

BAB I

PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi kecerdasan buatan (*Artificial Intelligence/AI*) dan visi computer (*computer vision*) telah memberikan kontribusi besar dalam berbagai aspek kehidupan, seperti layanan kesehatan, sistem pendidikan, dan bidang komunikasi. Salah satu keunggulan utama dari teknologi ini adalah kemampuannya dalam mengenali pola visual serta mengolah pergerakan secara *real-time*, yang menjadikannya semakin luas diterapkan dalam aplikasi-aplikasi berbasis interaksi antara manusia dan mesin. Salah satu framework yang cukup dikenal dalam pengolahan citra dan video adalah **MediaPipe**, yang dirancang untuk menganalisis pergerakan tubuh, ekspresi wajah, serta gestur tangan secara efisien. Dengan dukungan teknologi tersebut, berbagai hambatan dalam komunikasi visual dapat diminimalkan, termasuk dalam upaya menerjemahkan bahasa isyarat secara otomatis.

American Sign Language (ASL) merupakan salah satu bentuk bahasa isyarat yang paling banyak digunakan secara global, terutama di Kawasan Amerika Serikat dan negara-negara sekitarnya. Di Indonesia sendiri, terdapat dua sistem utama bahasa isyarat, yaitu Bahasa Isyarat Indonesia (BISINDO) dan Sistem Isyarat Bahasa Indonesia (SIBI), yang digunakan oleh komunitas tunarungu dalam aktivitas sehari-hari. Walaupun setiap negara memiliki karakteristik tersendiri dalam sistem bahasa isyarat, esensi dari komunikasi melalui gerakan tangan tetap memiliki kesamaan. Namun demikian, masih terdapat hambatan komunikasi antara pengguna bahasa isyarat dan individu yang tidak terbiasa dengan sistem tersebut. Hal ini mendorong kebutuhan akan solusi teknologi yang mampu mengenali dan menerjemahkan bahasa isyarat secara otomatis guna meningkatkan inklusivitas dan kemudahan berkomunikasi lintas kelompok.

Salah satu kendala yang cukup menonjol dalam membangun sistem penerjemah bahasa isyarat berbasis teknologi adalah pengenalan simbol

huruf dan angka dalam ASL. Simbol-simbol ini umum digunakan dalam konteks formal seperti pengejaan nama, penyebutan alamat, atau istilah teknis yang tidak memiliki padanan isyarat tertentu. Sayangnya, bentuk jari dalam alfabet manual ASL kerap memiliki kemiripan satu sama lain, sehingga menyulitkan sistem visi komputer dalam membedakan gerakan yang serupa. Faktor lain seperti sudut pengambilan gambar oleh kamera dan variasi personal dalam melakukan gestur turut memengaruhi tingkat akurasi sistem dalam mengenali isyarat secara tepat. Oleh sebab itu, dibutuhkan pendekatan yang lebih presisi dan adaptif dalam menangani pengenalan huruf dan angka pada bahasa isyarat.

Dengan memanfaatkan framework MediaPipe, sistem dapat mengidentifikasi pergerakan tangan secara lebih cepat dan presisi, sehingga mendukung proses penerjemahan ASL, khususnya untuk komponen huruf dan angka, secara lebih optimal. Teknologi ini menawarkan potensi besar dalam pengembangan berbagai aplikasi, seperti penerjemah otomatis, sistem pembelajaran inklusif, dan interaksi manusia dengan komputer berbasis gestur. Melalui penelitian ini, dikembangkan sebuah model penerjemah ASL berbasis MediaPipe dengan fokus pada klasifikasi huruf dan angka, yang diharapkan dapat memberikan solusi praktis dan terjangkau bagi komunitas tunarungu maupun masyarakat secara umum.

