

BAB III

METODE PENELITIAN

3.1. Tahapan Metode Penelitian

Penelitian ini dilakukan dengan menggunakan metode eksperimen. Metode penelitian eksperimen merupakan salah satu jenis metode penelitian kuantitatif yang digunakan untuk mencari hubungan sebab-akibat antara variabel independen dan dependen (Patzer, 1996). Penelitian ini bertujuan untuk membuat sistem penerjemah bahasa isyarat berbasis *machine learning*. Adapun harapannya dapat digunakan untuk membantu mengatasi permasalahan komunikasi yang dihadapi para pengidap tunarungu. Gambar 3.1 menjelaskan diagram alir tahapan metode penelitian.



Gambar 3. 1 - Diagram alir penelitian

Tahap awal dalam penelitian ini yaitu melakukan studi literatur. Hal ini dilakukan dengan cara mengamati, memahami, serta mengumpulkan informasi, konsep dan teori yang berkaitan dengan penelitian. Studi literatur ini mencakup pencarian informasi tentang *machine learning*, khususnya teknologi *computer vision*: mediapipe, dan teknologi *natural language processing*: Open AI dan *Web Speech API* yang digunakan dalam penelitian terkait. Informasi yang dikumpulkan didapatkan dari buku, jurnal, publikasi, serta sumber daya lain yang relevan.

Setelah melakukan studi literatur, langkah selanjutnya adalah identifikasi masalah dengan tujuan untuk mengidentifikasi suatu masalah dalam penelitian-penelitian sebelumnya yang berkaitan dengan pembuatan purwarupa sistem penerjemah bahasa isyarat guna mendapatkan solusi untuk pengembangan sistem yang lebih baik. Penelitian ini difokuskan pada pembuatan sistem penerjemah dengan menjadikan output sistem deteksi dari kata menjadi kalimat secara *real-time* dengan menggabungkan teknologi *Computer Vision* dan *Natural Language Processing*.

Setelah selesai mengidentifikasi masalah, selanjutnya dilakukan perancangan dan pembuatan sistem untuk menjawab/mengatasi permasalahan yang didapatkan sebelumnya. Sistem yang dibuat meliputi input, proses, dan output.

Dalam penelitian ini, pengujian dilakukan untuk mengetahui keberhasilan sistem yang telah dibuat. Parameter dari keberhasilan tersebut dilihat berdasarkan dua pengujian, yaitu uji fungsionalitas dan uji *usability*.

1. Uji Fungsionalitas

Uji fungsionalitas dilakukan untuk memastikan bahwa sistem yang dibuat dapat bekerja dengan baik. Beberapa hal yang akan diuji dalam uji fungsionalitas antara lain:

- Kemampuan deteksi gerakan tangan dan terjemahan bahasa isyarat ke dalam bentuk tulisan secara *real-time*.
- Kemampuan deteksi berbagai jenis gerakan tangan dan bahasa isyarat yang berbeda.
- Kemampuan deteksi suara untuk diubah ke dalam bentuk teks.

Terdapat Batasan pengujian yang perlu diperhatikan dalam uji fungsionalitas sistem penerjemah isyarat yang telah dibuat. Pertama, jarak antara objek yang akan diuji dengan kamera minimal 50 cm dan maksimal 100 cm. Hal ini penting agar objek dapat direkam dengan jelas dan mendapatkan hasil yang optimal. Selain itu, dalam ruangan pengujian, cahaya yang harus tersedia dalam ruangan tertutup yaitu minimal menggunakan lampu LED 9 watt, menghasilkan 920 lumen setara dengan lampu pijar 80 watt. Ketersediaan cahaya yang memadai memastikan kondisi pencahayaan yang cukup untuk mendapatkan hasil yang akurat dan konsisten selama proses pengujian.

Kedua, pengujian sistem hanya dilakukan dengan menggunakan laptop sebagai perangkat pengguna. Hal ini dilakukan untuk memastikan konsistensi dan kestabilan dalam pengujian, mengingat perbedaan antarmuka perangkat lain seperti *smartphone* atau tablet. Penggunaan laptop memberikan keseragaman dalam antarmuka yang digunakan dalam pengujian. Selain itu, menggunakan webcam untuk memastikan kualitas gambar yang optimal. Webcam biasanya memiliki kualitas yang lebih baik dalam menangkap detail dan gerakan dibandingkan dengan hanya menggunakan kamera laptop saja.

Terakhir, sistem penerjemah bahasa isyarat yang diuji harus terkoneksi dengan internet. Koneksi internet diperlukan untuk mengakses sumber daya atau layanan yang diperlukan dalam sistem.

2. Uji Usability

Uji *usability* dilakukan untuk memastikan bahwa sistem yang dibuat, mudah digunakan oleh pengguna dan dapat memenuhi kebutuhan pengguna. Beberapa hal yang akan diuji dalam uji *usability* antara lain:

- Kemudahan dalam navigasi dan penggunaan app/web
- Ketersediaan fitur-fitur yang dibutuhkan pengguna untuk berkomunikasi melalui bahasa isyarat
- Kepuasan pengguna terhadap aplikasi dan pengalaman pengguna dalam menggunakan app/web

Dalam melakukan uji *usability* terhadap sistem yang dibuat, dilakukan

pengujian secara langsung terhadap beberapa teman tunarungu sebagai responden dengan memberikan angket berupa pernyataan terkait dengan sistem yang sudah dibuat dan di-ujikan. Pernyataan dalam angket ini berasal dari skripsi dan jurnal yang telah melalui proses validasi oleh peneliti yang berwenang. Pertanyaan yang diberikan mencakup tanggapan responden mengenai aplikasi untuk mengenali ragam jenis gesture Bahasa Isyarat Indonesia dan tanggapan mengenai pentingnya aplikasi tersebut (Rachardi, 2020) ditunjukkan pada tabel 3.1 dan 3.2. Adapun pernyataan lainnya terkait dengan fungsi dari sistem yang telah dibuat, yaitu tanggapan responden mengenai efisiensi sistem penerjemah Bahasa Isyarat (Hikmatia A.E & Ihsan Zul, 2021) ditunjukkan pada Tabel 3.3.

Tabel 3. 1 - Tanggapan responden: Aplikasi untuk mengenali gesture BISINDO

Pernyataan	S	TS
Aplikasi untuk mengenali ragam jenis gesture Bahasa Isyarat Indonesia sangat berguna dalam mendukung komunikasi dua arah antara penyandang tunarungu-tunawicara dengan orang tanpa kebutuhan khusus		

Keterangan:

S = Setuju

TS = Tidak Setuju

Tabel 3. 2 - Tanggapan mengenai pentingnya aplikasi tersebut

Pernyataan	1	2	3	4	5
Aplikasi untuk mengenali ragam jenis gesture Bahasa Isyarat Indonesia sangat berguna dalam mendukung komunikasi dua arah antara penyandang tunarungu-tunawicara dengan orang tanpa kebutuhan khusus					

Keterangan:

1 = Sangat Penting

- 2 = Penting
 3 = Biasa
 4 = Tidak Penting
 5 = Sangat Tidak Penting

Tabel 3. 3 - Tanggapan responden: Efisiensi sistem penerjemah Bahasa Isyarat

Pernyataan	S	TS
Aplikasi dapat membantu komunikasi antara pengguna Bahasa Isyarat dan Bahasa Indonesia		
Aplikasi mudah digunakan		
Aplikasi mudah dipahami dan mempunyai tampilan yang menarik		

Keterangan:

S = Setuju

TS = Tidak Setuju

Setelah dilakukan pengujian pada sistem, tahap terakhir yaitu dilakukan analisis dan pengkajian secara lebih mendalam terhadap sistem yang telah dibuat untuk kemudian dapat ditarik kesimpulan berdasarkan rumusan masalah dan tujuan masalah.

3.2. Perangkat Penunjang Penelitian

Perangkat penunjang yang digunakan dalam penelitian ini dibagi menjadi dua yaitu perangkat keras (*hardware*) dan perangkat lunak (*software*). Perangkat keras yang digunakan terdiri dari Laptop, *Handphone*, dan Webcam (*optional*). Adapun perangkat lunak yang digunakan terdiri dari Visual Studio Code, Python, Javascript, OpenAI, dan Google Cloud.

Spesifikasi alat yang digunakan untuk membuat sistem dalam penelitian ini yaitu Laptop dengan spesifikasi prosesor AMD Ryzen 3 7320U, RAM 8 GB SSD 256. *Handphone* yang digunakan yaitu Xiaomi Redmi 9C dengan prosesor MediaTek Helio G35, Memori 3GB/32GB, Kamera Utama 13MP & Kamera Depan 5MP. Dan spesifikasi alat yang digunakan dalam pengujian sistem yaitu

Fanisa Nur Indah Sari, 2023

ALAT PENERJEMAH BAHASA ISYARAT BERBASIS MACHINE LEARNING UNTUK KOMUNIKASI DUA ARAH BAGI PENYANDANG DISABILITAS SENSORIK (TUNARUNGU DAN TUNAWICARA)

Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

menggunakan Laptop dengan prosesor Intel i3 6006U, RAM 8 GB, dan Hardisk 1 TB, serta menggunakan webcam dengan merek Xiaovv USB Webcam Web Camera 1080p.

3.3. Perancangan Sistem

Dalam perancangan sistem penerjemah bahasa isyarat ini, terdapat tiga tahapan utama yang meliputi kebutuhan, spesifikasi, dan desain dari sistem yang akan dibuat.

3.3.1 Kebutuhan

Dalam tahap ini melibatkan identifikasi terkait dengan kebutuhan sistem penerjemah bahasa isyarat. Kebutuhan utama dalam sistem ini yaitu mencakup:

- Sistem dapat berfungsi sebagai penerjemah bahasa isyarat untuk mendukung komunikasi dua arah antara penyandang tunarungu dengan orang tanpa kebutuhan khusus.
- Sistem dapat dipahami dengan mudah oleh pengguna.
- Sistem dapat diakses dengan mudah oleh pengguna.

