

BAB II

TINJAUAN PUSTAKA

2.1 Bahasa Isyarat

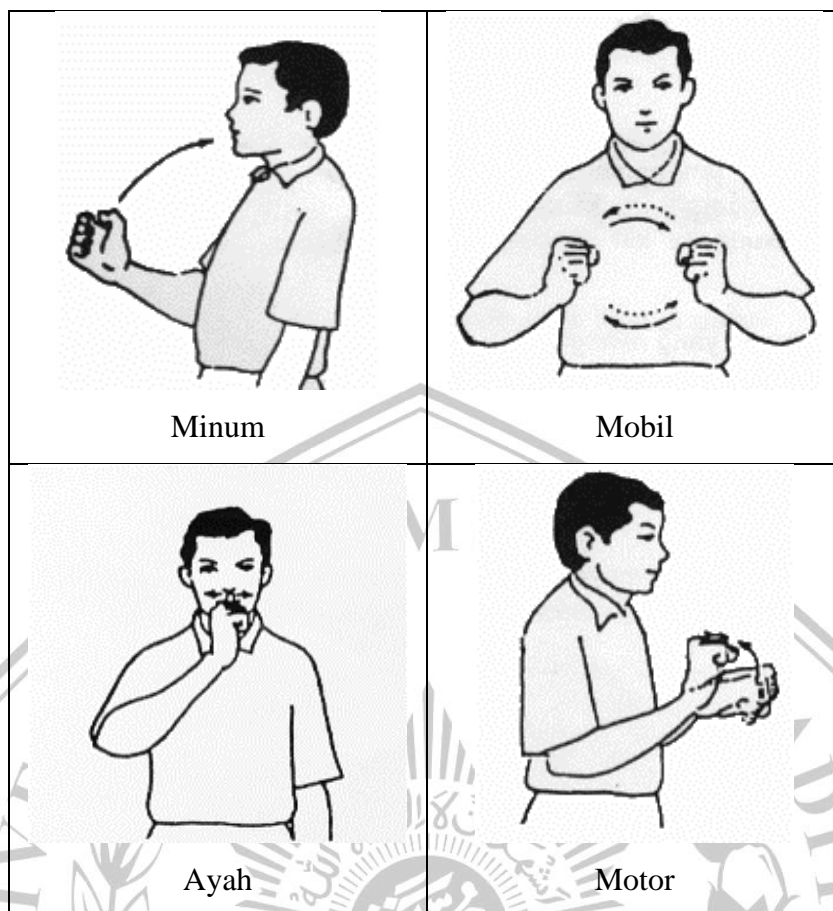
Bahasa isyarat (*sign language*) ialah sistem bahasa menggunakan gerakan tangan, jari, ekspresi wajah, dan tubuh untuk menyampaikan pesan.[8]. Bahasa isyarat adalah sistem bahasa yang memiliki tata bahasa sendiri dan digunakan oleh komunitas yang memiliki gangguan pendengaran atau mereka yang ingin berkomunikasi dengan mereka. Ini adalah bahasa yang sangat kaya dan kompleks, dengan struktur gramatikal dan kosakata yang berkembang sendiri. Di seluruh dunia, ada berbagai jenis bahasa isyarat yang digunakan oleh berbagai komunitas, seperti Bahasa Isyarat Amerika, Bahasa Isyarat Inggris Britania, Bahasa Isyarat Jepang, dan banyak lagi.

SIBI (Sistem Isyarat Bahasa Indonesia) merupakan salah satu dari dua pedoman bahasa isyarat yang digunakan oleh masyarakat tuna rungu yang ada di Indonesia. Penggunaan bahasa isyarat SIBI lebih dikhususkan pada instansi pendidikan khususnya sekolah luar biasa. Sedangkan, BISINDO dibuat oleh GERKATIN (Gerakan Kesejahteraan Tunarungu Indonesia) kemudian digunakan secara umum. SIBI yang terbilang lebih formal jika dibandingkan dengan BISINDO yang gaya bahasanya terkesan lebih santai. Beberapa anggota komunitas tuli di Indonesia menganggap SIBI sebagai bahasa isyarat yang lebih "baku" atau formal. Sedangkan, BISINDO lebih mencerminkan bahasa isyarat yang lebih alami dan otentik yang digunakan oleh komunitas tuli, dan lebih banyak dianggap sebagai bahasa yang mewakili identitas budaya mereka.

