

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Bahasa isyarat merupakan sarana komunikasi bagi teman tuli dengan banyak variasi di pada daerahnya, di Indonesia dimana BISINDO (Bahasa Isyarat Indonesia) menjadi bentuk paling banyak digunakan untuk berkomunikasi dalam sehari-hari, pembelajaran dan kegiatan sosial. Menurut data dari BPS (Badan Pusat Statistik) 2022, sekitar 933.893 jiwa penyandang disabilitas di Indonesia dengan perkiraan lebih dari 158.918 jiwa diantara memiliki keterbatasan penglihatan, pendengaran dan berbicara [1]. