

# Laporan Final Project

## Kecerdasan Buatan (Lanjut)

*Pengenalan Bahasa Isyarat Level Alfabet Secara Real-Time  
dengan Mediapipe dan Arsitektur Multi-Layer Perceptron  
(MLP)*



Kelompok 07

Anggota Kelompok:

1. 23.11.5695 | Ganendra Praharsa Pertiwibowo
2. 23.11.5729 | Zerick Syahputra

## **1. Latar Belakang**

Kehidupan sehari-hari penyandang tunarungu sering menghadapi berbagai tantangan, terutama dalam hal komunikasi sosial. Masalah ini tidak hanya memengaruhi cara mereka berinteraksi, tetapi juga memberikan dampak signifikan terhadap kesehatan psikologis dan integrasi sosial. Kesulitan berkomunikasi secara lisan sering kali membuat mereka kesulitan dalam bersosialisasi dengan masyarakat umum. Hal tersebut yang dapat berakibat pada masalah psikologis seperti kurangnya rasa percayaan diri serta keterasingan sosial [1]. Oleh karena itu, sangat penting bagi kita untuk mengembangkan solusi berbasis teknologi yang bisa menjadi jembatan untuk menutup jurang komunikasi antara teman-teman Tuli dan masyarakat luas.

Di Indonesia terdapat dua sistem bahasa isyarat utama, yaitu Sistem Isyarat Bahasa Indonesia (SIBI) dan Bahasa Isyarat Indonesia (BISINDO). Berbagai penelitian menunjukkan bahwa komunitas Tuli lebih nyaman dan responsif terhadap BISINDO karena bersifat alami serta selaras dengan budaya Tuli, berbeda dengan SIBI yang lebih formal dan terstruktur . Namun, rendahnya pemahaman masyarakat umum terhadap bahasa isyarat masih menjadi hambatan dalam komunikasi dua arah yang efektif.

Meskipun demikian, penelitian ini menggunakan American Sign Language (ASL) sebagai objek pengenalan isyarat karena ketersediaan dataset publik yang terstandarisasi dan banyak digunakan dalam penelitian berbasis visi komputer. Penggunaan ASL bertujuan untuk menguji keandalan metode yang diusulkan, sehingga hasil penelitian ini diharapkan dapat menjadi dasar pengembangan sistem serupa untuk bahasa isyarat lain, termasuk BISINDO.

Perkembangan teknologi computer vision dan machine learning telah membuka kesempatan besar untuk mengembangkan sistem pengenalan gesture tangan secara otomatis yang mampu mengenali bahasa isyarat. Beberapa penelitian lain juga telah menggunakan gambar dan video untuk mengidentifikasi gestur tangan sebagai representasi bahasa isyarat dengan memanfaatkan metode *machine learning* [3], [4].

Metode ini memungkinkan sistem untuk memahami gesture tangan hanya dengan menggunakan kamera tanpa perlu alat tambahan.

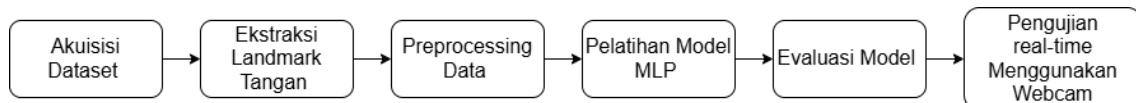
Salah satu teknologi canggih yang sering dipakai dalam sistem pengenalan gestur tangan adalah MediaPipe. Alat ini luar biasa karena bisa mengekstraksi 'titik-titik kunci' (landmark) tangan secara real-time dengan akurasi yang sangat tinggi. MediaPipe memungkinkan representasi bentuk dan posisi tangan dalam bentuk koordinat numerik, sehingga sangat sesuai untuk dikombinasikan dengan algoritma klasifikasi berbasis pembelajaran mesin [5], [6]. Penggunaan landmark tangan ini terbukti sangat efektif untuk sistem pengenalan bahasa isyarat, baik pada aplikasi real-time maupun perangkat mobile.