Bahasa Isyarat yang dipilih dalam penelitian ini ada pada Bahasa Isyarat pada kamus SIBI (Sistem Isyarat Bahasa Indonesia). Gambar- gambar didapatkan dari kamus SIBI yang ada pada SLB Negeri Batu. Berikut tabel 10 kata yang diambil dari kamus SIBI.

Tabel 2.1 Kamus SIBI 10 kata

Bahasa isyarat	
A line drawing of a man in a polo shirt. He is facing forward with his right hand raised, palm facing forward, and his left hand near his chest. A dashed line indicates the movement of his right hand.	A line drawing of a man in a polo shirt. He is facing forward with his right hand raised, index and middle fingers extended, pointing towards his forehead.
Halo	Ibu
A line drawing of a man in a polo shirt. He is facing forward with his right hand raised, palm facing forward, and his index finger pointing towards his chest.	A line drawing of a man in a polo shirt. He is shown in profile, facing right, with his right hand raised to his forehead, palm facing forward.
Saya	Bangun
A line drawing of a man in a polo shirt. He is shown in profile, facing right, with his right hand raised to his mouth, index finger pointing towards his lips.	A line drawing of a man in a polo shirt. He is facing forward with his hands raised, palms facing each other, and his fingers slightly curled.
Makan	Tidur



Sumber: SLB Negeri Batu

2.2 Pengolah Citra Digital

Pengolahan citra digital (*Digital Image Processing*) ialah perangkat lunak atau algoritma komputer yang digunakan untuk mengambil, memproses, mengedit, dan menganalisis gambar atau citra digital. Pengolah citra digital memungkinkan manipulasi gambar digital untuk berbagai tujuan, seperti perbaikan kualitas gambar, pengenalan pola, analisis medis, dan banyak aplikasi lainnya. Citra ialah representasi dua dimensi untuk ditampilkan tiga dimensi secara nyata [9].

Citra ialah fungsi kontinyu (*continue*) pada bidang dengan intensitas cahaya dua dimensi. Suatu citra harus dipresentasikan dengan numerik berisi nilai diskrit, supaya dapat diolah komputer digital. Representasi pada fungsi kontinyu yang diubah menjadi nilai diskrit dinamakan digitalisasi citra.

Sebuah citra digital diwakili oleh matriks dua dimensi $f(x,y)$ yang terdiri dari M kolom dan N baris, dimana perpotongan antara kolom dan baris dinamakan *pixel* (*picture element*) dari sebuah citra.

$$f(x,y) \approx \begin{bmatrix} f(0,0) & f(0,1) & \dots & f(0,M-1) \\ f(1,0) & f(1,1) & \dots & f(1,M-1) \\ \vdots & \vdots & \vdots & \vdots \\ f(N-1,0) & f(N-1,1) & \dots & f(N-1,M-1) \end{bmatrix}$$

Suatu citra $f(x,y)$ dalam fungsi matematis dapat dituliskan sebagai berikut:

$$0 \leq x \leq M-1$$

$$0 \leq y \leq N-1$$

$$0 \leq f(x,y) \leq G-1$$

dimana :
 M = jumlah piksel baris (*row*) pada array citra
 N = jumlah piksel kolom (*column*) pada array citra
 G = nilai skala keabuan (*graylevel*)

Proses deteksi pada citra digital yang digunakan untuk mendeteksi manusia sebagai sensor secara visual. Sistem deteksi manusia ialah proses deteksi visual untuk menangkap kehadiran manusia dalam sebuah video yang diproses oleh kamera[10].

Deteksi tubuh pada penelitian kali ini menggunakan teknologi *Mediapipe Holistic* mengabungkan model berbeda untuk elemen pose, wajah dan tangan. Dioptimalkan untuk area tertentu. *Holistic* memproyeksikan pose pada tubuh manusia menggunakan *detector BlazePose* dan *landmark* berikut. Menggunakan *landmark* pose inferred diperoleh *three crop regions of interested (ROI)* pada dua tangan tangan, wajah, dan juga badan. Kemudian, menggabungkan semua landmark dengan model pose.

