhook dan Pengujian Sistem

Sistem dihubungkan dengan webhook n8n untuk:

- Penyimpanan hasil terjemahan ke Google Spreadsheet

Setelah seluruh komponen terintegrasi, dilakukan pengujian performa fungsionalitas sistem secara menyeluruh.

### III.3 Sistem Penerjemahan Bahasa Isyarat Bisindo

Perancangan sistem dilakukan untuk membangun sistem penerjemahan Bahasa Isyarat Indonesia (BISINDO) dua arah berbasis MediaPipe dan Artificial Intelligence yang dirancang

berbasis web dan memproses input yang diberikan pengguna melalui mekanisme unggah gambar untuk mode gestur serta perekaman atau unggah file audio untuk mode suara.

Sistem dirancang dalam bentuk arsitektur modular agar setiap komponen dapat dikembangkan dan diuji secara terpisah sebelum diintegrasikan secara menyeluruh.

### III.3.1 Arsitektur Sistem

Arsitektur sistem terdiri dari enam modul utama sebagai berikut:

a. Modul Input Gestur (Unggah Gambar)

Modul ini menerima input berupa gambar gestur tangan yang diunggah oleh pengguna melalui antarmuka web. Gambar yang diterima kemudian diteruskan ke tahap deteksi tangan.

b. Modul Deteksi dan Ekstraksi Area Tangan (MediaPipe)

Modul ini menggunakan MediaPipe Hands untuk mendeteksi posisi tangan pada gambar yang diunggah. MediaPipe menghasilkan informasi posisi landmark tangan yang digunakan untuk menentukan area tangan. Area tersebut kemudian diproses lebih lanjut sebelum diklasifikasikan.

c. Modul Preprocessing Citra

Citra hasil deteksi tangan disesuaikan ukurannya menjadi  $128 \times 128$  piksel agar sesuai dengan konfigurasi input model MobileNetV2. Tahap ini memastikan konsistensi dimensi input terhadap model hasil pelatihan.

d. Modul Klasifikasi Huruf (Artificial Intelligence)

Modul ini menggunakan model deep learning berbasis Convolutional Neural Network (CNN) dengan arsitektur MobileNetV2 hasil fine-tuning yang disimpan dalam format model\_finetune.h5. Model melakukan klasifikasi terhadap citra tangan dan menghasilkan prediksi huruf alfabet A-Z.

e. Modul Konversi Suara dan Visualisasi

Modul ini terdiri dari dua bagian:

- Speech-to-Text (STT): Mengubah input suara yang direkam atau diunggah pengguna menjadi teks Bahasa Indonesia.
- Text-to-Speech (TTS): Mengubah teks hasil klasifikasi atau konversi menjadi keluaran suara.

- Visualisasi Huruf BISINDO: Menampilkan representasi gambar huruf BISINDO berdasarkan teks yang dihasilkan.

f. Modul Integrasi dan Penyimpanan Data

Modul ini memungkinkan hasil terjemahan dikirim dan disimpan melalui mekanisme integrasi berbasis webhook untuk keperluan dokumentasi komunikasi.

### III.3.2 Perancangan 2 Mode (Gesture dan Suara)

Sistem dirancang untuk mendukung dua skenario komunikasi utama:

a. Mode Gestur → Teks → Suara

Pada mode ini, sistem menerima input berupa gambar gestur tangan yang diunggah oleh pengguna. Gambar tersebut diproses oleh MediaPipe untuk mendeteksi area tangan, kemudian disesuaikan ukurannya menjadi 128×128 piksel sebelum diklasifikasikan oleh model MobileNetV2.

Hasil klasifikasi berupa huruf alfabet dirangkai menjadi teks. Teks tersebut selanjutnya dapat dikonversi menjadi suara menggunakan modul Text-to-Speech.

Mode ini dirancang untuk membantu penyandang tunarungu menyampaikan pesan kepada pengguna umum.

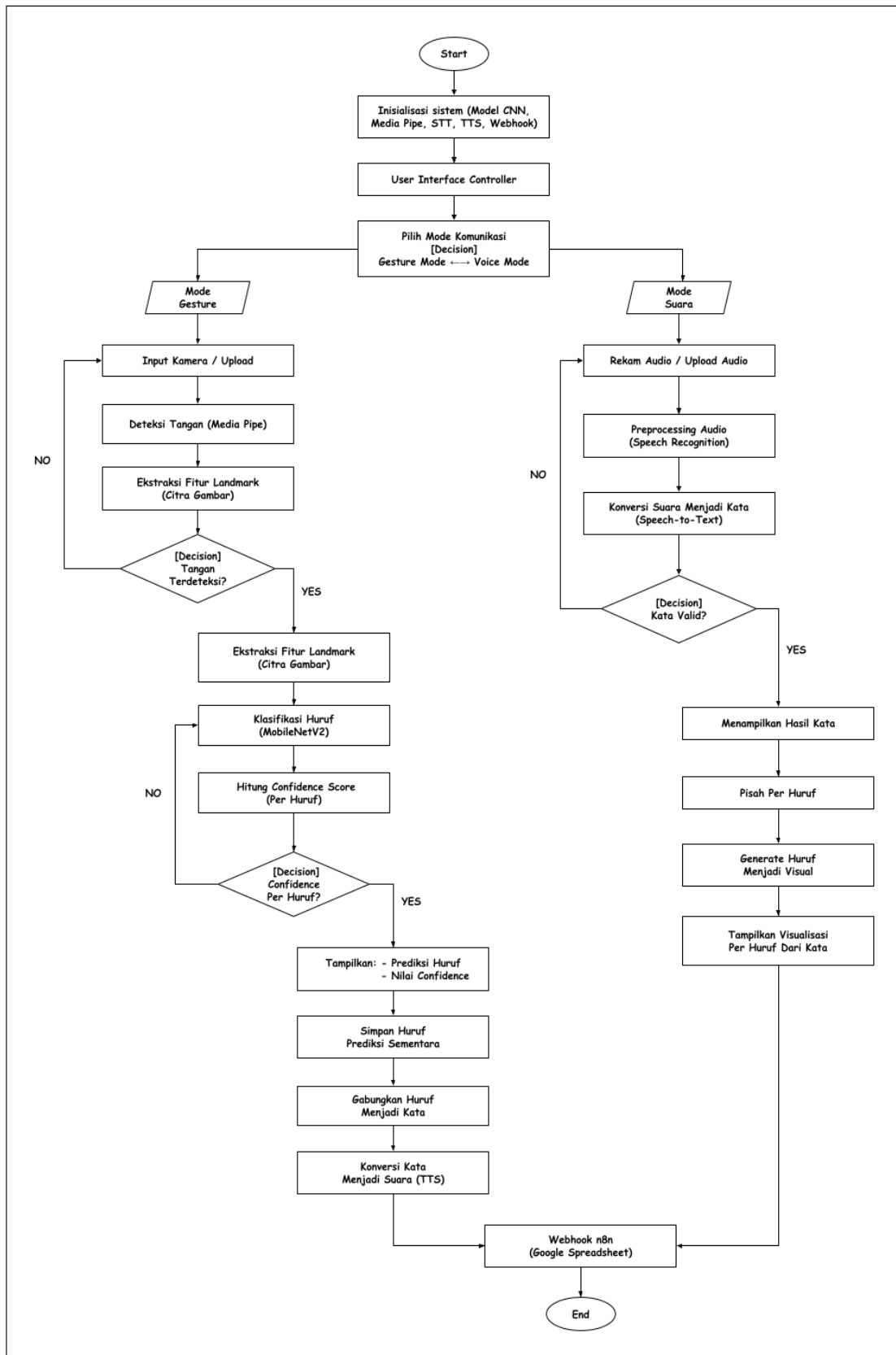
b. Mode Suara → Teks → Visual Huruf

Pada mode ini, sistem menerima input berupa suara yang direkam secara langsung melalui mikrofon atau diunggah dalam bentuk file audio. Audio tersebut diproses menggunakan modul Speech-to-Text untuk menghasilkan teks.

Teks yang dihasilkan kemudian divisualisasikan dalam bentuk representasi huruf BISINDO secara berurutan, sehingga pengguna tunarungu dapat memahami pesan yang disampaikan.

Mode ini mendukung komunikasi dua arah secara inklusif antara pengguna umum dan penyandang tunarungu.

### III.3.3 Perancangan Alur Sistem (System Flow)



Gambar III. 2 Perancangan Sistem Mode Gesture dan Suara

Alur data sistem dirancang berdasarkan dua mode komunikasi yang tersedia, sebagai berikut:

- a. Pengguna memilih mode komunikasi (Mode Gestur atau Mode Suara).
- b. Sistem menerima input sesuai mode yang dipilih:
  - Mode Gestur: unggahan gambar gestur tangan.
  - Mode Suara: rekaman atau unggahan file audio.
- c. Data diproses oleh modul yang sesuai:
  - Pada Mode Gestur, gambar diproses melalui MediaPipe untuk deteksi tangan, kemudian diklasifikasikan oleh model MobileNetV2.
  - Pada Mode Suara, audio diproses menggunakan modul Speech-to-Text untuk menghasilkan teks.
- d. Hasil pemrosesan berupa teks.
- e. Teks yang dihasilkan dapat:
  - Dikonversi menjadi suara menggunakan Text-to-Speech (pada Mode Gestur), atau
  - Divisualisasikan dalam bentuk huruf BISINDO (pada Mode Suara).
- f. Hasil akhir dapat dikirim dan disimpan melalui modul integrasi berbasis webhook untuk keperluan dokumentasi.

Perancangan alur data ini bertujuan untuk memastikan sistem bekerja secara terstruktur dan mendukung komunikasi dua arah secara konsisten.

### III.3.4 Pertimbangan Desain Sistem

Beberapa pertimbangan dalam perancangan sistem meliputi:

- a. Konsistensi Input: Model Ukuran citra distandarisasi menjadi 128×128 piksel untuk memastikan kesesuaian dengan konfigurasi pelatihan model.
- b. Akurasi Klasifikasi: Penggunaan MobileNetV2 dengan transfer learning dan fine-tuning bertujuan meningkatkan kemampuan generalisasi terhadap variasi gestur tangan.
- c. Efisiensi Komputasi: MobileNetV2 dipilih karena ringan dan cocok untuk implementasi berbasis web.
- d. Aksesibilitas: Sistem dirancang berbasis web agar dapat digunakan tanpa instalasi tambahan.

### III.4 Pengumpulan Data

Pengumpulan data dilakukan untuk membangun dataset citra gestur huruf Bahasa Isyarat Indonesia (BISINDO) yang digunakan dalam proses pelatihan model klasifikasi. Dataset disusun secara mandiri dengan tujuan memperoleh data yang sesuai dengan kebutuhan sistem yang dikembangkan.

#### III.4.1 Jenis Data

Data yang digunakan dalam penelitian ini berupa data citra (image dataset) yang merepresentasikan gestur tangan untuk huruf alfabet A sampai Z. Setiap kelas merepresentasikan satu huruf BISINDO sehingga total terdapat 26 kelas klasifikasi.

Data bersifat supervised learning, di mana setiap citra telah diberi label sesuai huruf yang direpresentasikan.

#### III.4.2 Proses Akuisisi Data

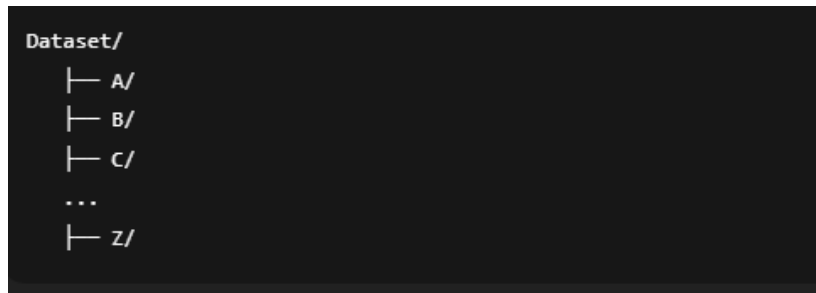
Proses pengambilan data dilakukan menggunakan kamera (webcam) untuk menangkap gestur tangan dalam berbagai posisi dan kondisi pencahayaan. Pengambilan gambar dilakukan secara bertahap untuk setiap huruf dengan memperhatikan variasi berikut:

- a. Variasi posisi tangan
- b. Variasi jarak tangan terhadap kamera
- c. Variasi sudut kemiringan tangan
- d. Variasi pencahayaan ruangan

Variasi ini bertujuan untuk meningkatkan keberagaman dataset sehingga model dapat belajar mengenali pola yang lebih general dan tidak hanya terbatas pada satu kondisi tertentu.

#### III.4.3 Struktur Dataset

Dataset disusun dalam struktur folder terpisah berdasarkan kelas huruf, dengan format sebagai berikut:



*Gambar III. 3 Struktur Dataset*

Struktur ini mempermudah proses pemanggilan data menggunakan generator pada tahap pelatihan model. Dataset disimpan pada Google Drive dan di-mount ke lingkungan Google Colab untuk mempermudah proses pelatihan.