Untuk proses klasifikasi, kita mengandalkan *Multi-Layer Perceptron* (MLP), sejenis arsitektur *Artificial Neural Network* (ANN) yang mampu memodelkan hubungan non-linear antar fitur secara efisien. MLP telah banyak digunakan secara luas pada berbagai tugas klasifikasi dan telah menunjukkan hasil yang memuaskan ketika digabungkan dengan fitur-fitur yang diambil secara representatif [7]. Sejumlah penelitian juga menunjukkan bahwa MLP mampu bersaing dengan algoritma machine learning lain dalam hal akurasi dan stabilitas prediksi, terutama pada data yang berbasis fitur numerik [8] [9].

## 2. Metode

### 2.1 Alur Pengerjaan (Flow Sistem)

Metode yang digunakan dalam penelitian ini dirancang bertahap, mulai dari akuisisi dataset hingga implementasi sistem pengenalan bahasa isyarat secara real-time. Alur umum sistem dapat dilihat sebagai berikut:



Gambar 2. 1 Alur Pengerjaan (Flow Sistem)

#### 1. Akuisisi Dataset

Penelitian ini menggunakan koleksi gambar American Sign Language (ASL) yang berisi gestur tangan statis dari setiap alfabet, ditambah kelas 'Space' dan 'Nothing'.

Kami hanya memanfaatkan data dari direktori /ASL\_Dataset/Train. Direktori /ASL\_Dataset/Test tidak digunakan karena rasionya terlalu kecil untuk menghasilkan data uji yang representatif. Oleh karena itu, semua proses pembagian (split) data menjadi data latih dan data uji dilakukan secara keseluruhan dari konten direktori /ASL\_Dataset/Train.

Struktur data sangat efisien: setiap kelas disimpan dalam folder terpisah (misalnya, folder "A", "B", dll.). Ini memungkinkan pelabelan data dilakukan secara otomatis berdasarkan nama folder, sehingga tidak memerlukan pelabelan manual tambahan.

## **2. Deteksi Landmark Tangan Menggunakan MediaPipe**

Setelah mendapatkan dataset, tahap berikutnya adalah deteksi tangan dan ekstraksi landmark dengan bantuan MediaPipe Hands. MediaPipe digunakan karena kemampuannya dalam mendeteksi tangan secara akurat dan efisien dengan pendekatan berbasis machine learning.

Pada tahap ini digunakan modul mp\_hands.Hands() yang menghasilkan:

- Deteksi keberadaan tangan
- 21 titik landmark tangan untuk setiap tangan
- Masing-masing landmark direpresentasikan dalam koordinat (x, y) yang telah dinormalisasi dalam rentang 0–1

Parameter penting yang digunakan dalam proses ini meliputi:

- min\_detection\_confidence, yang berfungsi mengatur tingkat kepercayaan minimum saat mendeteksi tangan
- min\_tracking\_confidence, yang berfungsi menjaga kestabilan pelacakan landmark di antara frame

Tahap ini bertujuan untuk mengubah data gambar mentah menjadi representasi numerik yang lebih informatif dan efisien untuk proses klasifikasi.

### **3. Ekstraksi Fitur dan Label (Preprocessing Data)**

Pada tahap preprocessing, setiap gambar yang berhasil terdeteksi tangan akan diekstraksi fitur landmark-nya. Dari 21 landmark tangan yang terdeteksi, diambil koordinat x dan y, sehingga total fitur yang dihasilkan adalah  $21 \text{ landmark} \times 2 = 42$  fitur

Setiap vektor fitur berukuran 42 dimensi kemudian dipasangkan dengan label kelas sesuai nama folder asal gambar tersebut. Proses ini dilakukan untuk seluruh data latih dan data uji.

Hasil ekstraksi fitur dan label selanjutnya disimpan dalam format file pickle, yaitu:

- train\_data.pickle untuk data latih
- test\_data.pickle untuk data uji

Penyimpanan ini bertujuan untuk mempercepat proses pelatihan dan evaluasi model tanpa perlu mengekstraksi ulang landmark setiap kali sistem dijalankan.

## **2.2 Algoritma yang Digunakan**

### **1. Arsitektur Multi-Layer Perceptron (MLP)**

Model klasifikasi yang digunakan dalam penelitian ini adalah Multi-Layer Perceptron (MLP), yaitu salah satu *Artificial Neural Network* (ANN) yang terdiri dari beberapa lapisan neuron.