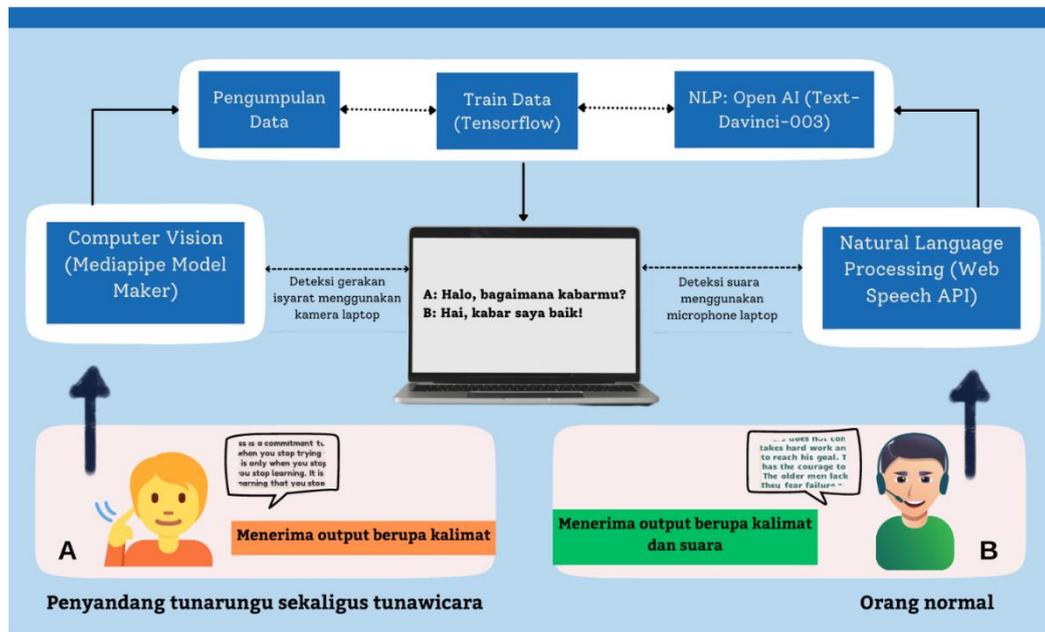
3.3.2 Spesifikasi

Berdasarkan kebutuhan yang telah diidentifikasi sebelumnya, spesifikasi sistem yang akan ditentukan dalam penelitian ini yaitu menggunakan platform *website* sebagai tampilan penggunaannya. Platform *website* dipilih karena memberikan kemudahan aksesibilitas, dapat diakses melalui berbagai perangkat dengan koneksi internet, dan memungkinkan pengguna untuk berinteraksi dengan sistem penerjemah bahasa isyarat secara efisien. Spesifikasi selanjutnya yang diharapkan dari sistem ini adalah mencapai target akurasi deteksi gerakan isyarat sebesar 92% dengan waktu respon sistem tidak lebih dari 1 detik.

3.3.3 Desain

Dalam tahap ini, dilakukan perancangan sistem penerjemah bahasa isyarat sesuai dengan kebutuhan dan spesifikasi yang telah ditentukan sebelumnya, Tahap desain ini memanfaatkan penggunaan teknologi *Computer Vision* (Mediapipe

Model Maker) dan *Natural Language Processing* (OpenAI: Text-Davinci-003 dan Web Speech API). Desain sistem secara keseluruhan dapat ditunjukkan pada Gambar 3.2.



Gambar 3. 2 - Desain sistem

Gambar 3.2 merupakan ilustrasi desain sistem beserta proses teknologi yang digunakannya ketika sistem sedang berjalan. Pengguna A merupakan penyandang tunarungu-tunawicara yang berkomunikasi melalui gerakan isyaratnya, sedangkan pengguna B merupakan orang normal yang berkomunikasi melalui bahasa lisannya. Ketika mereka berinteraksi satu sama lain, proses teknologi yang terjadi yaitu dimulai dari proses *gesture recognition* menggunakan teknologi *Computer Vision* (Mediapipe Model Maker) memungkinkan pengguna (penyandang tunarungu-tunawicara) untuk melakukan deteksi gerakan isyarat, dan selanjutnya untuk lawan bicaranya (orang normal) melalui proses *speech recognition* menggunakan teknologi *Natural Language Processing* (Web Speech API) memungkinkan untuk melakukan deteksi suara. Dari teknologi yang digunakan tersebut, terdapat proses yang harus dilalui dalam perancangan sistem ini, yaitu dimulai dari pengumpulan data yang akan digunakan untuk melatih model pengenalan gerakan isyarat menggunakan teknologi Mediapipe Model Maker. Selanjutnya, model akan dilatih menggunakan teknologi Tensorflow untuk mengenali dan mengklasifikasikan

gerakan isyaratnya, Dan langkah berikutnya adalah mengintegrasikannya dengan teknologi OpenAI. Teknologi OpenAI, seperti Text-Davinci-003, akan digunakan untuk mengubah output gerakan isyarat yang dikenali menjadi kalimat yang lebih lengkap dan bermakna.

3.4. Prinsip Kerja Sistem

Prinsip kerja dari sistem penerjemah bahasa isyarat yang akan diujikan dalam penelitian ini ditunjukkan pada Gambar 3.3.