#### **III.4.4 Pembagian Dataset**

Dataset dibagi menjadi:

- a. Data Training
- b. Data Validation

Pembagian ini dilakukan untuk memastikan model tidak hanya belajar dari data pelatihan, tetapi juga diuji pada data yang tidak digunakan saat training untuk mengukur kemampuan generalisasi.

Proporsi pembagian data dilakukan dengan pendekatan umum dalam pelatihan model klasifikasi citra.

#### **III.4.5 Pertimbangan Penggunaan Dataset Mandiri**

Penggunaan dataset mandiri dilakukan karena keterbatasan ketersediaan dataset publik khusus huruf BISINDO dengan format yang sesuai kebutuhan sistem. Dengan membangun dataset sendiri, penelitian ini dapat memastikan kesesuaian data dengan skenario implementasi sistem real-time yang dikembangkan.

Namun demikian, penggunaan dataset mandiri juga memiliki keterbatasan, seperti variasi subjek yang terbatas dan potensi bias terhadap kondisi tertentu. Hal ini menjadi salah satu pertimbangan dalam evaluasi performa model.

#### **III.5 Preprocessing Data Untuk Pelatihan**

Preprocessing data dilakukan untuk memastikan citra gestur yang dikumpulkan memiliki format dan karakteristik yang sesuai dengan kebutuhan model deep learning. Tahap ini

bertujuan meningkatkan kualitas data serta membantu model dalam proses pembelajaran fitur yang relevan.

### **III.5.1 Penyesuaian Ukuran Citra**

Seluruh citra pada dataset disesuaikan dengan ukuran input model klasifikasi. Penyesuaian ukuran dilakukan agar seluruh data memiliki dimensi yang seragam, sehingga dapat diproses secara optimal oleh arsitektur Convolutional Neural Network (CNN).

Ukuran citra disesuaikan dengan standar input arsitektur model pretrained yang digunakan dalam penelitian ini. Penyeragaman dimensi ini penting untuk menjaga konsistensi struktur tensor pada saat proses pelatihan.

### **III.5.2 Normalisasi Nilai Piksel**

Nilai piksel citra dinormalisasi untuk mengubah rentang intensitas warna menjadi skala yang lebih kecil. Normalisasi dilakukan untuk mempercepat proses konvergensi selama pelatihan model dan menjaga stabilitas perhitungan gradien.

Tanpa normalisasi, perbedaan skala nilai piksel dapat menyebabkan proses pelatihan menjadi kurang stabil atau membutuhkan waktu lebih lama untuk mencapai konvergensi.

### **III.5.3 Augmentasi Data**

Untuk meningkatkan kemampuan generalisasi model, dilakukan teknik augmentasi data pada dataset pelatihan. Augmentasi dilakukan dengan menghasilkan variasi citra baru dari data asli melalui transformasi tertentu, seperti:

- a. Rotasi ringan
- b. Pergeseran posisi (shift)
- c. Perubahan skala (zoom)
- d. Perubahan orientasi

Teknik ini bertujuan mensimulasikan variasi kondisi nyata, seperti perubahan sudut tangan dan posisi terhadap kamera, sehingga model tidak hanya belajar dari pola yang identik.

Augmentasi hanya diterapkan pada data pelatihan dan tidak pada data validasi, guna menjaga objektivitas proses evaluasi model.

### III.5.4 Encoding Label

Pada tahap ini, label huruf A–Z dikonversi ke bentuk numerik agar dapat diproses oleh model klasifikasi seperti Convolutional Neural Network (CNN). Model tidak dapat membaca data dalam bentuk teks, sehingga setiap huruf diberikan indeks angka (misalnya A=0, B=1, ..., Z=25). Proses ini disebut label encoding.

Selanjutnya, untuk bagian label numerik diubah menjadi representasi kategorikal menggunakan one-hot encoding. Dalam metode ini, setiap huruf direpresentasikan sebagai vektor biner sepanjang 26 kelas. Contohnya, huruf A (0) menjadi [1,0,0,...,0], sedangkan huruf C (2) menjadi [0,0,1,...,0].

Penggunaan one-hot encoding bertujuan untuk:

- ✓ Menghindari kesalahan interpretasi urutan numerik antar huruf.
- ✓ Mendukung fungsi aktivasi softmax pada layer output.
- ✓ Menyesuaikan dengan loss function seperti categorical crossentropy dalam klasifikasi multi-kelas.

Setelah model menghasilkan prediksi berupa vektor probabilitas, indeks dengan nilai tertinggi diambil menggunakan *argmax()*, kemudian dipetakan kembali ke huruf sesuai mapping awal.

Dengan demikian, encoding label berperan penting dalam memastikan model dapat mempelajari kelas huruf secara akurat dan mengembalikan hasil prediksi ke bentuk alfabet yang sesuai.

### III.5.5 Pembagian Data Training dan Validation

Dataset dibagi menjadi data pelatihan dan data validasi. Data pelatihan digunakan untuk memperbarui bobot model, sedangkan data validasi digunakan untuk memantau performa model selama proses pelatihan.

Pembagian ini dilakukan untuk mengurangi risiko overfitting serta memastikan model memiliki kemampuan generalisasi terhadap data yang tidak dilatih secara langsung.

## III.6 Metode Pelatihan Model

Model klasifikasi huruf BISINDO dalam penelitian ini dibangun menggunakan arsitektur Convolutional Neural Network (CNN) berbasis MobileNetV2 yang memanfaatkan bobot awal

(pretrained weights) dari dataset ImageNet melalui pendekatan Transfer Learning dan Fine-Tuning bertahap.

MobileNetV2 dipilih karena merupakan arsitektur CNN ringan (lightweight convolutional neural network) yang dirancang untuk efisiensi komputasi dan cocok untuk implementasi sistem penerjemah BISINDO. Dibandingkan dengan arsitektur CNN yang lebih kompleks seperti VGG atau ResNet, MobileNetV2 memiliki jumlah parameter yang lebih sedikit sehingga lebih efisien dijalankan pada lingkungan komputasi terbatas seperti Google Colab serta implementasi berbasis web menggunakan Gradio.

Arsitektur ini menggunakan mekanisme depthwise separable convolution dan inverted residual block yang memungkinkan pengurangan kompleksitas komputasi tanpa menurunkan performa secara signifikan. Dengan karakteristik tersebut, MobileNetV2 dinilai sesuai untuk mendukung sistem penerjemah bahasa isyarat yang memerlukan proses klasifikasi cepat dengan akurasi yang tetap optimal.

### III.6.1 Konfigurasi Arsitektur Model

Model dikembangkan menggunakan MobileNetV2 pretrained dengan konfigurasi sebagai berikut:

- a. Pretrained weights: ImageNet
- b. include\_top: False
- c. Input shape:  $128 \times 128 \times 3$
- d. Jumlah kelas output: 26 (huruf A-Z)

Lapisan fully connected tambahan ditambahkan pada bagian akhir model untuk menyesuaikan jumlah kelas keluaran. Layer output menggunakan fungsi aktivasi Softmax untuk menghasilkan probabilitas klasifikasi multi-kelas.

Pemilihan ukuran input  $128 \times 128$  dilakukan untuk meningkatkan efisiensi komputasi serta menyesuaikan kebutuhan sistem real-time tanpa mengurangi kemampuan model dalam mengekstraksi fitur visual gestur tangan.

### III.6.2 Tahap 1 Transfer Learning (Initial Training)

Transfer learning merupakan teknik pembelajaran mesin yang memanfaatkan model yang telah dilatih sebelumnya (pretrained model) untuk menyelesaikan tugas baru yang berbeda namun masih berada dalam domain yang serupa, yaitu klasifikasi citra.

Dalam penelitian ini, model pretrained digunakan sebagai feature extractor, di mana lapisan awal model berfungsi untuk mengekstraksi fitur umum seperti tepi, tekstur, dan bentuk. Lapisan akhir kemudian dimodifikasi untuk menyesuaikan jumlah kelas keluaran menjadi 26 kelas huruf BISINDO.

Pada tahap awal pelatihan:

- a. Seluruh layer MobileNetV2 dibekukan (freeze).
- b. Hanya layer klasifikasi tambahan yang dilatih.
- c. Fungsi loss yang digunakan adalah categorical crossentropy.
- d. Optimizer yang digunakan adalah Adam.
- e. Data yang digunakan merupakan citra hasil preprocessing berukuran  $128 \times 128$  piksel.

Tahap ini bertujuan untuk menyesuaikan layer akhir terhadap dataset huruf BISINDO tanpa mengubah bobot fitur dasar yang telah dipelajari dari dataset ImageNet. Model hasil tahap pertama disimpan sebagai bisindo.h5.

Proses pelatihan pada tahap awal dilakukan selama 15 epoch dengan batch size sebesar 32. Optimizer Adam digunakan dengan pengaturan learning rate standar pada tahap transfer learning. Konfigurasi ini dipilih untuk menjaga keseimbangan antara stabilitas pelatihan dan efisiensi waktu komputasi.

### III.6.3 Tahap 2 Fine-Tuning Bertahap

Setelah tahap awal selesai, dilakukan proses fine-tuning untuk meningkatkan performa klasifikasi.

Fine-tuning dilakukan dengan cara:

- a. Membuka (unfreeze) sebagian layer akhir MobileNetV2.
- b. Menggunakan learning rate yang lebih kecil untuk menjaga stabilitas pelatihan.
- c. Melatih ulang model agar fitur tingkat tinggi lebih adaptif terhadap karakteristik gestur huruf BISINDO.

Pendekatan bertahap ini mengurangi risiko perubahan bobot secara drastis yang dapat menyebabkan ketidakstabilan model.

Pada tahap fine-tuning, proses pelatihan dilakukan selama 3 epoch dengan batch size sebesar 32 serta menggunakan learning rate yang lebih kecil dibandingkan tahap

transfer learning. Penurunan learning rate bertujuan untuk menjaga stabilitas pembaruan bobot pada layer yang telah dibuka sehingga model dapat beradaptasi tanpa menyebabkan perubahan parameter yang terlalu drastis.

Model akhir hasil fine-tuning disimpan sebagai model\_finetune.h5 dan digunakan sebagai model utama dalam sistem penerjemahan dua arah.

### III.6.4 Pertimbangan Metodologis

Pemilihan transfer learning dan fine-tuning didasarkan pada beberapa pertimbangan:

- a. Dataset berskala terbatas sehingga pelatihan dari nol (training from scratch) berisiko menghasilkan overfitting.
- b. Model pretrained telah mempelajari fitur visual umum yang relevan untuk klasifikasi citra tangan.
- c. Pendekatan bertahap meningkatkan stabilitas konvergensi.
- d. Efisiensi waktu pelatihan lebih baik dibanding membangun arsitektur dari awal.

Dengan pendekatan ini, diharapkan model mampu mengenali pola gestur huruf BISINDO secara lebih akurat dan stabil.

### III.7 Metode Evaluasi

Evaluasi dilakukan untuk mengukur performa model klasifikasi huruf BISINDO serta menilai kesiapan sistem sebelum diimplementasikan dalam lingkungan secara langsung. Proses evaluasi dilakukan pada dua aspek utama, yaitu evaluasi model klasifikasi dan evaluasi sistem secara keseluruhan.

#### III.7.1 Evaluasi Model Klasifikasi

Evaluasi model dilakukan menggunakan data validasi yang tidak digunakan selama proses pembaruan bobot (training). Tujuannya adalah untuk mengukur kemampuan generalisasi model terhadap data yang belum pernah dilihat sebelumnya.

Metrik evaluasi yang digunakan meliputi:

- a. Accuracy Mengukur proporsi prediksi yang benar terhadap seluruh data pengujian. Metrik ini memberikan gambaran umum performa model dalam klasifikasi multi-kelas.

**Rumus:**

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

Untuk multi-kelas secara umum:

$$\text{Accuracy} = \frac{\text{Jumlah Seluruh Prediksi Benar}}{\text{Total Seluruh Data}}$$

- b. Precision Mengukur ketepatan model dalam memprediksi suatu kelas tertentu dibandingkan dengan seluruh prediksi pada kelas tersebut.

**Rumus:**

$$\text{Precision} = \frac{TP}{TP + FP}$$

**Keterangan:**

- ✓ TP (True Positive): data kelas tersebut yang diprediksi benar
- ✓ FP (False Positive): data kelas lain yang salah diprediksi sebagai kelas tersebut

- c. Recall Mengukur kemampuan model dalam mendeteksi seluruh data yang benar-benar termasuk dalam suatu kelas.

**Rumus:**

$$\text{Recall} = \frac{TP}{TP + FN}$$

**Keterangan:**

- ✓ FN (False Negative): data kelas tersebut yang salah diprediksi sebagai kelas lain

- d. F1-Score Merupakan rata-rata harmonis antara precision dan recall, yang memberikan keseimbangan antara kedua metrik tersebut.