Arsitektur MLP dibangun menggunakan Keras Sequential API dengan konfigurasi sebagai berikut:

- Input layer: 42 neuron, sesuai dengan jumlah fitur landmark tangan

- Hidden layer(s): beberapa lapisan tersembunyi dengan fungsi aktivasi non-linear (misalnya ReLU)
- Output layer: 28 neuron, merepresentasikan kelas A–Z, Space, dan Nothing, dengan fungsi aktivasi Softmax.

### 3. Dataset

#### 3.1 Sumber Dataset

Dataset yang digunakan dalam penelitian ini bersumber dari dataset publik yang diambil dari platform Kaggle, yang bernama *American Sign Language Alphabet Dataset*. Dataset tersebut dapat diakses melalui link berikut:

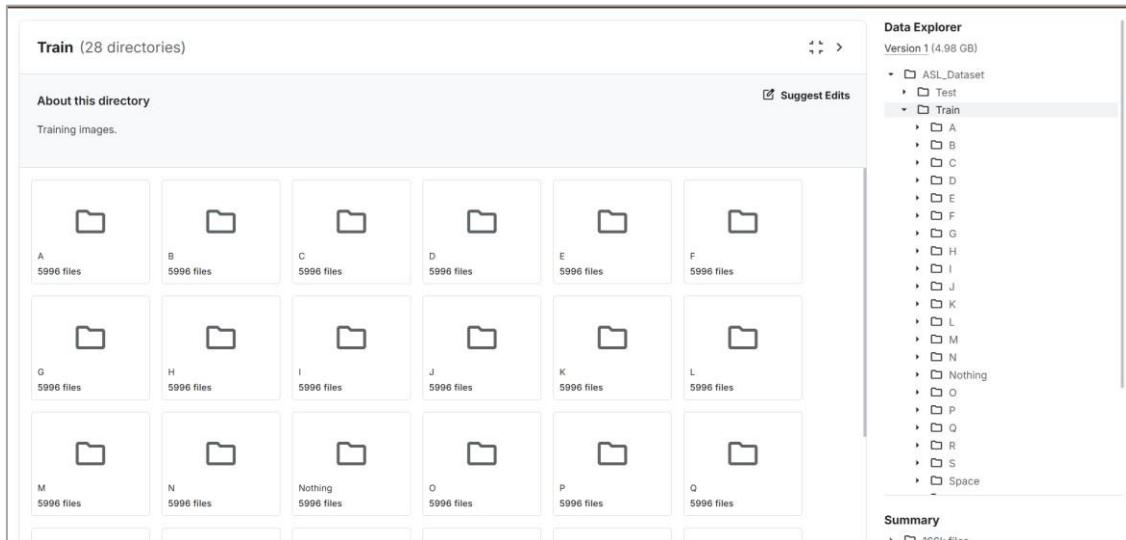
<https://www.kaggle.com/datasets/kapillondheamerican-sign-language>

Pemilihan dataset publik ini didasarkan pada pertimbangan mengenai ketersediaan data yang cukup besar, pengorganisasian data yang teratur, serta relevansinya dengan tujuan penelitian, yaitu pengenalan bahasa isyarat pada level alfabet.

#### 3.2 Gambaran Umum Dataset

Dataset *American Sign Language Alphabet* berisi kumpulan gambar tangan yang menunjukkan gerakan bahasa isyarat untuk setiap alfabet. Setiap gambar menampilkan satu tangan yang membentuk simbol dari alfabet tertentu, dengan latar belakang yang sederhana dan berbagai posisi tangan.

Contoh tampilan dataset bahasa isyarat alfabet dapat dilihat pada Gambar 3.1, yang memperlihatkan sampel gambar dari kelas alfabet yang berbeda.



Gambar 3. 1 Dataset American Sign Language Alphabet

Dataset ini terbagi menjadi beberapa kelas alfabet yang masing-masing disimpan dalam folder yang berbeda. Setiap folder mewakili satu label kelas, sehingga proses pelabelan data dapat dilakukan secara otomatis berdasarkan struktur folder.

