

BAB I

PENDAHULUAN

1.1. Latar Belakang

Komunikasi merupakan suatu proses penyampaian informasi dari satu pihak kepada pihak lain (Sholawati, 2022). Komunikasi memungkinkan manusia membangun suatu kerangka rujukan untuk menafsirkan situasi apapun yang mereka hadapi. Namun lain halnya dengan orang yang terlahir sebagai penyandang tunarungu. Penyandang tunarungu dan tunawicara berkomunikasi sehari-hari menggunakan bahasa isyarat. Bahasa isyarat yakni bahasa yang tidak menggunakan bunyi ucapan manusia atau tulisan dalam penerapannya (Pradikja, 2018). Bahasa Isyarat dilakukan dengan menggunakan gerakan-gerakan badan dan mimik muka sebagai simbol dari makna Bahasa. Pada tahun 2021 menurut data penyandang disabilitas tunarungu dan tunawicara dari KEMENSOS (Kementerian Sosial Republik Indonesia), jumlah penyandang tunarungu sebanyak 13.800 dan tunawicara sebanyak 5.580 orang. Bahasa isyarat termasuk bahasa yang unik, karena berbeda ditiap negara. Di Indonesia, terdapat dua standar bahasa isyarat, yaitu SIBI (Sistem Bahasa Isyarat Indonesia) dan BISINDO (Bahasa Isyarat Indonesia). BISINDO adalah bahasa yang digalakkan oleh Gerakan Kesejahteraan Tunarungu Indonesia (GERKATIN) dan dikembangkan oleh masyarakat tunarungu sendiri.

Pendidikan Sekolah Luar Biasa (SLB) memiliki peran yang sangat penting dalam mendukung perkembangan pembelajaran anak-anak penyandang tuna rungu. Melalui bahasa isyarat sebagian besar digunakan oleh para penyandang disabilitas, dan hanya sedikit hanya sedikit orang lain yang memahaminya, seperti keluarga, aktivis, dan guru di Sekolah Luar Biasa (SLB) (Pathak et al., 2022). Oleh karena itu dibutuhkan suatu metode yang dapat membaca gerakan isyarat tangan, tubuh, dan wajah untuk diterjemahkan sebagai teks dan suara, sehingga bisa berguna sebagai salah satu media yang menunjang pembelajaran anak-anak pada Sekolah Luar Biasa (SLB) untuk media pembelajaran dalam mengolah kosa kata dan

berkomunikasi dengan masyarakat lainnya. Perkembangan dunia teknologi era 5.0 telah menciptakan berbagai inovasi yang maju dan efektif di bidang *computer vision*. Penggunaan *Machine Learning* dan *Deep Learning* berperan besar dalam pengenalan dan pembuatan Pendeteksian Bahasa Isyarat Indonesia.

Setelah mengetahui permasalahan yang ada dan sebagai upaya untuk menunjang pembelajaran anak-anak pada SLB serta meningkatkan akurasi dari penelitian sebelumnya. Pada penelitian yang dilakukan (Damatraseta, 2021) menggunakan algoritma *Convolutional Neural Network* (CNN) dengan masing-masing arsitekturnya yaitu *LeNet-5* dan *Alexnet*. Meski dapat berjalan secara realtime, hasil akurasi training yang didapatkan hanya sebesar 59.54 persen pada model data *Alexnet*, sedangkan *LeNet* hanya 19 persen. Pada pengembangan berikutnya dilakukan oleh (Pathak, 2021) juga membuat sistem dapat berjalan secara realtime menggunakan algoritma yang sama yaitu *Convolutional Neural Network* (CNN) dengan model arsitekturnya *SSD Mobilenet-V2*. Hasil yang didapatkan pada proses training memperoleh nilai akurasi 70-80 persen. Hasil akurasi yang didapatkan cukup tinggi, namun dataset yang digunakan masih berupa citra gambar sedangkan sebagian dari bahasa isyarat menggunakan gerakan yang dinamis serta latar belakang yang digunakan masih tidak terkendali sehingga menyebabkan penurunan akurasi pendeteksian. Oleh sebab itu, pada penelitian ini mengangkat algoritma lain yaitu, *LSTM (Long Short-Term Memory)* yang berbasis *RNN* dan dapat mendeteksi secara *realtime* dengan tensorflow (Auliya, 2022). *LSTM network* menjadi pilihan terbaik berkat kemampuannya untuk menyimpan ingatan dalam waktu yang lama pada saat yang sama, dan korelasi yang kompleks antara data memberikan informasi yang sangat berguna. Dalam penelitian “Penerapan Convolutional Long Short-Term Memory Untuk Klasifikasi Teks Berita Bahasa Indonesia (2021)” menunjukkan hasil dimana metode LSTM memiliki akurasi lebih baik untuk mengklasifikasi hasil teks berita yaitu sebesar 90,87% dibandingkan dengan metode CNN dengan akurasi sebesar 89,85%. LSTM dianggap lebih baik daripada RNN dan ARIMA dalam mempelajari