**Rumus:**

$$F1 - \text{Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

- e. Confusion Matrix Digunakan untuk menganalisis distribusi kesalahan klasifikasi antar kelas huruf. Confusion matrix membantu mengidentifikasi huruf yang sering tertukar oleh model.

Confusion Matrix tidak memiliki satu rumus tunggal, tetapi berbentuk tabel yang menunjukkan jumlah:

- ✓ True Positive (TP)
- ✓ False Positive (FP)
- ✓ False Negative (FN)
- ✓ True Negative (TN)

Pada klasifikasi huruf A-Z (26 kelas), confusion matrix berbentuk matriks  $26 \times 26$ , di mana:

- ✓ Baris = kelas sebenarnya
- ✓ Kolom = kelas prediksi

Elemen diagonal menunjukkan jumlah prediksi benar, sedangkan elemen di luar diagonal menunjukkan kesalahan klasifikasi (huruf yang tertukar).

Penggunaan beberapa metrik ini bertujuan untuk memperoleh gambaran performa model secara komprehensif, tidak hanya berdasarkan accuracy semata.

### III.7.2 Evaluasi Sistem Secara Fungsional

Selain evaluasi model klasifikasi, sistem juga dirancang untuk diuji secara fungsional sebelum digunakan dalam implementasi secara langsung pengujiannya.

Pengujian fungsional mencakup:

- a. Kemampuan sistem menerima input gestur gambar dan Confidence per huruf
- b. Kemampuan sistem menerima input suara diubah menjadi kata dan sistem bisa memvisualisasikan perhuruf
- c. Integrasi antar modul dalam arsitektur sistem dengan n8n

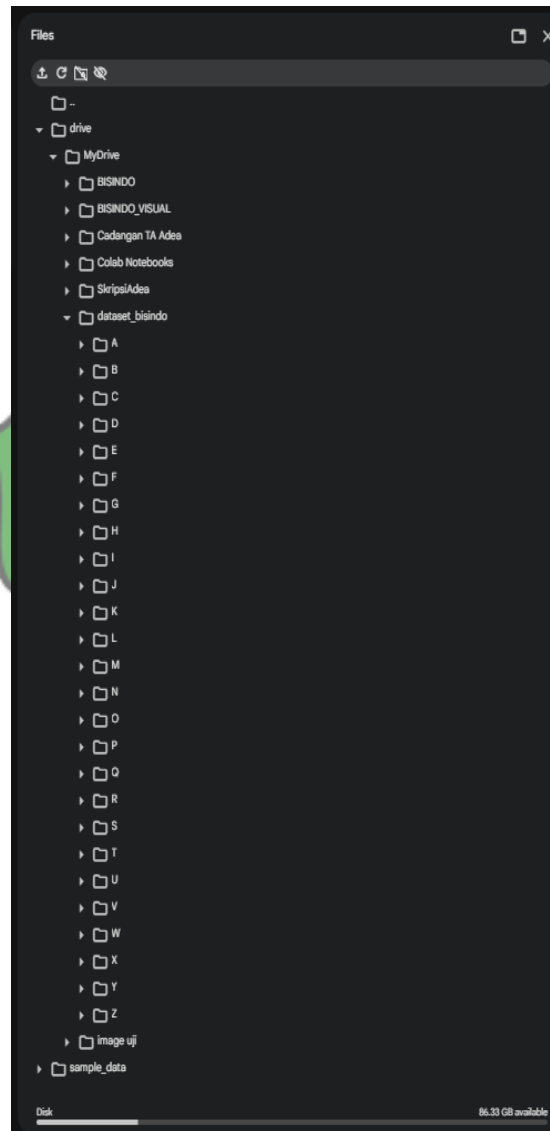
Evaluasi ini bertujuan memastikan bahwa seluruh modul yang dirancang pada tahap perancangan dapat bekerja secara terintegrasi.

## BAB IV

### IMPLEMENTASI DAN HASIL PENELITIAN

#### IV.1 Implementasi Dataset dan Preprocessing

Pada tahap ini dilakukan implementasi persiapan dataset sebelum proses pelatihan model. Dataset huruf BISINDO A-Z yang telah dikumpulkan disusun dalam bentuk folder per kelas, kemudian diakses melalui Google Colab untuk proses pelatihan.



*Gambar IV. 1 Folder Dataset di Google Drive*

Dataset dibagi menjadi data training dan validation menggunakan teknik split otomatis melalui parameter `validation_split` pada `ImageDataGenerator`. Selain itu, diterapkan teknik augmentasi data untuk meningkatkan variasi citra dan kemampuan generalisasi model.

Implementasi augmentasi dan pembagian data dilakukan menggunakan kode berikut:

```
from tensorflow.keras.preprocessing.image import ImageDataGenerator

# ===== Ukuran gambar dan batch =====

IMG_SIZE = (128, 128)
BATCH_SIZE = 32

# ===== Augmentasi & rescale untuk training =====
train_datagen = ImageDataGenerator(
    rescale=1./255,
    rotation_range=15,
    width_shift_range=0.1,
    height_shift_range=0.1,
    zoom_range=0.1,
    brightness_range=[0.8, 1.2],
    horizontal_flip=True,
    fill_mode='nearest',
    validation_split=0.2
)

# ===== Generator untuk training =====
train_generator = train_datagen.flow_from_directory(
    DATASET_DIR,
    target_size=IMG_SIZE,
    batch_size=BATCH_SIZE,
    class_mode='categorical',
    subset='training',
    shuffle=True
)

# ===== Generator untuk validation =====
val_generator = train_datagen.flow_from_directory(
    DATASET_DIR,
    target_size=IMG_SIZE,
    batch_size=BATCH_SIZE,
    class_mode='categorical',
    subset='validation',
    shuffle=False
)
```

Berdasarkan hasil eksekusi `flow_from_directory`, sistem secara otomatis membaca struktur folder dataset dan menghasilkan pembagian data sebesar 80% untuk training dan 20% untuk validation.

Ukuran citra ditetapkan sebesar  $128 \times 128$  piksel dengan tiga kanal warna (RGB). Pemilihan ukuran ini bertujuan menjaga keseimbangan antara detail visual gestur tangan dan efisiensi komputasi selama proses pelatihan.

```
for file in os.listdir(label_path):
    if file.lower().endswith(('.jpg', '.png', '.jpeg')):
        class_count[label] += 1

for k in sorted(class_count):
    print(k, class_count[k])

print("\nTotal data:", sum(class_count.values()))

*** A 165
    B 169
    C 169
    D 170
    E 169
    F 173
    G 173
    H 173
    I 173
    J 169
    K 173
    L 172
    M 173
    N 173
    O 173
    P 172
    Q 173
    R 171
    S 173
    T 173
    U 173
    V 172
    W 170
    X 174
    Y 174
    Z 174

Total data: 4466
```

**Gambar IV. 2 Banyak Gambar Per Label Huruf BISINDO**

```
# ===== Generator untuk validation =====
val_generator = train_datagen.flow_from_directory(
    DATASET_DIR,
    target_size=IMG_SIZE,
    batch_size=BATCH_SIZE,
    class_mode='categorical',
    subset='validation',          # ambil 20% untuk validation
    shuffle=False
)

Found 3587 images belonging to 26 classes.
Found 879 images belonging to 26 classes.
```

**Gambar IV. 3 Output Pembagian Data Training dan Validation**

Augmentasi dilakukan melalui rotasi, pergeseran, zoom, variasi kecerahan, serta horizontal flip. Penerapan augmentasi ini bertujuan meningkatkan kemampuan model dalam mengenali variasi posisi tangan, pencahayaan, dan orientasi gestur.



*Gambar IV. 4 Hasil Beberapa Gambar Augmentasi Data*

## IV2 Implementasi Pelatihan Model

### IV21 Implementasi Transfer Learning

Setelah proses persiapan dataset dan augmentasi selesai dilakukan, tahap berikutnya adalah implementasi model klasifikasi menggunakan pendekatan transfer learning dengan arsitektur MobileNetV2 pretrained.

Model dibangun menggunakan bobot awal dari ImageNet dengan konfigurasi `include_top=False`, sehingga hanya bagian feature extractor yang digunakan. Layer klasifikasi tambahan kemudian ditambahkan untuk menyesuaikan jumlah kelas keluaran sebanyak 26 huruf BISINDO (A-Z).

Implementasi pembuatan model dilakukan menggunakan kode berikut:

```
from tensorflow.keras.applications import MobileNetV2
from tensorflow.keras.layers import Dense, GlobalAveragePooling2D
from tensorflow.keras.models import Model
```

```
# Load pretrained MobileNetV2
```

```
base_model = MobileNetV2(
    weights='imagenet',
    include_top=False,
    input_shape=(128, 128, 3)
)
```

```
# Freeze seluruh layer backbone
```

```
base_model.trainable = False
```

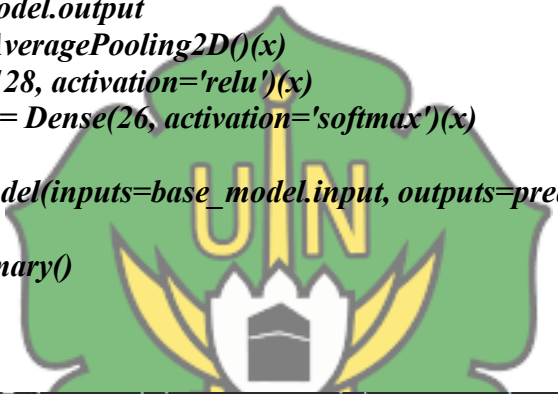
```
# Tambahkan layer klasifikasi
```

```
x = base_model.output
x = GlobalAveragePooling2D()(x)
x = Dense(128, activation='relu')(x)
predictions = Dense(26, activation='softmax')(x)
```

```
model = Model(inputs=base_model.input, outputs=predictions)
```

```
model.summary()
```

**OUTPUT:**



...	(ReLU)			
block_16_project (Conv2D)	(None, 4, 4, 320)	307,200	block_16_depthwi...	
block_16_project_BN (BatchNormalizatio...	(None, 4, 4, 320)	1,280	block_16_project...	
Conv_1 (Conv2D)	(None, 4, 4, 1280)	409,600	block_16_project...	
Conv_1_bn (BatchNormalizatio...	(None, 4, 4, 1280)	5,120	Conv_1[0][0]	
out_relu (ReLU)	(None, 4, 4, 1280)	0	Conv_1_bn[0][0]	
global_average_poo... (GlobalAveragePool...	(None, 1280)	0	out_relu[0][0]	
dense_6 (Dense)	(None, 128)	163,968	global_average_p...	
dropout_3 (Dropout)	(None, 128)	0	dense_6[0][0]	
dense_7 (Dense)	(None, 26)	3,354	dropout_3[0][0]	
Total params: 2,425,306 (9.25 MB)				
Trainable params: 1,785,882 (6.81 MB)				
Non-trainable params: 639,424 (2.44 MB)				

Gambar IV. 5 Ringkasan Arsitektur MobileNetV2

Berdasarkan hasil `model.summary()`, terlihat bahwa seluruh layer backbone MobileNetV2 berada dalam kondisi non-trainable, sedangkan layer klasifikasi tambahan bersifat trainable. Hal ini menunjukkan bahwa pada tahap transfer learning, model hanya memperbarui bobot pada layer klasifikasi tanpa mengubah bobot fitur dasar yang telah dipelajari dari dataset ImageNet.