Secara umum, karakteristik dataset ini adalah sebagai berikut:

- Tersedia dalam bentuk gambar statis
- Setiap gambar mewakili satu kelas alfabet dalam bahasa isyarat
- Menampilkan variasi dalam orientasi dan posisi tangan
- Sesuai untuk proses ekstraksi fitur landmark tangan menggunakan Mediapipe.

## 4. Hasil Pengujian

Proses pengujian dibagi menjadi dua tahap, yaitu evaluasi model menggunakan data uji (test data) dan pengujian langsung secara real-time menggunakan webcam secara real-time.

### 4.1 Skenario Pengujian

#### 4.1.1 Evaluasi Model dengan Data Test

Pada skenario ini, model MLP diuji menggunakan data test yang telah dipisahkan dari data pelatihan. Data test terdiri dari gambar bahasa

isyarat alfabet yang direpresentasikan dalam bentuk vektor fitur berukuran 42, hasil ekstraksi koordinat (x, y) dari 21 landmark tangan menggunakan MediaPipe Hands.

Pengujian dilakukan dengan memberikan seluruh data test ke dalam model untuk memperoleh hasil prediksi kelas. Hasil prediksi tersebut kemudian dibandingkan dengan label sebenarnya untuk menghitung nilai akurasi, serta mengevaluasi performa model menggunakan precision, recall, dan f1-score melalui classification report. Skenario ini bertujuan untuk mengukur kemampuan generalisasi model terhadap data yang tidak digunakan selama proses pelatihan.

#### **4.1.2 Pengujian Real-time Menggunakan Webcam**

Skenario pengujian real-time dilakukan untuk menguji performa sistem dalam kondisi penggunaan langsung. Pengujian ini menggunakan webcam sebagai input video secara real-time.

Pada setiap frame kamera:

- MediaPipe Hands mendeteksi tangan dan mengekstraksi landmark tangan
- Landmark tangan diubah menjadi vektor fitur berukuran 42
- Vektor fitur diproses oleh model MLP untuk memprediksi kelas bahasa isyarat
- Sistem menampilkan hasil prediksi secara langsung pada layar berupa huruf hasil prediksi beserta nilai confidence, bounding box, dan visualisasi landmark tangan

Skenario ini bertujuan untuk menguji responsivitas, stabilitas prediksi, serta kesiapan sistem untuk digunakan secara interaktif.

## 4.2 Hasil Evaluasi Model dengan Data Test

Berdasarkan hasil pengujian menggunakan data test, Multilayer Perceptron (MLP) menunjukkan performa yang sangat baik. Model berhasil mengenali hampir seluruh kelas huruf dalam bahasa isyarat dengan tingkat akurasi yang tinggi.

Hasil evaluasi memperlihatkan bahwa:

- Model mencapai akurasi sebesar 99,98%
- Hampir seluruh kelas huruf berhasil diprediksi dengan sangat baik
- Classification report menunjukkan nilai precision, recall, dan f1-score sebesar 1,00 untuk seluruh kelas huruf
- Huruf dengan jumlah sampel relatif sedikit, seperti huruf D dan kelas Space, tetap dapat diprediksi dengan baik tanpa kesalahan klasifikasi

Hasil ini menunjukkan bahwa fitur landmark tangan yang diekstraksi menggunakan MediaPipe mampu merepresentasikan bentuk bahasa isyarat alfabet secara efektif dan mendukung performa klasifikasi yang tinggi.

```
584/584 [=====] - 1s 2ms/step - loss: 0.0012 - accuracy: 0.9998
Test Accuracy: 99.98%
```

Gambar 4. 1 Persentase Test Accuracy Model

Classification Report:				
	precision	recall	f1-score	support
A	1.00	1.00	1.00	877
B	1.00	1.00	1.00	887
C	1.00	1.00	1.00	635
D	1.00	1.00	1.00	239
E	1.00	1.00	1.00	865
F	1.00	1.00	1.00	872
G	1.00	1.00	1.00	869
H	1.00	1.00	1.00	873
I	1.00	1.00	1.00	869
J	1.00	1.00	1.00	886
K	1.00	1.00	1.00	898
L	1.00	1.00	1.00	748
M	1.00	1.00	1.00	371
N	1.00	1.00	1.00	77
O	1.00	1.00	1.00	525
P	1.00	1.00	1.00	435
Q	1.00	1.00	1.00	417
R	1.00	1.00	1.00	889
S	1.00	1.00	1.00	768
Space	1.00	1.00	1.00	321
T	1.00	1.00	1.00	791
U	1.00	1.00	1.00	670
V	1.00	1.00	1.00	853
W	1.00	1.00	1.00	882
X	1.00	1.00	1.00	865
Y	1.00	1.00	1.00	845
Z	1.00	1.00	1.00	459
accuracy			1.00	18686
macro avg	1.00	1.00	1.00	18686
weighted avg	1.00	1.00	1.00	18686