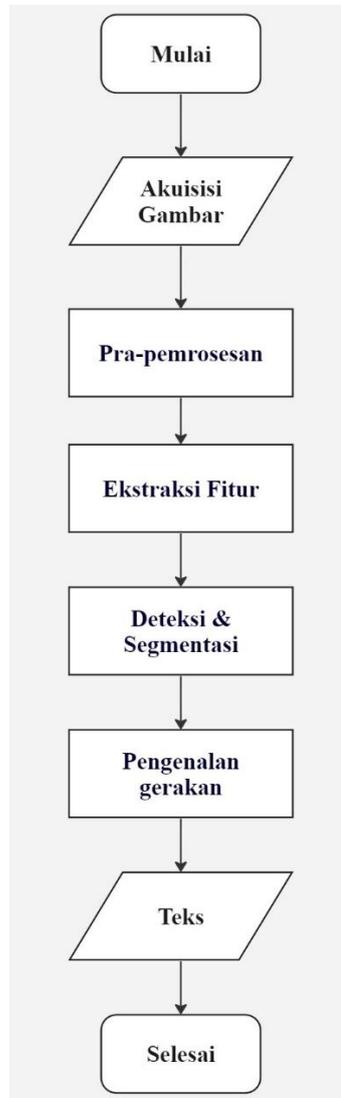
Gambar 3. 3 - Prinsip kerja sistem

Pada saat sistem digunakan, tampilan pengguna akan muncul dengan pilihan *input* berupa *gesture recognition* untuk deteksi gerakan dan *speech recognition* untuk deteksi suara. Pengguna dapat memilih salah satu dari dua opsi, yaitu menggunakan *gesture recognition* dengan menunjukkan isyarat tangan atau menggunakan *speech recognition* untuk mendeteksi suara dengan menggunakan mikrofon. Setelah memilih jenis *input*, sistem akan melakukan kompilasi atau pemrosesan awal terhadap data masukan yang diterima. Jika input yang dipilih adalah deteksi isyarat, sistem akan mengidentifikasi dan mengekstrak gerakan isyarat dari gambar atau video yang diproses. Jika input yang dipilih adalah deteksi suara, sistem akan menerjemahkan suara menjadi teks menggunakan teknologi pengenalan suara. Selanjutnya sistem akan mendokumentasikan hasil deteksi. Jika input yang dipilih adalah deteksi isyarat, sistem akan menghasilkan teks atau kalimat yang mewakili arti dari gerakan isyarat tersebut. Jika input yang dipilih adalah deteksi suara, sistem akan menyimpan hasil transkripsi teks dari suara yang direkam. Hasil deteksi ini dapat digunakan untuk ditampilkan kepada pengguna untuk mengetahui informasi dari input yang diberikan.

Fanisa Nur Indah Sari, 2023

ALAT PENERJEMAH BAHASA ISYARAT BERBASIS MACHINE LEARNING UNTUK KOMUNIKASI DUA ARAH BAGI PENYANDANG DISABILITAS SENSORIK (TUNARUNGU DAN TUNAWICARA)
Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

Dalam implementasinya, teknologi yang digunakan dalam *gesture recognition* dan *speech recognition* memiliki proses yang berbeda dalam melakukan fungsi kerjanya masing-masing. Proses kerja dari *gesture recognition* ditunjukkan pada Gambar 3.4.



Gambar 3. 4 – Diagram alir tahapan metode CV: Gesture Recognition

1. Akuisisi Gambar (*Image Acquisition*)

Tahap ini melibatkan pengambilan citra atau video melalui perangkat kamera dalam Laptop sebagai input awal untuk diproses dalam sistem *Computer Vision*. Tahapan yang dilakukan dalam *image acquisition* melibatkan operasi-operasi seperti memilih sumber citra (seperti kamera), mengatur parameter kamera

(*optional*), mengambil citra gerakan tangan, melakukan pemrosesan awal citra (seperti perubahan ukuran atau konversi ke skala abu-abu), dan menyimpan atau menyampaikan citra untuk diproses lebih lanjut dalam sistem penerjemah bahasa isyarat.

2. Pra-pemrosesan (*Pre-processing*)

Citra yang diambil melalui tahap *image acquisition* kemudian diproses dalam pre-processing untuk memperbaiki kualitas citra. Tahap pra-pemrosesan melibatkan serangkaian operasi untuk mempersiapkan citra gerakan tangan sebelum diproses lebih lanjut. Tahapan *Pre-processing* dalam *Computer Vision* melibatkan serangkaian operasi seperti perubahan ukuran (*resize*), pengurangan noise (*denoising*), pengkonversian ke citra grayscale (*grayscale*), peningkatan kontras (*contrast enhancement*), dan normalisasi (*normalization*) untuk mempersiapkan citra sebelum analisis lebih lanjut.