Model kemudian dikompilasi dengan konfigurasi sebagai berikut:

```
model.compile(  
    optimizer='adam',  
    loss='categorical_crossentropy',  
    metrics=['accuracy']  
)
```

Pelatihan model dilakukan selama 15 epoch dengan batch size 32, menggunakan dataset yang telah melalui proses augmentasi dan pembagian data sebelumnya.

```
history = model.fit(  
    train_generator,  
    validation_data=val_generator,  
    epochs=15  
)  
OUTPUT:
```



```

Epoch 9: val_loss improved from 2.10580 to 1.92675, saving model to /content/drive/MyDrive/BISINDO/model_terbaik.h5
WARNING:absl:You are saving your model as an HDF5 file via `model.save()` or `keras.saving.save_model(model)`. This file format is c
113/113 ----- 114s 1s/step - accuracy: 0.7157 - loss: 0.9097 - val_accuracy: 0.4721 - val_loss: 1.9267
Epoch 10/15
113/113 ----- 0s 840ms/step - accuracy: 0.7468 - loss: 0.8149
Epoch 10: val_loss did not improve from 1.92675
113/113 ----- 112s 990ms/step - accuracy: 0.7467 - loss: 0.8151 - val_accuracy: 0.4539 - val_loss: 1.9911
Epoch 11/15
113/113 ----- 0s 844ms/step - accuracy: 0.7653 - loss: 0.7350
Epoch 11: val_loss did not improve from 1.92675
113/113 ----- 113s 1s/step - accuracy: 0.7653 - loss: 0.7351 - val_accuracy: 0.4710 - val_loss: 1.9429
Epoch 12/15
113/113 ----- 0s 828ms/step - accuracy: 0.7684 - loss: 0.6997
Epoch 12: val_loss did not improve from 1.92675
113/113 ----- 110s 977ms/step - accuracy: 0.7684 - loss: 0.6997 - val_accuracy: 0.4164 - val_loss: 2.3903
Epoch 13/15
113/113 ----- 0s 849ms/step - accuracy: 0.8023 - loss: 0.6258
Epoch 13: val_loss did not improve from 1.92675
113/113 ----- 113s 998ms/step - accuracy: 0.8023 - loss: 0.6256 - val_accuracy: 0.4721 - val_loss: 2.0652
Epoch 14/15
113/113 ----- 0s 826ms/step - accuracy: 0.8111 - loss: 0.5538
Epoch 14: val_loss did not improve from 1.92675
113/113 ----- 110s 974ms/step - accuracy: 0.8110 - loss: 0.5541 - val_accuracy: 0.4016 - val_loss: 2.5511
Epoch 14: early stopping
Restoring model weights from the end of the best epoch: 9.
WARNING:absl:You are saving your model as an HDF5 file via `model.save()` or `keras.saving.save_model(model)`. This file format is c
✓ Model terbaik tersimpan di: /content/drive/MyDrive/BISINDO/model_terbaik.h5
✓ Model terakhir tersimpan di: /content/drive/MyDrive/BISINDO/model_terakhir.h5

```

*Gambar IV. 6 Proses Pelatihan Transfer Learning*

Setelah proses pelatihan selesai, model disimpan dalam format .h5 untuk digunakan pada tahap fine-tuning:

`model.save('bisindo.h5')`

## IV22 Implementasi Fine-Tuning

Setelah model hasil transfer learning disimpan sebagai bisindo.h5, tahap berikutnya adalah proses fine-tuning untuk meningkatkan kemampuan model dalam mengenali karakteristik gestur huruf BISINDO secara lebih spesifik.

Pada tahap ini, sebagian layer akhir dari MobileNetV2 dibuka (unfreeze) agar bobotnya dapat diperbarui kembali. Berbeda dengan tahap sebelumnya, fine-tuning dilakukan dengan learning rate yang lebih kecil untuk menjaga stabilitas pembaruan bobot.

Implementasi fine-tuning dilakukan dengan langkah berikut:

```

# Load model hasil transfer learning
from tensorflow.keras.models import load_model
from tensorflow.keras.optimizers import Adam

```

```

model = load_model('bisindo.h5')

# Membuka sebagian layer akhir
for layer in model.layers[-20:]: # contoh membuka 20 layer terakhir
    layer.trainable = True

# Compile ulang dengan learning rate lebih kecil
model.compile(
    optimizer=Adam(learning_rate=1e-5),
    loss='categorical_crossentropy',
    metrics=['accuracy']
)

model.summary()

```

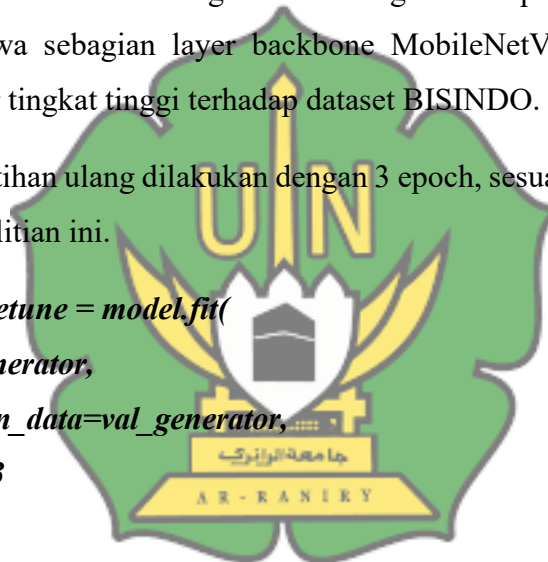
Berdasarkan hasil `model.summary()` setelah proses unfreeze, terlihat bahwa jumlah parameter trainable meningkat dibandingkan tahap transfer learning. Hal ini menunjukkan bahwa sebagian layer backbone MobileNetV2 kini ikut dilatih untuk menyesuaikan fitur tingkat tinggi terhadap dataset BISINDO.

Proses pelatihan ulang dilakukan dengan 3 epoch, sesuai dengan konfigurasi pada implementasi penelitian ini.

```

history_finetune = model.fit(
    train_generator,
    validation_data=val_generator,
    epochs=3
)

```



## OUTPUT:

```
print(f"✅ Fine-tune selesai. Model fine-tune tersimpan di: {FINE_TUNE_MODEL_PATH}")

*** WARNING:absl:Compiled the loaded model, but the compiled metrics have yet to be built. `model.compile_metrics` will be empty until you t
✅ Model terbaik berhasil di-load.
Epoch 1/3
113/113 ————— 0s 814ms/step - accuracy: 0.7420 - loss: 0.8317
Epoch 1: val_loss improved from inf to 1.69433, saving model to /content/drive/MyDrive/BISINDO/model_finetune.h5
WARNING:absl:You are saving your model as an HDF5 file via `model.save()` or `keras.saving.save_model(model)`. This file format is consid
113/113 ————— 123s 987ms/step - accuracy: 0.7420 - loss: 0.8317 - val_accuracy: 0.5006 - val_loss: 1.6943
Epoch 2/3
113/113 ————— 0s 828ms/step - accuracy: 0.7520 - loss: 0.7589
Epoch 2: val_loss improved from 1.69433 to 1.59704, saving model to /content/drive/MyDrive/BISINDO/model_finetune.h5
WARNING:absl:You are saving your model as an HDF5 file via `model.save()` or `keras.saving.save_model(model)`. This file format is consid
113/113 ————— 114s 999ms/step - accuracy: 0.7521 - loss: 0.7589 - val_accuracy: 0.5199 - val_loss: 1.5970
Epoch 3/3
113/113 ————— 0s 825ms/step - accuracy: 0.7490 - loss: 0.7615
Epoch 3: val_loss improved from 1.59704 to 1.53032, saving model to /content/drive/MyDrive/BISINDO/model_finetune.h5
WARNING:absl:You are saving your model as an HDF5 file via `model.save()` or `keras.saving.save_model(model)`. This file format is consid
113/113 ————— 112s 992ms/step - accuracy: 0.7490 - loss: 0.7615 - val_accuracy: 0.5472 - val_loss: 1.5303
Restoring model weights from the end of the best epoch: 3.
✅ Fine-tune selesai. Model fine-tune tersimpan di: /content/drive/MyDrive/BISINDO/model_finetune.h5
```

**Gambar IV. 7 Proses Pelatihan Fine-Tuning**

Setelah proses fine-tuning selesai, model disimpan sebagai model final:

**`model.save('model_finetune.h5')`**

Model `model_finetune.h5` inilah yang kemudian digunakan pada sistem penerjemahan dua arah berbasis web.

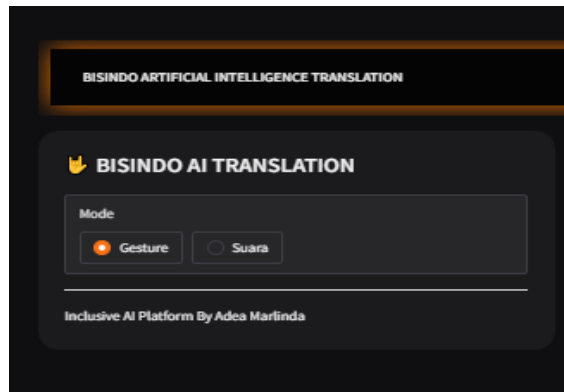
### IV3 Implementasi Sistem Penerjemahan Dua Arah

Setelah model `model_finetune.h5` diperoleh, tahap berikutnya adalah implementasi sistem penerjemahan dua arah berbasis web menggunakan Gradio sebagai antarmuka utama.

Berbeda dengan tahap perancangan pada BAB III, bagian ini menjelaskan realisasi sistem berdasarkan file implementasi yang dibuat pada Google Colab.

Sistem terdiri dari dua mode utama:

- Mode Gestur: Upload Gambar Gesture → Teks → TTS
- Mode Suara: Rekam/Upload Audio → Teks → Visual Huruf



*Gambar IV. 8 Pilihan Mode BISINDO*

### IV31 Implementasi Mode Gestur

Pada mode ini, pengguna mengunggah (upload) citra gestur huruf BISINDO melalui antarmuka web. Citra yang diunggah kemudian diproses menggunakan model `model_finetune.h5` untuk menghasilkan prediksi huruf.

Implementasi pemanggilan model dilakukan sebagai berikut:

```
import numpy as np
from tensorflow.keras.models import load_model
from tensorflow.keras.preprocessing import image

model = load_model("model_finetune.h5")

def predict_image(img):
    img = img.resize((128, 128))
    img_array = image.img_to_array(img)
    img_array = img_array / 255.0
    img_array = np.expand_dims(img_array, axis=0)

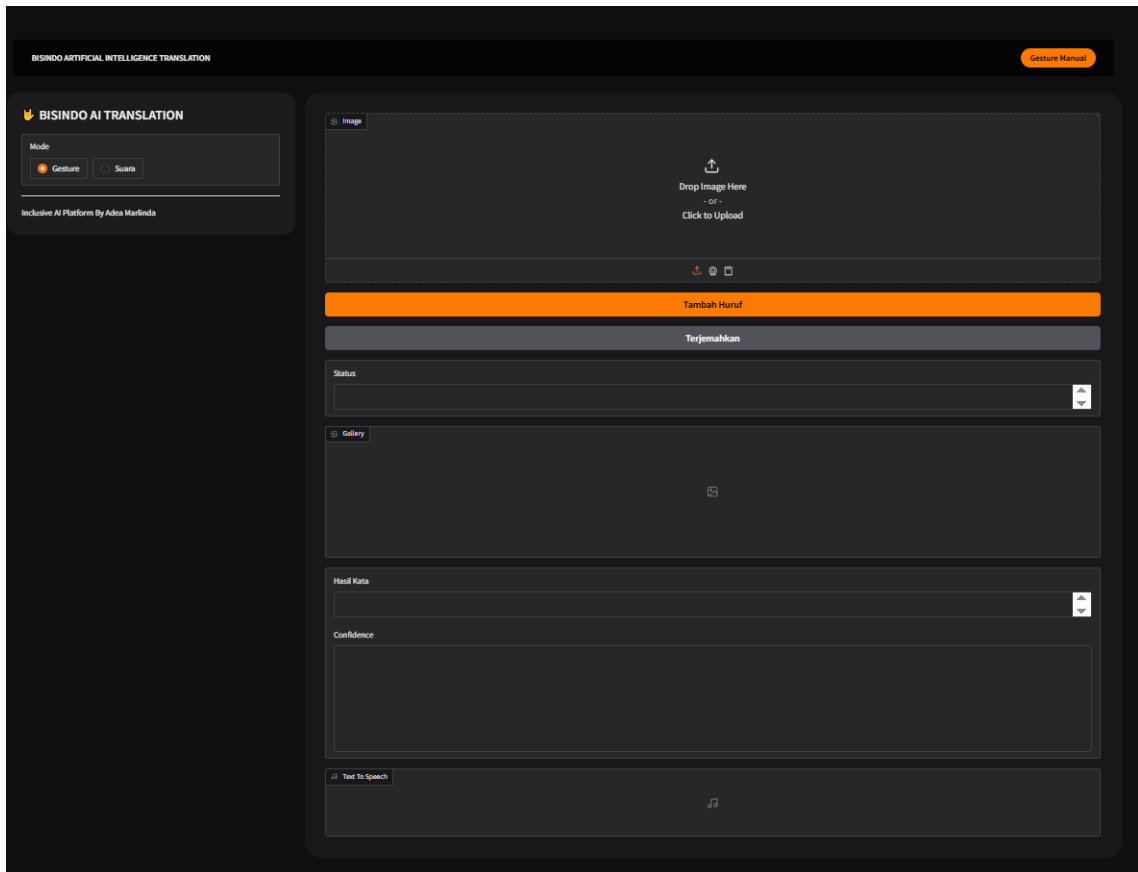
    prediction = model.predict(img_array)
    class_index = np.argmax(prediction)

    return class_index
```