Gambar 4. 2 Classification Report Hasil Evaluasi Model

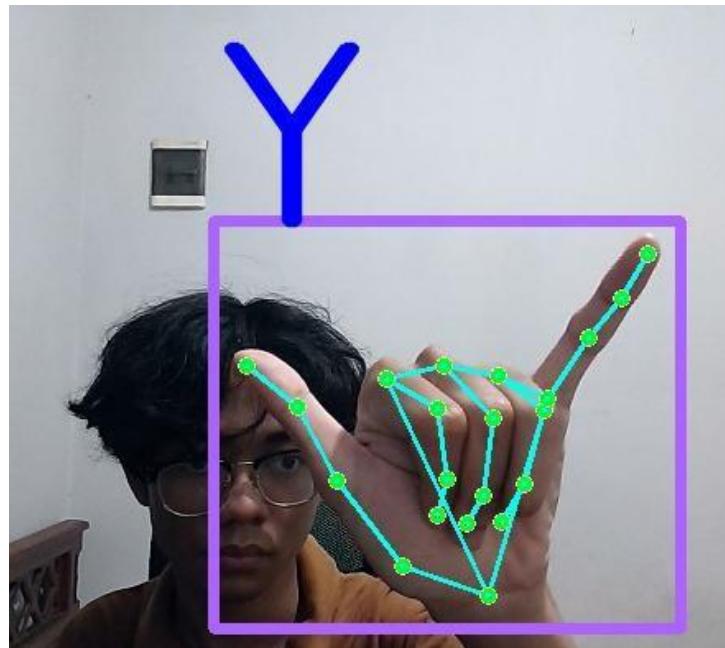
### 4.3 Hasil Pengujian Real-time dengan Webcam

Pengujian real-time dilakukan menggunakan kamera webcam untuk mengevaluasi kinerja sistem pengenalan bahasa isyarat alfabet dalam kondisi penggunaan langsung. Pada pengujian ini, sistem memproses setiap frame video secara real-time untuk mendeteksi tangan, mengekstraksi landmark tangan menggunakan Mediapipe Hands, dan melakukan klasifikasi huruf menggunakan model Multilayer Perceptron (MLP) yang telah dilatih sebelumnya.

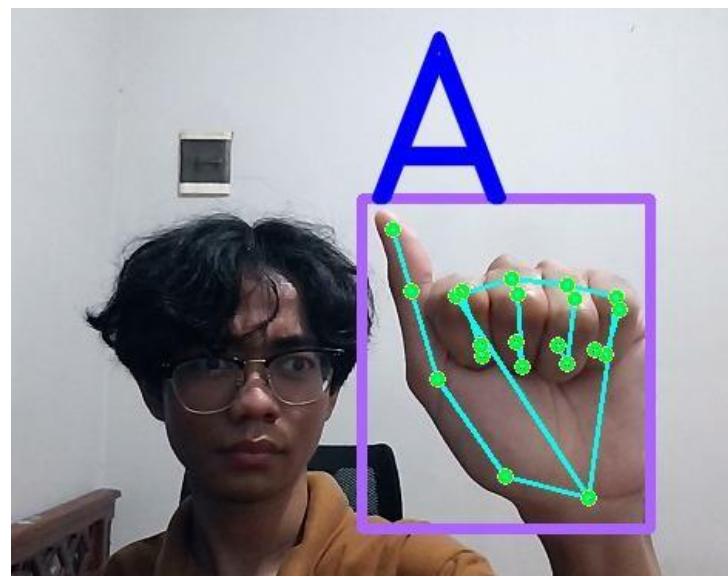
Hasil pengujian real-time menunjukkan bahwa sistem mampu menampilkan prediksi huruf bahasa isyarat secara langsung di layar, disertai dengan bounding box yang mengelilingi tangan pengguna.