1.2 Rumusan Masalah

Berdasarkan uraian pada latar belakang, maka permasalahan yang dibahas dalam penelitian ini dapat dirumuskan sebagai berikut:

- a. Bagaimana merancang sistem penerjemah bahasa isyarat untuk pengenalan huruf dan angka dalam ASL dengan pendekatan *Hand Gesture Recognition*?
- b. Bagaimana proses pengolahan data visual dari gerakan tangan dapat dilakukan secara tepat dan efisien menggunakan framework MediaPipe?

- c. Bagaimana membangun arsitektur model Convolutional Neural Network (CNN) yang mampu mengklasifikasikan gestur ASL secara optimal?
- d. Bagaimana metode evaluasi yang digunakan untuk mengukur kinerja sistem penerjemah bahasa isyarat yang dikembangkan?

1.3 Tujuan

Adapun tujuan dari penelitian ini adalah sebagai berikut:

- Mengembangkan sebuah aplikasi yang mampu mengenali bahasa isyarat menggunakan teknologi MediaPipe dan bahasa pemrograman Python.
- Mengimplementasikan MediaPipe untuk mengekstraksi ciri-ciri gerakan tangan secara langsung (*real-time*).
- Merancang dan melatih model Convolutional Neural Network (CNN) guna melakukan klasifikasi terhadap gestur dalam bahasa isyarat ASL.
- Melakukan proses pengujian serta analisis performa model berdasarkan tingkat akurasi dan sejumlah metrik evaluasi lainnya.

1.4 Manfaat

a. Bagi Masyarakat

Penelitian ini diharapkan mampu meningkatkan kesadaran publik serta mendorong terciptanya lingkungan yang lebih inklusif, khususnya dalam menjembatani komunikasi antara masyarakat umum dan penyandang tunarungu. Dengan adanya teknologi ini, akses komunikasi menjadi lebih terbuka dan mudah terjangkau.

b. Bagi Universitas

Temuan dari penelitian ini dapat dimanfaatkan sebagai referensi akademik dalam pengembangan teknologi pengenalan bahasa isyarat berbasis MediaPipe dan Python. Selain itu, hasilnya dapat dijadikan bahan

pembelajaran bagi mahasiswa di masa mendatang serta memacu lahirnya inovasi dalam bidang *computer vision* dan pengenalan pola.

c. Bagi Penulis

Proyek penelitian ini menjadi sarana bagi penulis untuk meningkatkan kemampuan teknis, khususnya dalam penerapan teknologi pengenalan visual dan pembelajaran mesin (*machine learning*). Pengalaman ini juga menjadi landasan untuk melakukan pengembangan lebih lanjut dalam ranah interaksi antara manusia dan komputer.

1.5 Batasan Masalah

- a. Penelitian ini menggunakan model MediaPipe Hands untuk mendeteksi isyarat alfabet dari huruf A hingga Y serta angka 1 hingga 10. Adapun huruf J dan Z tidak disertakan karena karakteristik gerakannya bersifat dinamis, sehingga memerlukan proses pelacakan gerakan yang lebih kompleks dan tidak dapat dijangkau dalam cakupan penelitian ini.
- b. Dataset yang digunakan mencakup 34 kelas gestur tangan. Data ini dibagi menjadi tiga bagian, yaitu 80% untuk proses pelatihan (training), 10% untuk validasi (validation), dan 10% sisanya untuk pengujian (testing), agar model yang dikembangkan dapat dilatih secara optimal dan terukur.
- c. Sistem pengenalan bahasa isyarat dalam penelitian ini dirancang menggunakan kamera webcam sebagai perangkat input utama untuk menangkap citra gerakan tangan dari pengguna secara langsung.
- d. Fokus utama dari penelitian ini adalah mengoptimalkan performa model agar mampu memproses input video secara langsung (real-time) dengan latensi yang rendah, sehingga sistem dapat memberikan respons yang cepat dan akurat dalam mengenali gestur yang ditampilkan.