3. Ekstraksi Fitur (*Feature Extraction*)

Pada tahap ini, fitur-fitur penting dari gerakan tangan diekstraksi dari citra yang telah di *pre-processing*. Fitur-fitur ini dapat mencakup pose tangan, bentuk tangan, gerakan tangan, *landmark* tangan, atau pola gerakan yang spesifik. Tahapan *Feature Extraction* melibatkan operasi-operasi seperti ekstraksi tepi (*edge extraction*) untuk mengidentifikasi kontur tangan, deteksi fitur spesifik untuk menemukan pola gerakan tangan yang khas, ekstraksi deskriptor untuk merepresentasikan fitur-fitur penting dalam citra tangan, transformasi domain frekuensi untuk mengungkapkan karakteristik frekuensi dalam citra, dan ekstraksi fitur berbasis *deep learning* menggunakan jaringan saraf tiruan untuk secara otomatis mengekstraksi fitur-fitur yang relevan. Ekstraksi fitur bertujuan untuk menggambarkan dan merepresentasikan gerakan tangan dalam bentuk fitur yang dapat digunakan untuk membedakan berbagai jenis *gesture*.

4. Deteksi & Segmentasi (*Detection & Segmentation*)

Tahap ini melibatkan deteksi dan segmentasi gerakan tangan dalam citra. Deteksi bertujuan untuk mengidentifikasi keberadaan gerakan tangan dalam citra secara keseluruhan. Kemudian, segmentasi digunakan untuk memisahkan tangan dari latar belakang atau objek lain dalam citra, untuk memfokuskan analisis pada gerakan tangan. Tahapan deteksi dan segmentasi melibatkan

operasi-operasi seperti deteksi objek (*object detection*) untuk mengidentifikasi tangan, segmentasi piksel (*pixel segmentation*) untuk memisahkan tangan dari latar belakang, segmentasi berdasarkan instans (*instance segmentation*) untuk memisahkan tangan individu, segmentasi berdasarkan semantik (*semantic segmentation*) untuk memahami konten visual dalam konteks yang lebih luas, dan segmentasi berdasarkan kontur (*contour segmentation*) untuk ekstraksi kontur atau tepi tangan yang signifikan.

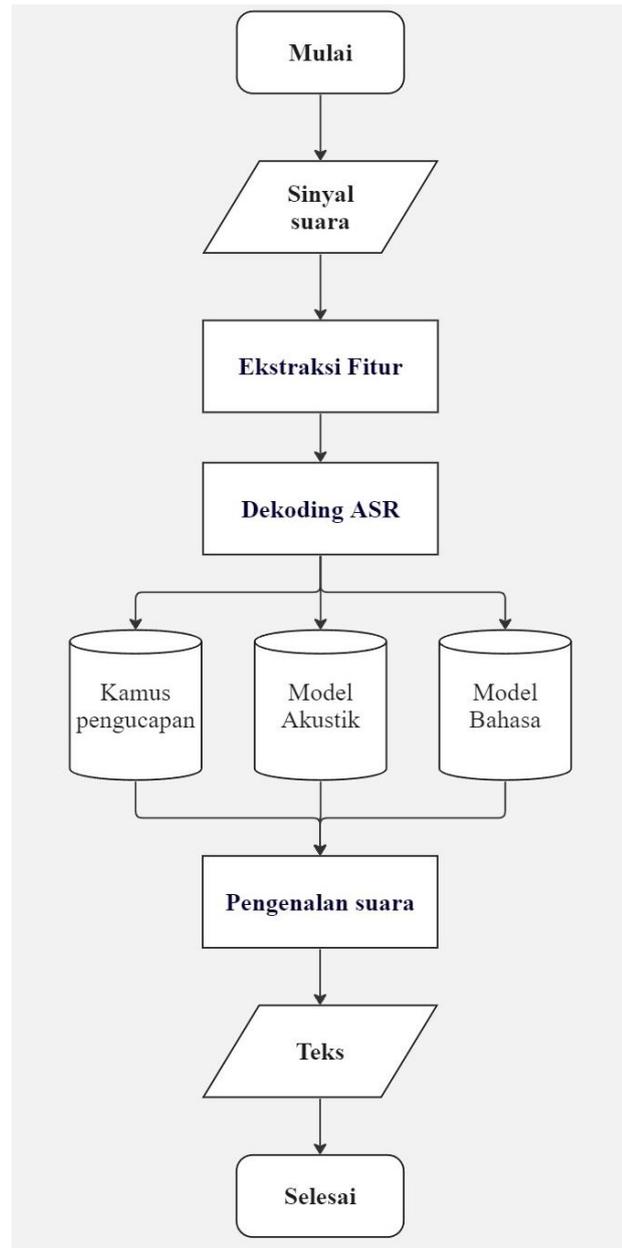
5. Pengenalan Gerakan (*Gesture Recognition*)