Namun demikian, pada beberapa alfabet yang memiliki gestur tangan yang mirip, seperti U dan R, hasil prediksi dapat berubah apabila posisi jari tidak dibentuk dengan jelas dan konsisten. Pada kondisi selain tersebut, sistem mampu menghasilkan prediksi yang stabil meskipun terjadi sedikit pergerakan atau perubahan posisi tangan selama pengujian realtime menggunakan webcam.

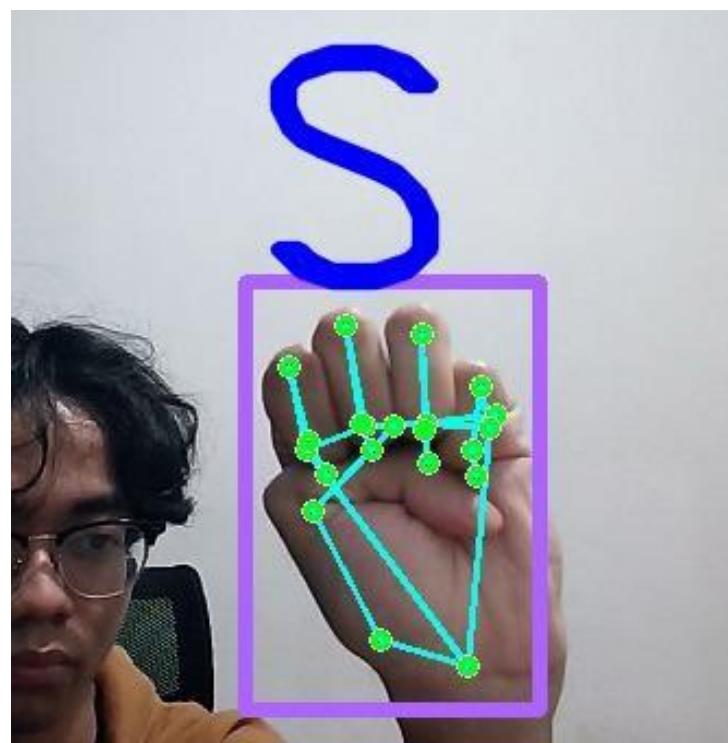
Sebagai ilustrasi, Gambar 4.3 hingga Gambar 4.6 menampilkan contoh hasil pengujian real-time pada beberapa huruf bahasa isyarat alfabet yang berbeda.



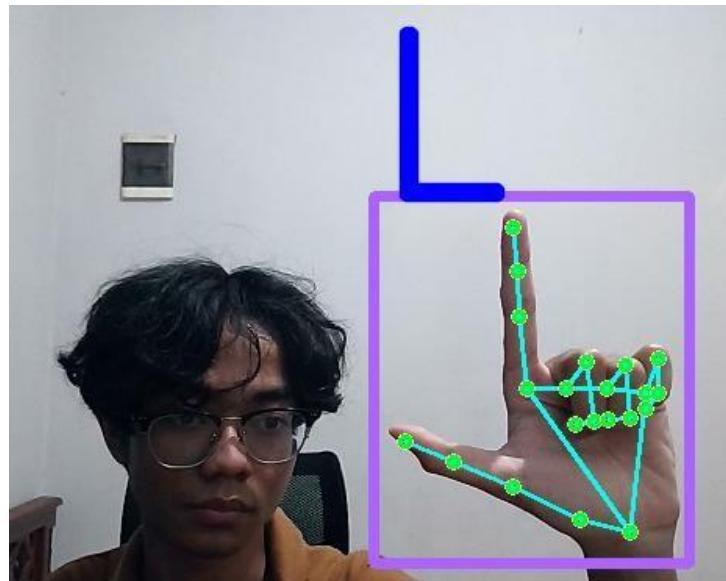
Gambar 4. 3 Demonstrasi Real-time Huruf Y



Gambar 4. 4 Demonstrasi Real-time Huruf A



Gambar 4. 5 Demonstrasi Real-time Huruf S



Gambar 4. 6 Demonstrasi Real-time Huruf L

## 5. Analisa Hasil

Hasil pengujian menunjukkan bahwa sistem pengenalan bahasa isyarat alfabet berbasis MediaPipe Hands dan Multilayer Perceptron (MLP) memiliki performa sangat baik. Model mencapai akurasi 99,98% pada data uji, dengan nilai precision, recall, dan f1-score mendekati 1,00 pada hampir seluruh kelas, yang menandakan kemampuan pembelajaran dan generalisasi yang tinggi.