1.6 Metodologi Penelitian

1.6.1. Tempat dan Waktu Penelitian

Penelitian ini dilaksanakan di beberapa lokasi berbeda. Kegiatan studi literatur dan analisis awal dilakukan di Perpustakaan Universitas Bhinneka Nusantara, sedangkan tahapan perancangan sistem, pelatihan model, serta pengujian dilakukan di Laboratorium Rekayasa Perangkat Lunak UBHINUS dan/atau secara daring dari tempat tinggal peneliti. Data sekunder berupa dataset bahasa isyarat ASL diperoleh melalui repositori daring yang terpercaya.

Durasi penelitian berlangsung selama enam bulan, dimulai dari Desember 2024 hingga Mei 2025. Rincian jadwal kegiatan disajikan pada tabel berikut:

Tabel 1.1 Waktu Penelitian

Kegiatan	Tahun 2024/2025					
	Des	Jan	Feb	Mar	Apr	Mei
Studi Literatur & Perencanaan						
Pengumpulan Data						
Pengembangan sistem (Desain & Implementasi Awal dengan MediaPipe)						
Pelatihan Model						
Evaluasi dan Pengujian Model						
Analisis Hasil dan Penyusunan Laporan						

1.6.2. Bahan dan Alat Penelitian

Untuk mendukung validasi penelitian dan replikasi oleh peneliti lain, berikut adalah spesifikasi bahan serta alat yang digunakan:

- **Bahan Penelitian:**

1. Dataset ASL: Dataset publik berisi citra gestur alfabet dan angka ASL (contoh: *Sign Language Multi-Class Classification with CNN*), mencakup jumlah kelas, ukuran gambar, dan format file.
 2. Literatur Ilmiah: Berbagai sumber akademik seperti jurnal, buku, dan artikel terkait topik *computer vision*, pengenalan gestur tangan, ASL, MediaPipe, serta metode *machine learning*.
- **Alat Penelitian:**
 - *Perangkat Keras:*
 - Laptop/komputer dengan prosesor minimal Intel Core i5, RAM 8GB, dan GPU (NVIDIA GeForce GTX atau setara).
 - Kamera digital beresolusi minimal 720p untuk keperluan pengambilan gambar primer.
 - *Perangkat Lunak:*
 - Sistem operasi: Windows 10/11, Linux (Ubuntu), atau macOS.
 - Bahasa pemrograman: Python ≥ 3.7 .
 - Pustaka/Framework utama: MediaPipe, TensorFlow, Keras, OpenCV, NumPy, Pandas, Scikit-learn.
 - Lingkungan pengembangan: Jupyter Notebook, Google Colab, atau Visual Studio Code.

1.6.3. Teknik Pengumpulan Data dan Informasi

Pengumpulan data dilakukan melalui dua jalur utama:

1. **Data Citra Gestur ASL:**
 - *Dataset Publik:* Mengakses dan mengunduh dataset citra bahasa isyarat dari repositori terbuka (misalnya Kaggle). Jika

dalam format video, dilakukan ekstraksi ke dalam bentuk *frame*.

- *Pengambilan Data Primer*: Dilakukan pengambilan gambar gestur langsung dengan kamera digital apabila diperlukan variasi tambahan. Proses ini mencakup pemilihan gestur, partisipan, pencahayaan, dan latar belakang yang sesuai.
- 2. **Pra-pemrosesan Data:**
 - Seleksi kualitas gambar.
 - Pemotongan area tangan (ROI).
 - Penyesuaian ukuran gambar (resize) ke dimensi standar.
 - Teknik augmentasi (rotasi, pencahayaan, penambahan noise) bila diperlukan.
- 3. **Ekstraksi Fitur dengan MediaPipe:**
 - Setiap gambar akan diproses menggunakan solusi *Hand Tracking* dari MediaPipe untuk memperoleh koordinat 21 titik landmark (x, y, z).
- 4. **Studi Literatur:**
 - Mengumpulkan referensi dari jurnal, buku, dan dokumentasi terkait topik penelitian.