Pada tahap ini, fitur-fitur yang diekstraksi dan gerakan tangan yang telah dideteksi dan di-segmentasi digunakan untuk mengenali dan menginterpretasikan gesture yang dilakukan. Pengenalan gesture melibatkan perbandingan fitur-fitur yang diekstraksi dengan model atau database gesture yang telah dilatih sebelumnya. Setelah pengenalan gesture, tahap interpretasi dilakukan untuk mengartikan gesture tersebut dalam bentuk kata untuk selanjutnya dijadikan kalimat.

6. Teks

Hasil dari proses pengenalan gerakan tangan adalah teks yang berisi interpretasi dari gerakan tangan yang telah dikenali. Teks ini berupa kata yang mewakili makna atau arti dari gerakan tangan tersebut, dan bisa digunakan untuk berkomunikasi dengan pengguna melalui bahasa manusia.

Selanjutnya, proses kerja dari *speech recognition* ditunjukkan pada Gambar 3.5. *Speech recognition* merupakan teknologi yang dapat mengubah ucapan manusia menjadi teks tertulis melalui proses perekaman suara, pra-pemrosesan, ekstraksi fitur, dan analisis menggunakan model akustik dan model bahasa, menghasilkan output berupa teks transkripsi dari ucapan asli.



Gambar 3. 5 - Diagram alir tahapan metode NLP: Speech Recognition

- 1) Sinyal Suara (*Speech Signal*): Pada tahap ini, suara atau ucapan manusia direkam menggunakan mikrofon atau sumber suara lainnya. Ucapan ini kemudian diubah menjadi sinyal suara digital yang dapat diproses oleh komputer.
- 2) Ekstraksi Fitur (*Feature Extraction*): Pada tahap ini, fitur-fitur akustik ekstraksi dilakukan pada sinyal suara digital yang diperoleh sebelumnya. Fitur-fitur ini

dapat melibatkan teknik seperti *Mel-Frequency Cepstral Coefficients* (MFCC), *Linear Predictive Coding* (LPC), atau *Filter Bank*. Tujuannya adalah untuk mengidentifikasi dan mengekstrak pola-pola penting dari sinyal suara yang dapat digunakan untuk pengenalan suara.

- 3) Dekoding ASR (*ASR Decoding*): Tahap ini melibatkan proses *decoding* dalam sistem *Automatic Speech Recognition* (ASR). ASR decoding terdiri dari tiga komponen utama:
 - a. Kamus pengucapan (*Pronunciation dictionary*): Ini adalah kamus yang berisi kumpulan kata dan transkripsi fonetik mereka. Kamus ini menyediakan informasi tentang bagaimana kata-kata tersebut diucapkan secara fonetik.
 - b. Model akustik (*Acoustic model*): Model ini menghubungkan fitur-fitur ekstraksi yang diperoleh pada tahap sebelumnya dengan representasi fonetik dalam kamus. *Acoustic model* ini digunakan untuk menentukan probabilitas bahwa suatu suara tertentu merupakan representasi fonetik dari suatu kata.
 - c. Model bahasa (*Language model*): Model ini menyediakan informasi statistik tentang urutan kata yang mungkin dalam bahasa yang dikenali. *Language model* membantu dalam mengenali urutan kata yang paling mungkin berdasarkan konteks dan kemungkinan kemunculan kata-kata.
- 4) Pengenalan Suara (*Speech Recognition*): Tahap akhir adalah proses pengenalan suara yang sebenarnya. Dalam tahap ini, probabilitas yang dihitung dari *acoustic model* dan *language model* digunakan untuk menghasilkan teks yang merupakan transkripsi dari ucapan yang dikenali.
- 5) Teks: Setelah proses pengenalan suara selesai, hasilnya berupa teks yang merupakan transkripsi dari ucapan yang dikenali. Teks ini dituliskan sebagai output dari sistem ASR.