Representasi 42 koordinat landmark tangan (x, y) dari MediaPipe terbukti efektif dalam menangkap pola bentuk isyarat, sehingga mendukung kinerja MLP dalam klasifikasi, termasuk pada kelas dengan jumlah data terbatas seperti huruf D dan Space.

Pada pengujian real-time menggunakan webcam, sistem mampu melakukan prediksi secara responsif dan relatif stabil selama tangan terdeteksi dengan baik. Namun, pada gestur alfabet yang memiliki kemiripan bentuk, seperti U dan R, perubahan kecil posisi jari dapat memengaruhi hasil prediksi.

Secara keseluruhan, kombinasi MediaPipe sebagai ekstraktor fitur dan MLP sebagai pengklasifikasi merupakan pendekatan yang efektif untuk pengenalan bahasa isyarat alfabet.

## 6. Kesimpulan

Berdasarkan hasil perancangan, implementasi, dan pengujian sistem pengenalan bahasa isyarat alfabet yang telah dilakukan, dapat disimpulkan beberapa hal berikut ini:

- a) Sistem pengenalan bahasa isyarat alfabet berbasis MediaPipe Hands dan arsitektur Multilayer Perceptron (MLP) berhasil dibangun dan diimplementasikan dengan baik.
- b) Proses ekstraksi fitur menggunakan 21 landmark tangan MediaPipe yang direpresentasikan dalam koordinat (x, y) terbukti efektif dalam merepresentasikan bentuk bahasa isyarat alfabet.
- c) Model MLP yang dilatih menggunakan fitur landmark tangan mampu mencapai akurasi pengujian sebesar 99,98% pada data test, dengan nilai precision, recall, dan f1-score yang sangat tinggi pada hampir seluruh kelas.
- d) Sistem mampu melakukan pengenalan bahasa isyarat alfabet secara real-time menggunakan webcam dengan hasil prediksi yang cukup stabil dan responsif.
- e) Perubahan kecil pada posisi atau gerakan tangan secara umum tidak memengaruhi hasil prediksi sistem. Namun, pada beberapa alfabet dengan bentuk gestur yang mirip, seperti U dan R, diperlukan posisi dan konfigurasi jari yang jelas dan tepat agar tidak terjadi kesalahan klasifikasi. Hal ini menunjukkan bahwa sensitivitas lebih dipengaruhi oleh kemiripan antar huruf isyarat dibandingkan oleh perubahan posisi landmark tangan.

Dengan demikian, dapat disimpulkan bahwa sistem yang dikembangkan memiliki potensi untuk menjadi solusi awal dalam membantu proses komunikasi bahasa isyarat, khususnya pada level alfabet, serta dapat dikembangkan lebih lanjut untuk mendukung bahasa isyarat yang lebih kompleks dan aplikatif.