1.6.4. Analisis Data

Analisis data dilakukan dengan pendekatan kuantitatif deskriptif dan inferensial:

- **Deskriptif**: Menyajikan statistic dasar dari dataset, distribusi kelas, dan hasil pra-pemrosesan.
- **Evaluasi Model**:

Model yang dikembangkan akan dianalisis berdasarkan:

 - *Accuracy*
 - *Precision*
 - *Recall*

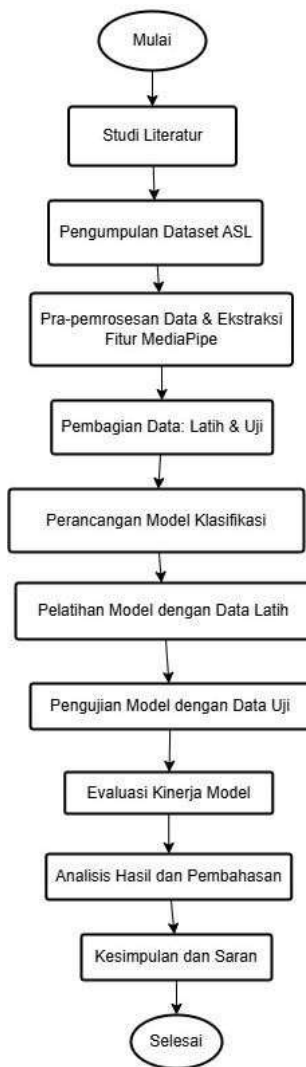
- *F1-Score*
- *Confusion Matrix*

Analisis juga mencakup perbandingan performa model pada konfigurasi atau arsitektur berbeda bila relevan.

1.6.5. Prosedur Penelitian

Langkah-langkah penelitian dirancang secara sistematis dan dijabarkan sebagai berikut:

1. Studi literatur untuk memahami dasar teori dan penelitian terdahulu.
2. Pengumpulan dataset (baik dari repositori publik maupun pengambilan primer).
3. Pra-pemrosesan data dan ekstraksi koordinat landmark tangan menggunakan MediaPipe.
4. Perancangan dan pengembangan model klasifikasi (CNN atau alternatif lain).
5. Pelatihan model dengan data latih.
6. Pengujian model menggunakan data uji yang terpisah.
7. Evaluasi performa model dengan metrik klasifikasi.
8. Analisis hasil untuk mengidentifikasi kekuatan dan keterbatasan model.
9. Penarikan kesimpulan dan penyusunan rekomendasi.
10. Penyusunan laporan akhir secara sistematis.



Gambar 1.1 Diagram Alir Penelitian

1.7 Sistematika Penulisan

Laporan Tugas Akhir ini disusun dalam lima bab utama yang masing-masing memuat pokok bahasan sebagai berikut:

- **Bab I – Pendahuluan**

Memuat uraian umum mengenai latar belakang masalah, perumusan masalah, tujuan dan manfaat penelitian, batasan ruang lingkup, metode yang digunakan, serta gambaran sistematika penulisan.

- **Bab II – Tinjauan Pustaka**

Berisi landasan teori yang mendasari penelitian, termasuk penjelasan konsep-konsep utama seperti computer vision, bahasa isyarat ASL, MediaPipe, serta kajian pustaka dari penelitian terdahulu yang relevan.

- **Bab III – Metodologi Penelitian**

Menjelaskan langkah-langkah penelitian secara sistematis, termasuk lokasi, bahan dan alat, teknik pengumpulan data, tahapan implementasi, dan pemrosesan data.

- **Bab IV – Hasil dan Pembahasan**

Menyajikan hasil pengembangan dan pengujian sistem, serta analisis terhadap performa model berdasarkan data yang diperoleh.

- **Bab V – Kesimpulan dan Saran**

Memberikan ringkasan hasil penelitian, menjawab rumusan masalah, serta menyampaikan saran untuk pengembangan lebih lanjut di masa depan.