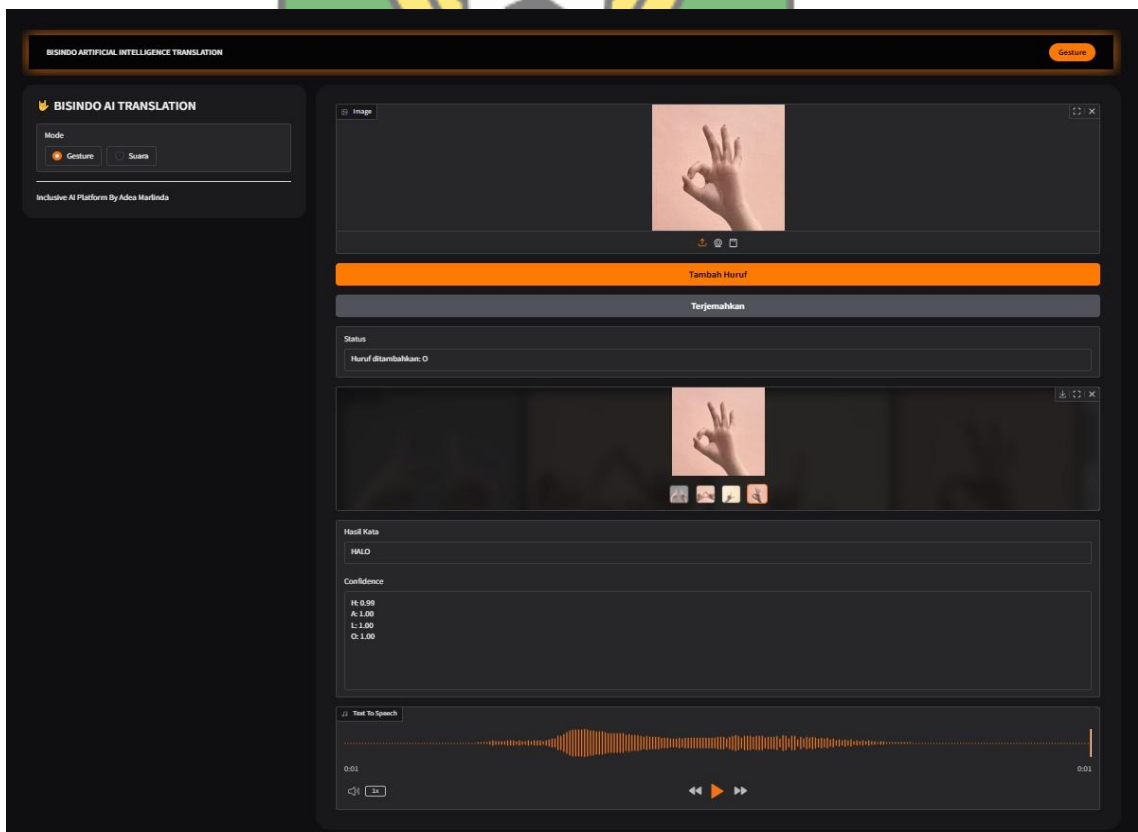
Proses yang terjadi pada fungsi di atas meliputi:

- Resize citra ke ukuran 128×128 piksel
- Normalisasi piksel (0-1)
- Ekspansi dimensi untuk batch input
- Prediksi menggunakan model
- Pengambilan indeks kelas dengan probabilitas tertinggi

**OUTPUT:**



*Gambar IV. 9 Antarmuka Upload Citra Gestur*



*Gambar IV. 10 Hasil Prediksi Huruf BISINDO*

### IV32 Implementasi Mode Suara

Pada mode ini, pengguna dapat merekam atau mengunggah file suara. Sistem mengubah suara menjadi teks menggunakan modul Speech-to-Text (STT). Teks tersebut kemudian divisualisasikan dalam bentuk huruf BISINDO.

#### *Implementasi STT:*

```
import speech_recognition as sr
```

```
def speech_to_text(audio_file):
```

```
    recognizer = sr.Recognizer()
```

```
    with sr.AudioFile(audio_file) as source:
```

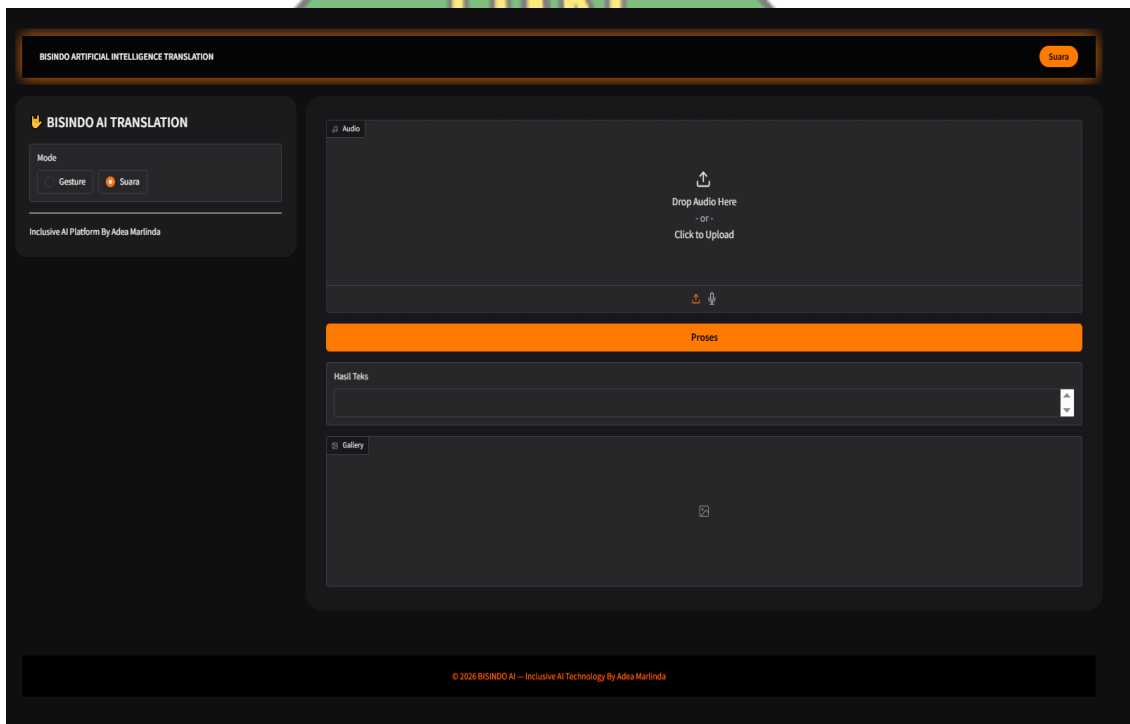
```
        audio = recognizer.record(source)
```

```
    text = recognizer.recognize_google(audio, language="id-ID")
```

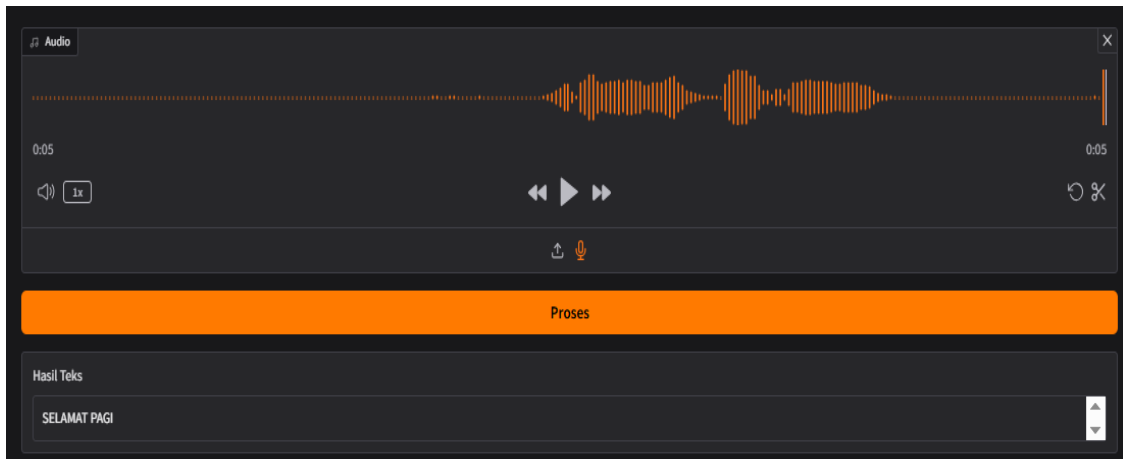
```
    return text
```

Setelah teks diperoleh, sistem menampilkan representasi huruf secara visual sesuai alfabet.

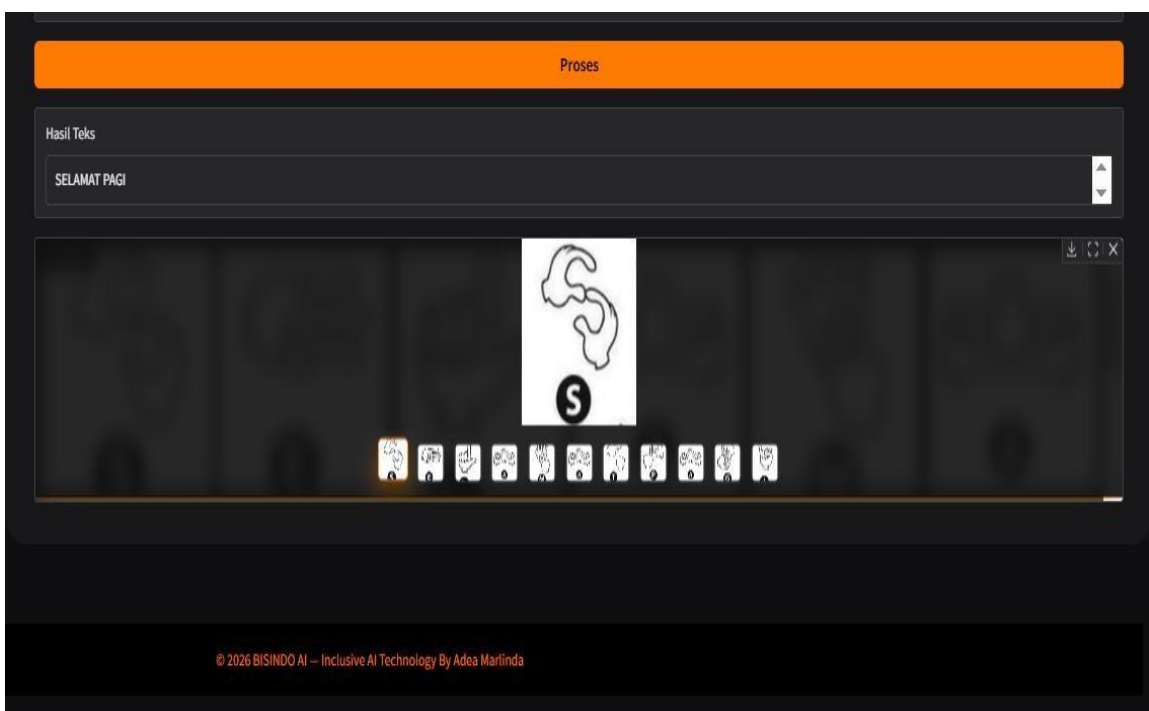
#### *OUTPUT:*



*Gambar IV. 11 Antarmuka Rekam/Upload Suara*



*Gambar IV. 12 Hasil Konversi Suara ke Teks*



*Gambar IV. 13 Visualisasi Huruf BISINDO*

### IV33 Integrasi Webhook n8n

Sistem juga diintegrasikan dengan webhook n8n untuk menyimpan hasil terjemahan serta mengirim notifikasi.

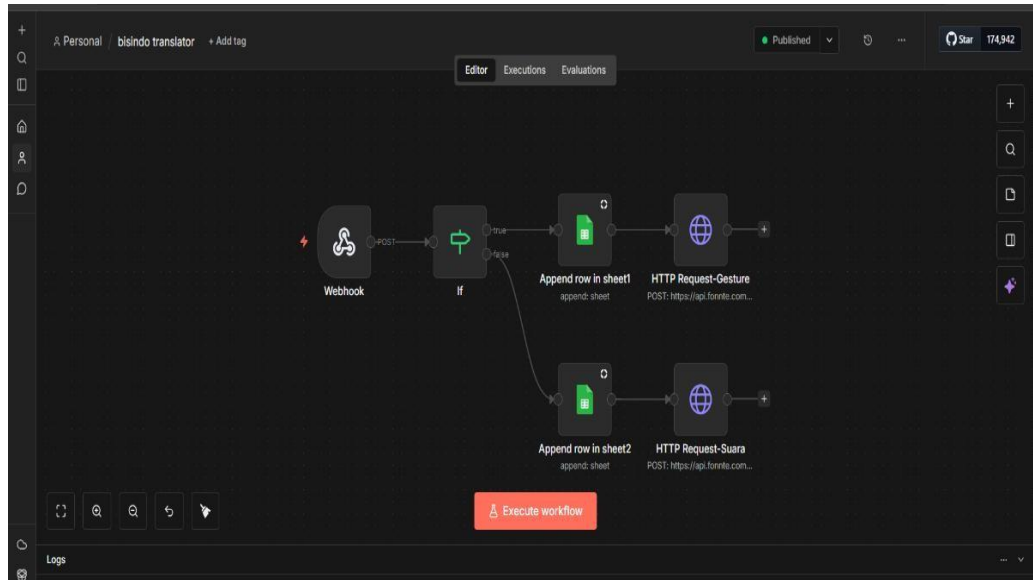
Implementasi pengiriman data dilakukan menggunakan HTTP request:

```
import requests
def send_to_webhook(text):
    url = "WEBHOOK_URL"
    data = {"hasil_terjemahan": text}
    requests.post(url, json=data)
```