## 7. Referensi

- [1] S. Mambela, "TINJAUAN UMUM MASALAH PSIKOLOGIS DAN MASALAH SOSIAL INDIVIDU PENYANDANG TUNANETRA," *J. Buana Pendidik.*, vol. 14, no. 25, pp. 65–73, 2018.
- [2] R. A. Mursita, "Respon Tunarungu Terhadap Penggunaan Sistem Bahasa Isyarat Indonesia ( Sibi ) Dan," *Inklusi*, vol. 2, no. 2, pp. 221–232, 2015.
- [3] A. Nugroho, R. Setiawan, A. Harris, and Beny, "Deteksi Bahasa Isyarat Bisindo Menggunakan Metode Machine Learning," *Process. J. Ilm. Sist. Informasi, Teknol. Inf. dan Sist. Komput.*, vol. 18, no. 2, pp. 152–158, 2023.
- [4] I. Suyudi, S. Sudadio, and S. Suherman, "Pengenalan Bahasa Isyarat Indonesia menggunakan Mediapipe dengan Model Random Forest dan Multinomial Logistic Regression," *J. Ilmu Siber dan Teknol. Digit.*, vol. 1, no. 1, pp. 65–80, 2023, doi: 10.35912/jisted.v1i1.1899.
- [5] G. B. Prananta, H. A. Azzikri, and C. Rozikin, "DETEKSI DAN PENGENALAN GESTUR TANGAN SECARA REAL-TIME MENGGUNAKAN JARINGAN SARAF TIRUAN KONVOLUSIONAL," *Method. J. Tek. Inform. dan Sist. Inf.*, vol. 9, no. 2, pp. 30–34, 2023, doi: 10.46880/mtk.v9i2.1911.
- [6] I. R. W. Putra, H. Pratikno, and Musayyanah, "Sistem Deteksi Simbol pada SIBI (Sistem Isyarat Bahasa Indonesia) Menggunakan Mediapipe Berbasis Android," *JCONES - J. Control Netw. Syst.*, vol. 10, no. 2, pp. 50–59, 2021, [Online]. Available: <https://repository.dinamika.ac.id/id/eprint/5814/9/17410200032-2021-UNIVERSITASDINAMIKA.pdf>
- [7] K. F. Sugiantari, P. H. Suputra, L. J. E. Dewi, and G. B. P. Putra, "Pendekatan MLP Dalam Klasifikasi Bahasa Isyarat: Analisis Jarak Euclidean Landmark Tangan," *JIKA (Jurnal Inform.)*, vol. 9, no. 2, p. 209, 2025, doi: 10.31000/jika.v9i2.13368.
- [8] M. F. Alam, A. Nuryaman, P. H. Khotimah, A. Parlina, and A. Sihombing, "Optimizing Multi-Layer Perceptron Performance in Sentiment Classification through Neural Network Feature Extraction," *BACA J. Dokumentasi dan Inf.*, vol. 46, no. 1, pp. 1–14, 2025, doi: 10.55981/baca.2025.8240.
- [9] D. W. Setiawan, N. T. Lapatta, and D. W. Nugraha, "Performance Comparison of Multilayer Perceptron ( MLP ) and Random Forest for Early Detection of Cardiovascular Disease," vol. 9, no. 6, pp. 3012–3024, 2025.

## 8. Kontribusi & distribusi anggota kelompok

No.	NIM	Nama	Kontribusi
1	23.11.5695	Ganendra Praharsa Pertiwibowo	<ul style="list-style-type: none"> <li>Melakukan pencarian dan pemilihan dataset bahasa isyarat alfabet yang sesuai sebagai data penelitian.</li> <li>Merancang dan mengembangkan kode program sistem pengenalan bahasa isyarat alfabet menggunakan MediaPipe dan arsitektur MLP.</li> <li>Melaksanakan proses pelatihan dan pengujian model, baik menggunakan data test maupun pengujian real-time dengan webcam</li> <li>Melakukan analisis terhadap hasil pengujian sistem untuk mengevaluasi performa dan kestabilan model.</li> <li>Melakukan pengecekan keseluruhan laporan untuk memastikan kesesuaian isi, konsistensi metode, serta kebenaran hasil yang disajikan.</li> <li>Membuat video demonstrasi project sebagai media dokumentasi dan presentasi hasil sistem.</li> </ul>
2	23.11.5729	Zerick Syahputra	<ul style="list-style-type: none"> <li>Mengusulkan dan merumuskan ide awal project pengenalan bahasa isyarat alfabet sebagai topik final project mata kuliah Kecerdasan Buatan Lanjut.</li> <li>Menyusun dan menyesuaikan format laporan sesuai dengan pedoman penulisan yang ditetapkan.</li> <li>Menyusun dan mengintegrasikan seluruh isi laporan, mulai dari pendahuluan, metode, hingga hasil dan kesimpulan.</li> </ul>

No.	NIM	Nama	Kontribusi
			<ul style="list-style-type: none"> <li>• Menyusun dan mendesain materi presentasi dalam bentuk slide PowerPoint untuk keperluan pemaparan hasil project.</li> </ul>