2.2 Mediapipe

MediaPipe adalah *Framework* untuk membangun *pipeline machine learning* untuk memproses data deret waktu seperti video, audio, dll. *Framework* lintas *platform* ini berfungsi di *Desktop/Server, Android, iOS*, dan perangkat tersemat seperti *Raspberry Pi* dan *Jetson Nano*. *MediaPipe* membangun saluran *ML multi-modal* untuk pengembang adalah kerangka kerja. Kerangka node dan tepi

untuk melacak titik-titik koordinat yang disebut *landmarks* pada bagian-bagian tubuh[11]. Namun, meskipun *Mediapipe* memiliki banyak kelebihan, platform ini juga memiliki beberapa keterbatasan. Misalnya, pemrosesan media yang berat dapat membutuhkan banyak sumber daya komputer, yang dapat menyebabkan masalah performa. Oleh karena itu, penting bagi pengembang untuk mempertimbangkan keterbatasan ini sebelum menggunakan *Mediapipe* dalam proyek mereka.

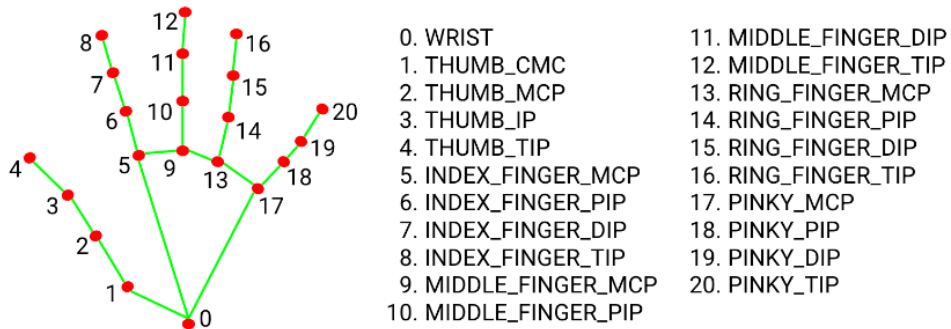
Landmark adalah titik-titik penting yang digunakan sebagai acuan dalam proses pengenalan dan analisis suatu objek. Dalam konteks *Mediapipe*, *landmark* memiliki peran yang sangat penting dalam pengenalan dan analisis wajah manusia. *Landmark* dalam *Mediapipe* ditemukan dengan menggunakan algoritma deteksi wajah dan pengenalan pola. Algoritma ini dapat menentukan posisi dari titik-titik penting pada wajah, seperti sudut mata, ujung hidung, dll. Titik-titik penting ini kemudian digunakan sebagai acuan dalam proses pembuatan model 3D wajah. *Landmark* adalah titik-titik penting yang digunakan sebagai acuan dalam proses pengenalan dan analisis suatu objek. Dalam konteks *Mediapipe*, deteksi dari wajah, tangan, dan tubuh diekstraksi sebagai titik kunci (*keypoints*) dan *landmark* menggunakan teknik augmentasi data bawaan dari rangkaian frame input yang diambil dari *webcam* [12].

2.3 *Mediapipe Holistic*

MediaPipe Holistic adalah salah satu solusi yang disediakan oleh kerangka kerja *MediaPipe* yang dikembangkan oleh *Google*. Ini adalah solusi yang dirancang untuk melakukan analisis tubuh manusia dalam video secara *real-time*. *Mediapipe Holistic* mengimplementasikan *Mediapipe Pose*, *Mediapipe Face Mesh*, dan *Mediapipe Hands* masing-masing menghasilkan total 543 *landmark* (33 *landmark* untuk pose, 468 *landmark* untuk wajah, dan 21 *landmark* untuk setiap tangan) menyediakan analisis *holistic* terhadap seluruh tubuh manusia dalam video [13].

Mediapipe Holistic mengintegrasikan beberapa komponen *Mediapipe* yang berbeda termasuk deteksi pada wajah, deteksi pada pose, dan deteksi pada tangan sebagai berikut.