Webhook kemudian:

- Menyimpan data ke Google Spreadsheet

**OUTPUT:**



**Gambar IV. 14 Workflow Integrasi Webhook n8n**

The screenshot shows a Google Spreadsheet with the following data:

Timestamp	Mode	Hasil Teks	Confidence	TTS
2026-02-17 20:34:48	Gesture	HALO	H: 1.00 A: 1.00 L: 1.00 O: 1.00	<a href="#">Play Audio TTS</a>
2026-02-18 18:40:09	Gesture	ARSYAH	A: 1.00 R: 0.99 S: 0.92 Y: 0.82	<a href="#">Play Audio TTS</a>
2026-02-18 18:50:57	Gesture	SAHAR	A: 1.00 H: 0.99 A: 1.00 R: 0.99	<a href="#">Play Audio TTS</a>

**Gambar IV. 15 Hasil Penyimpanan ke Google Spreadsheet Gesture**

A	B	C	D	E
Timestamp	Mode	Voice	Teks	Visual
2026-02-17 20:35:35	Voice	<a href="#">Play Audio STT</a>	TERIMA KASIH	/content/drive/MyDrive/BISINDO_VISUAL/I.PNG /content/drive/MyDrive/BISINDO_VISUAL/E.PNG /content/drive/MyDrive/BISINDO_VISUAL/R.PNG /content/drive/MyDrive/BISINDO_VISUAL/J.PNG /content/drive/MyDrive/BISINDO_VISUAL/M.PNG /content/drive/MyDrive/BISINDO_VISUAL/A.PNG /content/drive/MyDrive/BISINDO_VISUAL/K.PNG /content/drive/MyDrive/BISINDO_VISUAL/JA.PNG /content/drive/MyDrive/BISINDO_VISUAL/S.PNG /content/drive/MyDrive/BISINDO_VISUAL/I.PNG /content/drive/MyDrive/BISINDO_VISUAL/H.PNG
2026-02-18 18:41:41	Voice	<a href="#">Play Audio STT</a>	BELAJAR	/content/drive/MyDrive/BISINDO_VISUAL/B.PNG /content/drive/MyDrive/BISINDO_VISUAL/E.PNG /content/drive/MyDrive/BISINDO_VISUAL/J.PNG /content/drive/MyDrive/BISINDO_VISUAL/A.PNG /content/drive/MyDrive/BISINDO_VISUAL/J.PNG /content/drive/MyDrive/BISINDO_VISUAL/J.PNG /content/drive/MyDrive/BISINDO_VISUAL/J.PNG

*Gambar IV. 16 Hasil Penyimpanan ke Google Spreadsheet Suara*

#### IV4 Evaluasi dan Pengujian Sistem

Evaluasi dilakukan untuk menilai performa model klasifikasi huruf BISINDO serta memastikan sistem penerjemahan dua arah dapat berjalan dengan baik sesuai dengan rancangan pada BAB III.

Evaluasi dilakukan pada dua aspek utama:

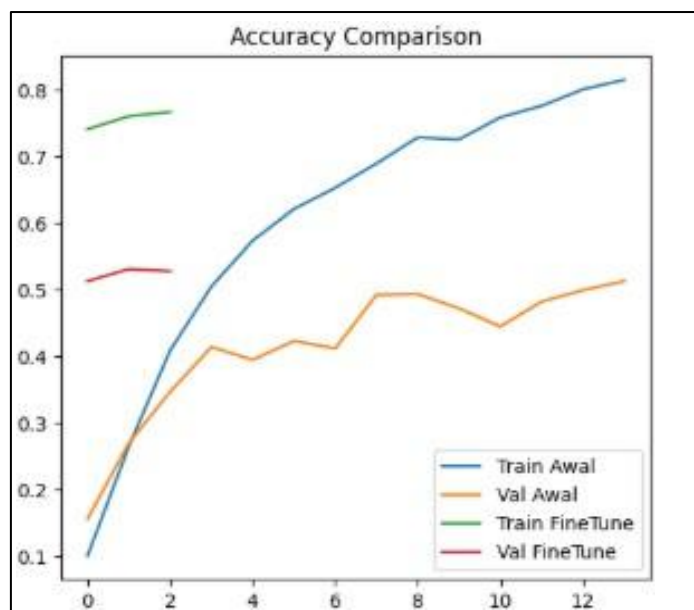
- Evaluasi performa model klasifikasi
- Pengujian fungsional sistem

##### IV4.1 Evaluasi Model Klasifikasi

Evaluasi model dilakukan menggunakan data validation yang tidak digunakan selama proses pembaruan bobot (training). Pengukuran dilakukan berdasarkan metrik yang telah dijelaskan pada BAB III.

###### a. Accuracy

Berdasarkan hasil pelatihan, model menunjukkan nilai accuracy training dan validation yang stabil. Grafik berikut menunjukkan perkembangan akurasi selama proses pelatihan:



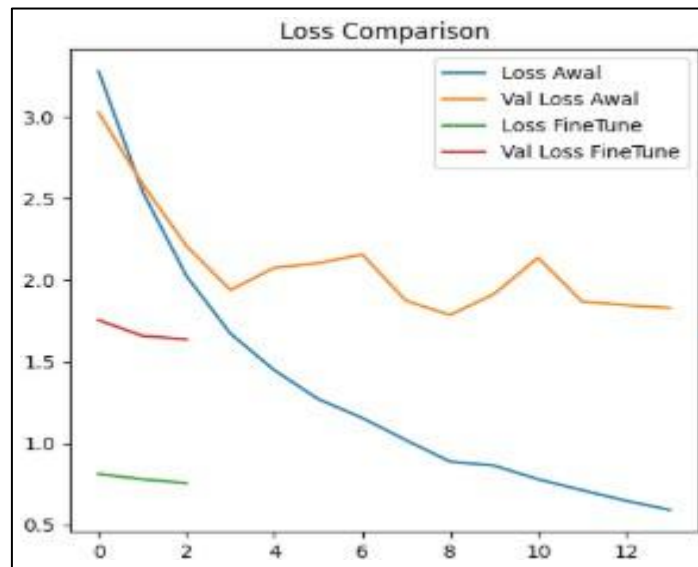
*Gambar IV. 17 Grafik Akurasi Model*

Berdasarkan Gambar IV.17 terlihat bahwa akurasi model mengalami peningkatan secara bertahap pada setiap epoch, baik pada data pelatihan maupun data validasi. Pola peningkatan yang konsisten menunjukkan bahwa model berhasil mempelajari fitur-fitur penting dari data gestur huruf A-Z secara progresif. Perbedaan antara training accuracy dan validation accuracy relatif kecil, sehingga dapat disimpulkan bahwa model tidak mengalami overfitting secara signifikan dan memiliki kemampuan generalisasi yang cukup baik.

Pada akhir proses pelatihan, model memperoleh akurasi sebesar **76%**. Nilai ini menunjukkan bahwa sebagian besar data uji berhasil diklasifikasikan dengan benar. Meskipun masih terdapat beberapa kelas huruf yang mengalami kesalahan prediksi, secara keseluruhan performa model sudah cukup stabil untuk sistem klasifikasi multi-kelas dengan 26 kategori. Hasil ini juga didukung oleh analisis confusion matrix dan classification report yang menunjukkan dominasi nilai pada diagonal utama serta distribusi kesalahan yang masih dalam batas wajar.

Dengan demikian, model dapat dikatakan telah mampu mengenali pola gestur secara efektif, meskipun masih terdapat ruang untuk peningkatan melalui optimasi arsitektur, penambahan data, atau penyempurnaan proses pelatihan.

b. Loss



Gambar IV. 18 Grafik Loss Model

Grafik loss menunjukkan tren penurunan seiring bertambahnya epoch. Penurunan loss yang konsisten menunjukkan bahwa model berhasil mempelajari pola fitur gestur huruf BISINDO dengan baik.

c. Confusion Matrix

```
Classification Report:
precision  recall  f1-score  support
A         0.83    0.58    0.68     33
B         0.30    0.82    0.44     33
C         0.50    0.24    0.33     33
D         0.50    0.56    0.53     34
E         0.33    0.58    0.42     33
F         0.92    0.32    0.48     34
G         0.91    0.62    0.74     34
H         0.27    0.85    0.41     34
I         0.47    0.41    0.44     34
J         0.47    0.67    0.55     33
K         0.32    0.71    0.44     34
L         1.00    0.15    0.26     34
M         0.50    0.76    0.60     34
N         0.40    0.35    0.38     34
O         0.85    0.32    0.47     34
P         0.50    0.41    0.45     34
Q         0.67    0.18    0.28     34
R         0.29    0.56    0.38     34
S         0.87    0.38    0.53     34
T         0.89    0.24    0.37     34
U         0.48    0.35    0.41     34
V         0.71    0.29    0.42     34
W         0.71    0.59    0.65     34
X         0.89    0.47    0.62     34
Y         0.46    0.32    0.38     34
Z         0.87    0.59    0.70     34
```

Gambar IV. 19 Clasification Report Accuracy

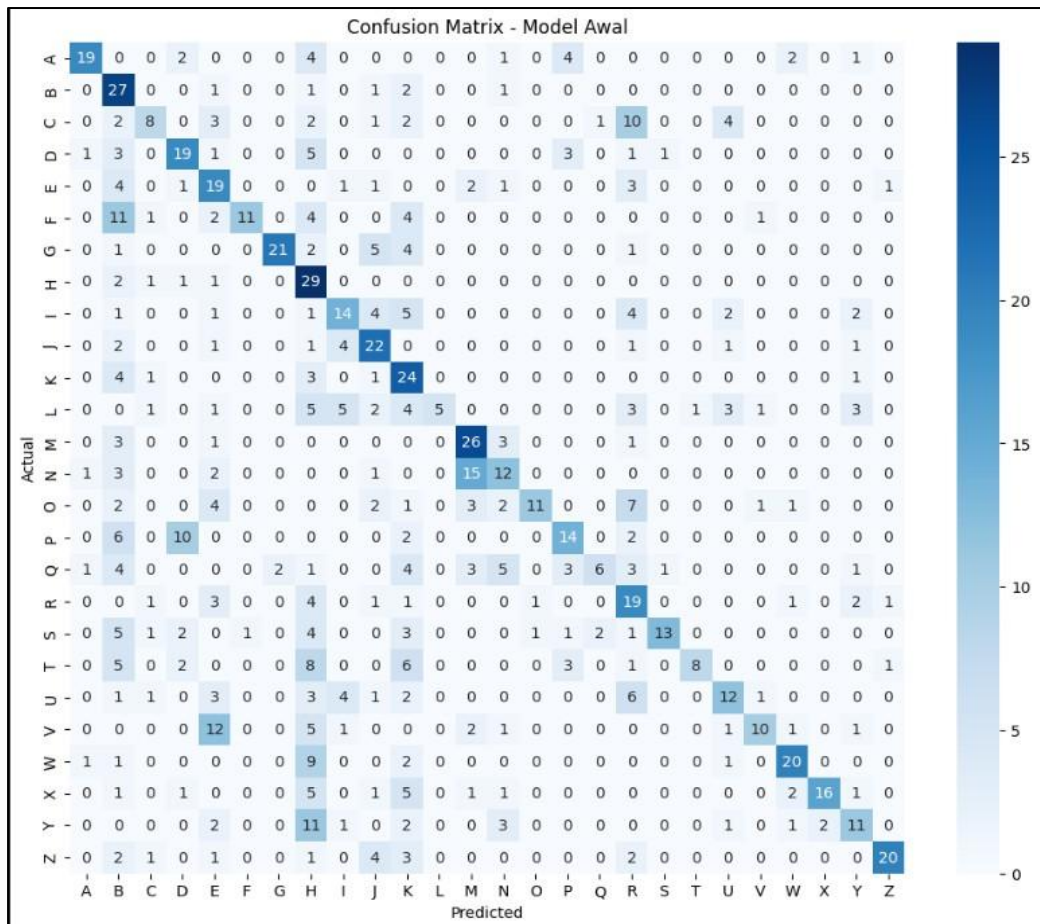
Gambar IV.19 menunjukkan hasil evaluasi model menggunakan classification report. Terlihat bahwa performa model berbeda pada tiap huruf. Beberapa huruf seperti G dan Z memiliki F1-score tinggi, sedangkan huruf seperti L dan Q masih rendah. Karena jumlah data tiap kelas seimbang, perbedaan ini kemungkinan disebabkan oleh kemiripan bentuk gestur antar huruf.