ketergantungan jangka panjang dalam memberikan keseimbangan antara akurasi model dan prekabilitas model (Yifan Yao, 2022).

1.2. Rumusan Masalah

Berdasarkan latar belakang di atas, maka rumusan masalah yang akan dibahas sebagai berikut :

1. Bagaimana pengaruh penggunaan MediaPipe Holistics dan LSTM dalam mendeteksi gerakan BISINDO menjadi kosa kata?
2. Bagaimana penerapan fitur TTS dalam deteksi gerakan BISINDO sehingga respons suara yang dihasilkan sesuai dengan jenis gerakan yang dikenali?
3. Bagaimana akurasi, presisi, recall, dan F1 score setelah menggunakan Long Short-Term Memory dan MediaPipe Holistics?

1.3. Tujuan

Berdasarkan latar belakang dan rumusan masalah diatas, maka diperoleh tujuan penelitian ini sebagai berikut :

1. Menganalisis pengaruh Mediapipe Holistics dan LSTM dalam untuk menerjemahkan gerakan BISINDO menjadi kosa kata.
2. Mengimplementasikan gTTS untuk fitur Text-to-Speech agar output dari Teks bisa menjadi Suara.
3. Mengetahui performa akurasi yang didapatkan pada deteksi gerakan BISINDO menggunakan Long Short-Term Memory dan MediaPipe Holistics.

1.4. Manfaat

Adapun manfaat yang dapat diambil dari penelitian ini sebagai berikut :

1. Bagi pengguna yaitu sebagai media pembelajaran anak-anak pada Sekolah Luar Biasa (SLB) untuk mengolah kata-kata menjadi sebuah kosa kata menjadi kalimat.
2. Bagi penulis yaitu menambah portofolio dan pengetahuan serta penerapan mengenai deteksi gerakan BISINDO menggunakan Long Short-Term Memory dan MediaPipe Holistics.
3. Bagi Mahasiswa yaitu bisa menjadi referensi untuk pengembangan

mengenai deteksi gerakan BISINDO.

1.5. Batasan Masalah

Pada penelitian ini, batasan masalah yang dibahas adalah sebagai berikut :

1. Penelitian hanya terbatas pada pengujian laboratorium (tidak dikembangkan sebagai aplikasi mobile, web, dekstop ataupun aplikasi multiplatform).
2. Pengujian secara realtime menggunakan webcam pada device.
3. Algoritma yang digunakan adalah *Long Short Term Memory*.
4. Mendeteksi landmark tangan, wajah dan pose tubuh menggunakan MediaPipe Holistics.
5. Ouput kalimat dibatasi dengan 3-4 kalimat.
6. Pembuatan dataset dengan setengah badan serta dengan background statis.
7. Standar acuan menggunakan BISINDO dengan 35 gerakan untuk dataset yaitu : saya, kamu, ayah, ibu, kakak, adik, teman, senang, sedih, marah, takut, membantu, cinta, minta maaf, terimakasih, bermain, belajar, tolong, rumah, sekolah, taman, kelas, dan, dengan, dari.
8. Penggunaan gTTs bergantung pada koneksi internet yang stabil.