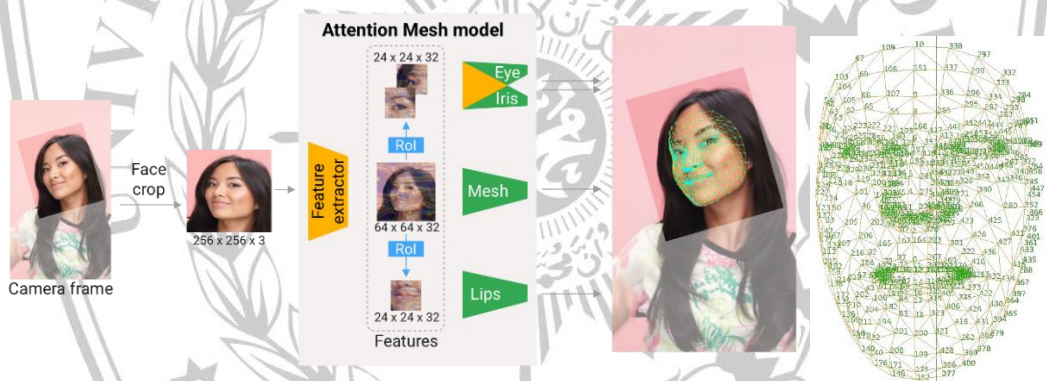
2.4.1 Mediapipe Hand Tracking



Gambar 2.1 Mediapipe Hand Tracking

MediaPipe Hands berfungsi mendeteksi telapak tangan dan jari menggunakan *Machine Learning* (ML) untuk menyimpulkan 21 *landmark* 3D tangan dengan *single frame*.

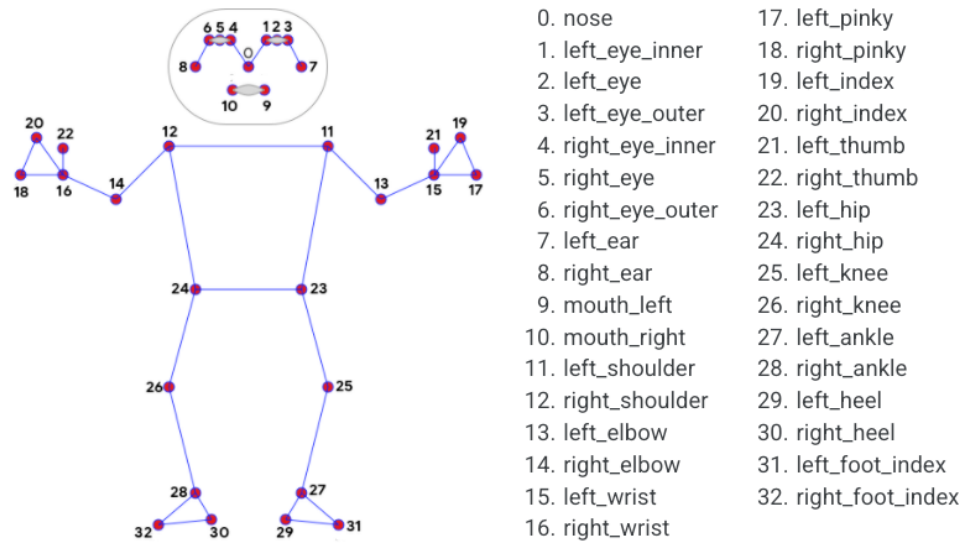
2.4.2 Face Mesh



Gambar 2.2 Mediapipe FaceMesh

MediaPipe Face Mesh adalah solusi yang memperkirakan 468 *landmark* wajah 3D secara *real-time*. Sistem deteksi ini menggunakan pembelajaran mesin (ML) untuk menyimpulkan permukaan wajah 3D, hanya membutuhkan input kamera tunggal tanpa memerlukan sensor kedalaman khusus. Memanfaatkan arsitektur model yang ringan bersama dengan akselerasi GPU di seluruh rangkaian, solusi ini menghadirkan kinerja waktu nyata yang penting untuk pengalaman langsung.

2.4.3 Human Pose Detection and Tracking



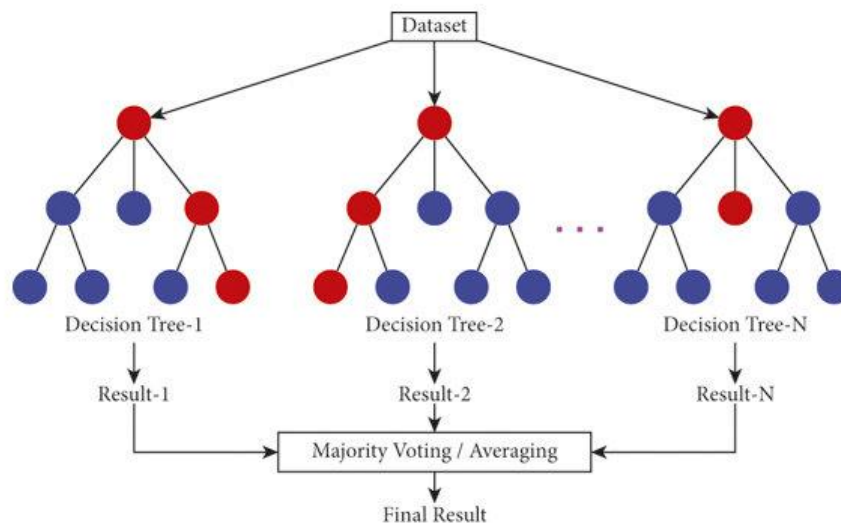
Gambar 2.3 Mediapipe Pose Detection

MediaPipe Pose adalah solusi ML untuk deteksi pose tubuh dengan fidelitas tinggi, 33 landmark 3D, dan masker segmentasi latar belakang di seluruh tubuh dari bingkai video RGB menggunakan model *BlazePose*.