*Tabel IV. 1 Clasification Report*

No	Huruf	Precision	Recall	F1-Score
1	A	0.83	0.58	0.68
2	B	0.30	0.82	0.44
3	C	0.50	0.24	0.33
4	D	0.50	0.56	0.53
5	E	0.33	0.58	0.42
6	F	0.92	0.32	0.48
7	G	0.91	0.62	0.74
8	H	0.27	0.85	0.41
9	I	0.47	0.41	0.44
10	J	0.47	0.67	0.55
11	K	0.32	0.71	0.44
12	L	1.00	0.15	0.26
13	M	0.50	0.76	0.60
14	N	0.40	0.35	0.38
15	O	0.85	0.32	0.47
16	P	0.50	0.41	0.45
17	Q	0.67	0.18	0.28
18	R	0.29	0.56	0.38
19	S	0.87	0.38	0.53
20	T	0.89	0.24	0.37
21	U	0.48	0.35	0.41
22	V	0.71	0.29	0.42
23	W	0.71	0.59	0.65
24	X	0.89	0.47	0.62
25	Y	0.46	0.32	0.38
26	Z	0.87	0.59	0.70

Berdasarkan tabel di atas, performa model dalam mengklasifikasikan huruf A-Z menunjukkan variasi yang cukup signifikan antar kelas.

- ✓ Huruf dengan Performa Terbaik (F1-Score Tinggi)  
Huruf G (0.74), Z (0.70), A (0.68), dan W (0.65) memiliki F1-Score relatif tinggi. Hal ini menunjukkan keseimbangan yang baik antara precision dan recall, sehingga model cukup stabil dalam mengenali huruf-huruf tersebut.
- ✓ Huruf dengan Precision Tinggi tetapi Recall Rendah  
Contohnya huruf L (Precision 1.00, Recall 0.15). Artinya ketika model memprediksi L, hasilnya hampir selalu benar, tetapi model sangat jarang mendeteksi seluruh data L yang sebenarnya ada. Ini menunjukkan model terlalu konservatif dalam memprediksi huruf L.
- ✓ Huruf dengan Recall Tinggi tetapi Precision Rendah  
Contohnya huruf H (Recall 0.85, Precision 0.27) dan B (Recall 0.82, Precision 0.30). Artinya model mampu mendeteksi sebagian besar huruf tersebut, namun sering keliru karena banyak data huruf lain yang ikut diprediksi sebagai huruf tersebut.
- ✓ Huruf dengan Performa Rendah  
Huruf seperti C, L, Q, T, Y memiliki F1-Score relatif rendah ( $<0.40$ ). Hal ini menunjukkan model masih sering tertukar dalam membedakan huruf-huruf tersebut, kemungkinan karena kemiripan bentuk gestur.



Gambar IV. 20 Confusion Matrix Klasifikasi Huruf BISINDO

Gambar tersebut menampilkan Confusion Matrix Model Awal yang digunakan untuk menganalisis performa model dalam mengklasifikasikan huruf A-Z. Confusion matrix menunjukkan perbandingan antara label aktual (Actual) pada sumbu vertikal dan label prediksi (Predicted) pada sumbu horizontal.

1) Struktur Confusion Matrix

- ✓ Setiap baris menunjukkan huruf yang sebenarnya (data asli).
- ✓ Setiap kolom menunjukkan huruf hasil prediksi model.
- ✓ Nilai pada diagonal utama gambar (dari kiri atas ke kanan bawah) menunjukkan jumlah prediksi yang benar.
- ✓ Nilai di luar diagonal menunjukkan kesalahan klasifikasi (huruf tertukar).

Semakin besar nilai pada diagonal utama dan semakin kecil nilai di luar diagonal, maka semakin baik performa model.

## 2) Analisis Hasil

- ✓ Prediksi Benar (Diagonal Dominan) Beberapa huruf memiliki nilai diagonal yang cukup tinggi, seperti: H, M, W, Z. Hal ini menunjukkan bahwa model cukup baik dalam mengenali huruf-huruf tersebut.
  - ✓ Kesalahan Klasifikasi (Off-Diagonal) Terlihat adanya penyebaran nilai di luar diagonal pada beberapa huruf. Hal ini menandakan bahwa model masih sering tertukar dalam mengenali huruf tertentu.
- Contohnya:

- Beberapa huruf dengan bentuk gestur yang mirip memiliki distribusi kesalahan yang lebih tinggi.
- Huruf dengan recall rendah pada classification report sebelumnya tampak memiliki banyak prediksi yang menyebar ke kolom huruf lain.
- ✓ Distribusi Warna Intensitas warna biru menunjukkan jumlah prediksi.
  - Warna biru tua pada diagonal menandakan prediksi benar yang tinggi.
  - Warna lebih terang atau angka kecil di luar diagonal menunjukkan jumlah kesalahan.

## 3) Interpretasi Umum

Berdasarkan confusion matrix ini dapat disimpulkan bahwa:

- ✓ Model sudah mampu mengklasifikasikan sebagian huruf dengan cukup baik.
- ✓ Namun masih terdapat kesalahan klasifikasi pada beberapa huruf tertentu.
- ✓ Kesalahan ini kemungkinan disebabkan oleh:
  - Kemiripan bentuk gestur antar huruf,
  - Variasi pencahayaan atau posisi tangan,
  - Keterbatasan jumlah atau variasi dataset,
  - Model awal yang belum dioptimasi secara maksimal.

Kesimpulan akhir pada gambar confusion matrix, nilai diagonal menunjukkan prediksi yang benar. Terlihat beberapa huruf sudah cukup dominan di diagonal, namun masih terdapat penyebaran nilai di luar diagonal yang menunjukkan

kesalahan klasifikasi. Hal ini menandakan model awal masih perlu optimasi lebih lanjut

#### **IV.4.2 Pengujian Fungsional Sistem**

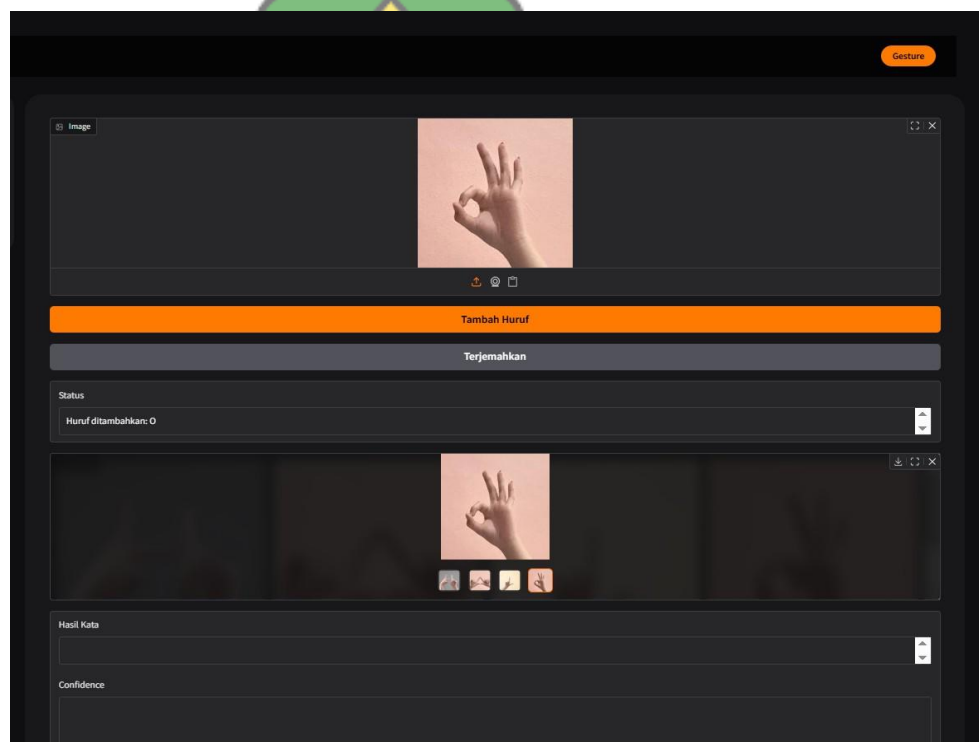
Bagian ini berbeda dengan evaluasi model, karena disini penilaian dilakukan untuk mengetahui apakah sistem berjalan sesuai skenario penggunaan. Adapun pengujian nya sebagai berikut:

a. Pengujian Mode Upload Gestur

Pengujian dilakukan dengan mengunggah beberapa citra huruf BISINDO yang berbeda.

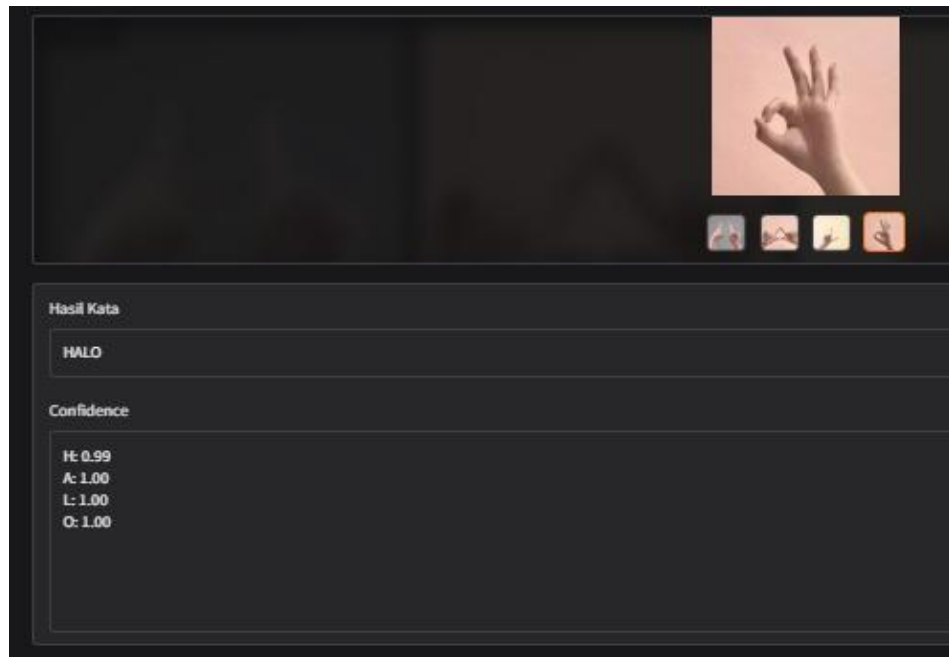
Aspek yang diuji:

- Kemampuan sistem menerima file gambar



*Gambar IV. 21 Hasil Uji Sistem Dapat menerima File Gambar*

- Confidence prediksi perhuruf



*Gambar IV. 22 Hasil Uji Confidence Perhuruf*

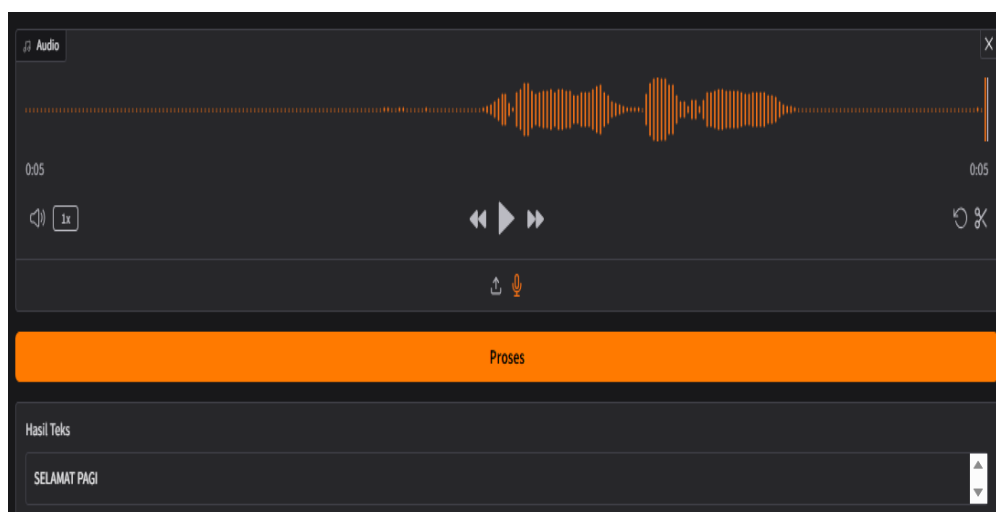
Sistem berhasil menerima input citra dan menghasilkan prediksi huruf dalam waktu kurang dari beberapa detik. Hasil prediksi konsisten dilihat dari confidence perhuruf.

b. Pengujian Mode Suara

Pengujian dilakukan dengan merekam dan mengunggah file suara berbahasa Indonesia.