2.4 Klasifikasi *Random Forest*

Random Forest merupakan salah satu algoritma *machine learning* sebagai pengklasifikasian *dataset* dengan jumlah yang cukup besar. Algoritma *Random Forest* memiliki kelebihan menghasilkan *accuracy* yang sangat bagus [14]. *Random Forest* ialah suatu metode *ensemble learning* yang dapat digunakan dalam *machine learning* untuk meningkatkan kinerja dan ketepatan prediksi model. *Ensemble learning* menggabungkan prediksi dari beberapa model supaya dapat menghasilkan hasil yang jauh lebih baik daripada yang dapat dicapai oleh model tunggal.

Klasifikasi dilakukan dengan cara menggabungkan *tree* dalam sistem *decision tree* dengan aturan training *dataset*. Penentuan klasifikasi menggunakan *Random Forest* dihitung berdasarkan hasil *voting* dan *tree*. Berikut adalah gambar analisa *Random Forest*.



Gambar 2.4 *Random Forest Voting Data*

Algoritma *Random Forest* bekerja dengan urutan langkah-langkah berikut:

1. Pemilihan sampel acak yang diambil dari dataset pelatihan. Proses ini dikenal sebagai "*bootstrapping*" atau "*bagging*" (*Bootstrap Aggregating*).
2. Pembentukan Pohon Keputusan (*Decision Tree*) untuk setiap sampel yang diambil secara acak, sebuah pohon keputusan dibangun pada subset tersebut. Pada setiap node dalam pohon keputusan, hanya sebagian kecil fitur yang dipertimbangkan untuk pembagian node berikutnya. Jumlah fitur ini biasanya merupakan akar kuadrat dari total fitur. Hal ini membantu mencegah *overfitting* dan membuat model lebih seragam.
3. Dilakukan proses klasifikasi, ini melibatkan *voting* mayoritas dari pohon-pohon tersebut. Pilihan yang paling banyak diambil sebagai prediksi akhir. Pada kasus *regresi*, prediksi dapat berupa *mean* atau *median* dari prediksi semua pohon.
4. Penggabungan hasil melalui *majority voting* pada klasifikasi atau *averaging* untuk *regresi*. Dengan cara ini, *Random Forest* dapat mengatasi *overfitting* yang mungkin terjadi pada pohon keputusan tunggal.

2.5 *OpenCV*

OpenCV (*Open Source Computer Vision*) ialah software pustaka open source untuk *computer vision* dan *machine learning* yang dapat digunakan secara *real-time* [15]. Penerapan *OpenCV* dapat digunakan untuk deteksi objek, pelacakan objek, pengenalan karakter menggunakan metode *Artificial Intelligence*. *OpenCV*

menyediakan beberapa algoritma sederhana yang terkait dengan *Computer Vision Library* ini ditulis dalam bahasa *C*, *C++*, *Python*, *Java*, dan lainnya. *OpenCV* dapat bekerja dengan baik secara *real-time*. *OpenCV* juga dapat dijalankan dengan berbagai sistem operasi yaitu *Windows*, *Linux*, dan *Mac OS* [16].

2.6 Python

Python ialah bahasa pemrograman tingkat tinggi yang sangat populer dan sering dipakai untuk pengembangan pada perangkat lunak, pengembangan web, ilmu data, kecerdasan buatan. *Python* merupakan pemrograman yang sering digunakan oleh *developer* untuk meningkatkan aplikasi berbasis *mobile*, *desktop*, dan *web* [17].

Bahasa pemrograman *Python* menggunakan struktur data tingkat tinggi, sintaks mudah dipahami, dan modul siap untuk digunakan sehingga membuatnya mudah dipelajari. *Python* memiliki banyak pustaka (*libraries*) dan modul yang mempermudah pengembangan. Salah satu pustaka yang terkenal adalah *NumPy* untuk pemrosesan matriks, *Pandas* untuk analisis data, *Matplotlib* untuk visualisasi data, dan *TensorFlow* untuk pembelajaran mesin dan kecerdasan buatan.