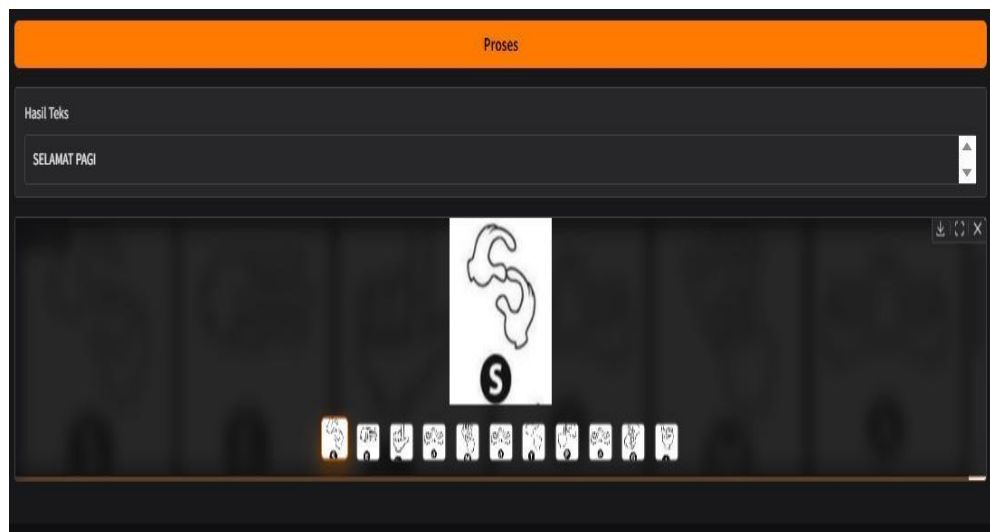
Aspek yang diuji:

- Hasil konversi suara ke teks



*Gambar IV. 23 Hasil Uji Konversi Suara ke Text*

- Visualisasi huruf sesuai teks



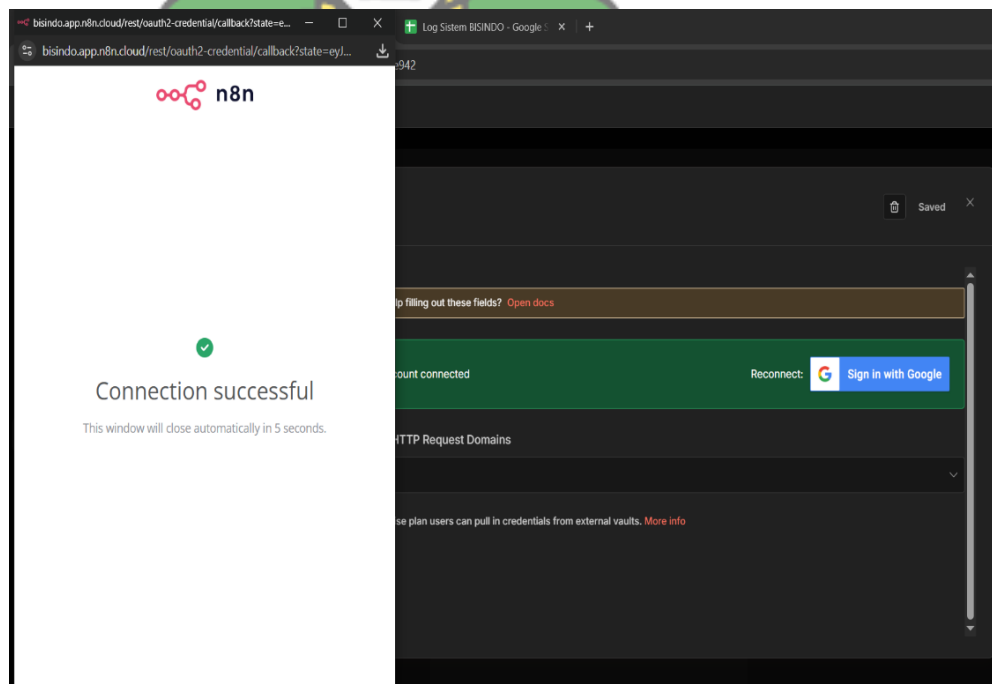
*Gambar IV. 24 Hasil Uji STT dan visualisasi huruf*

Sistem berhasil mengubah input suara menjadi teks dan menampilkan visualisasi huruf sesuai dengan teks hasil konversi.

c. Pengujian Integrasi n8n (Webhook)

Pengujian dilakukan dengan memastikan hasil terjemahan terkirim ke:

- Google Spreadsheet



*Gambar IV. 25 Hasil Koneksi n8n*



performa model, stabilitas sistem, serta keterkaitan antara metode yang digunakan dan hasil yang diperoleh.

#### **IV51 Analisis Performa Model Klasifikasi**

Berdasarkan hasil pelatihan dan evaluasi yang telah ditampilkan pada subbab sebelumnya, model MobileNetV2 berbasis transfer learning menunjukkan performa yang stabil pada data validasi.

Berikut beberapa poin analisis yang didapatkan oleh penulis:

a. Transfer Learning Efektif untuk Dataset Terbatas

Penggunaan pretrained model ImageNet membantu model mengenali pola visual dasar seperti tepi dan bentuk tangan. Hal ini terbukti dari peningkatan akurasi yang konsisten meskipun jumlah dataset relatif terbatas.

b. Fine-Tuning Meningkatkan Adaptasi Fitur

Proses membuka sebagian layer pada tahap fine-tuning memungkinkan model menyesuaikan fitur tingkat tinggi terhadap karakteristik gestur huruf BISINDO. Meskipun fine-tuning hanya dilakukan selama 3 epoch, terjadi peningkatan stabilitas prediksi.

c. Ukuran Input  $128 \times 128$  Efisien dan Cukup Representatif

Resolusi  $128 \times 128$  mampu mempertahankan detail gestur tangan tanpa membebani komputasi secara berlebihan, sehingga cocok untuk sistem berbasis web.

d. Kemungkinan Kesalahan Klasifikasi

Berdasarkan analisis confusion matrix (jika ada), kesalahan cenderung terjadi pada huruf yang memiliki kemiripan bentuk gestur. Hal ini menunjukkan bahwa variasi sudut tangan dan pencahayaan masih memengaruhi hasil prediksi.

#### **IV52 Analisis Implementasi Sistem**

Selain performa model, analisis juga dilakukan terhadap implementasi sistem secara keseluruhan.

a. Mode Upload Gestur

Sistem mampu menerima citra gestur dan menghasilkan prediksi huruf dengan waktu respons yang relatif cepat. Hal ini menunjukkan bahwa integrasi model dengan antarmuka Gradio berjalan dengan baik.

Karena sistem menggunakan metode upload gambar, stabilitas prediksi lebih terkontrol dibandingkan sistem berbasis kamera real-time, karena:

- Tidak ada fluktuasi frame
- Tidak ada noise pergerakan tangan
- Pengguna dapat memilih gambar terbaik sebelum diproses

Ini justru menjadi kelebihan sistem ini (upload gesture) dibanding real-time streaming.

b. Mode Suara → Teks → Visual Huruf

Implementasi Speech-to-Text memungkinkan sistem mengubah suara menjadi teks Bahasa Indonesia dengan cukup baik pada kondisi suara yang jelas.

Namun, performa STT sangat dipengaruhi oleh:

- Kualitas mikrofon
- Kebisingan lingkungan
- Kejelasan artikulasi

Artinya, akurasi sistem secara keseluruhan tidak hanya bergantung pada model CNN, tetapi juga pada kualitas input suara.

c. Integrasi Webhook

Integrasi webhook memungkinkan sistem tidak hanya berhenti pada proses terjemahan, tetapi juga melakukan pencatatan otomatis dan notifikasi.

Dari hasil pengujian:

- Data berhasil tersimpan di Google Spreadsheet
- Tidak terjadi kegagalan pengiriman selama pengujian

Ini menunjukkan bahwa sistem telah berkembang dari sekadar model AI menjadi sistem komunikasi terintegrasi.

#### **IV.5.3 Kesesuaian dengan Tujuan Penelitian**

Jika dikaitkan dengan tujuan penelitian pada BAB I, sistem yang dikembangkan telah:

- a. Mampu mengenali huruf BISINDO A-Z
- b. Mendukung komunikasi dua arah
- c. Berjalan melalui antarmuka web
- d. Terintegrasi dengan sistem pencatatan otomatis

Dengan demikian, tujuan penelitian dapat dikatakan tercapai.

#### **IV54 Keterbatasan Sistem**

Beberapa keterbatasan yang didapatkan oleh penulis akan sistem ini yaitu:

- a. Sistem klasifikasi masih berbasis huruf alfabet, belum kata atau kalimat utuh.
- b. Sistem gestur menggunakan upload gambar, belum berbasis deteksi gerakan dinamis.
- c. Performa model bergantung pada variasi dataset yang tersedia.
- d. Akurasi STT dipengaruhi kondisi lingkungan suara.



## BAB V

### KESIMPULAN DAN SARAN

#### V1 Kesimpulan

Berdasarkan hasil perancangan dan implementasi sistem penerjemahan Bahasa Isyarat Indonesia (BISINDO) dua arah berbasis Artificial Intelligence, dapat disimpulkan bahwa:

- a. Model klasifikasi huruf BISINDO menggunakan arsitektur MobileNetV2 dengan pendekatan transfer learning dan fine-tuning berhasil diimplementasikan untuk mengenali 26 huruf (A-Z) dengan ukuran input 128×128 piksel.
- b. Sistem dua arah berbasis web mampu mengubah gestur (upload gambar) menjadi teks dan suara, serta mengubah suara (rekam/upload audio) menjadi teks dan visual huruf BISINDO.
- c. Integrasi webhook melalui n8n memungkinkan pencatatan hasil terjemahan secara otomatis, sehingga sistem berfungsi sebagai prototipe komunikasi yang terintegrasi.

Secara umum, sistem yang dikembangkan telah memenuhi tujuan penelitian dalam mendukung komunikasi dua arah secara inklusif.

#### V2 Saran

Meskipun sistem telah berhasil diimplementasikan, masih terdapat beberapa hal yang dapat dikembangkan lebih lanjut:

- a. Mengembangkan klasifikasi pada level kata atau kalimat.
- b. Mengimplementasikan deteksi gestur berbasis video real-time.
- c. Menambah variasi dataset untuk meningkatkan akurasi dan stabilitas model.
- d. Mengoptimalkan performa Speech-to-Text pada kondisi lingkungan yang bervariasi.

## DAFTAR PUSTAKA

- [1] S. A. Sanjaya and H. F. Ilone, "BISINDO Sign Language Recognition: A Systematic Literature Review of Deep Learning Techniques for Image Processing," *International Journal of Computer Science (IJCS)*, 2023.
- [2] M. N. Eriyadi *et al.*, "Object Detection of BISINDO Sign Language Letters Using Residual Network," *Jurnal Informatika Universitas Muhammadiyah Surakarta*, 2024.
- [3] E. L. Kelana *et al.*, "Integrating the CNN Model with the Web for Indonesian Sign Language (BISINDO) Recognition," *Jurnal Politeknik Negeri Batam*, 2024.
- [4] A. Nugroho *et al.*, "Deteksi Bahasa Isyarat BISINDO Menggunakan Metode Machine Learning," *E-Journal Universitas Muhammadiyah Malang*, 2023.
- [5] A. Munandar *et al.*, "Indonesian Sign Language (BISINDO) Alphabet Detection System Using YOLO Algorithm," *Proceedings of Universitas Malikussaleh*, 2024.
- [6] I. V. Lemmuela *et al.*, "Dynamic Sign Language Recognition in Bahasa using MediaPipe, Long Short-Term Memory, and Convolutional Neural Network," *E-Journal Universitas Airlangga*, 2025.
- [7] A. Saleh, "A Comparative Analysis of CNN and SVM for Static Sign Language Recognition Using MediaPipe Landmarks," *Jurnal Universitas Negeri Surabaya*, 2025.
- [8] Widyagama "Gesture Recognition untuk Deteksi Bahasa Isyarat BISINDO: Pendekatan Mediapipe dan Random Forest," *Jurnal Teknologi Informasi Universitas Widyagama*, 2023.
- [9] B. R. Fajri *et al.*, "Media Interaktif Pengenalan Bahasa Isyarat BISINDO," *Teknologi Informasi dan Pendidikan Universitas Negeri Padang*, 2024.
- [10] R. Z. Fadillah *et al.*, "Model Penerjemah Bahasa Isyarat Indonesia (BISINDO) Menggunakan Pendekatan Transfer Learning," *Neliti Journal Database*, 2022.
- [11] E. M. Yuniarno, "Sign Language Recognition Based on Geometric Features Using Deep Learning," *E-Journal Universitas Pendidikan Ganesha*, 2025.
- [12] M. F. Abdillah and D. Andamisari, "Strategi Komunikasi Komunitas Gerakan Kesejahteraan Tunarungu Indonesia dalam Menyosialisasikan Bahasa Isyarat Indonesia di Jakarta," *Jurnal Komunikasi STIAM I*, 2023.
- [13] J. Jiang, "Sign Language Recognition Methods: Applications and Advances of Deep Learning Technology," *Darcy & Roy Press*, 2024.
- [14] IAES "Real-time Recognition of Indonesian Sign Language SIBI Using CNN-SVM Model Combination," *International Journal of Electrical and Computer Systems (IJECS)*, 2025.

- [15] I. D. A. Rachmawati *et al.*, “Deep Transfer Learning for Sign Language Image Classification: A BISINDO Dataset Study,” *Journal Binus*, 2023.
- [16] S. Isnaniah *et al.*, “The Use of Sign Language in Deaf Indonesian Classrooms in Surakarta,” *Jurnal Ilmiah Universitas Muhammadiyah Malang*, 2023.

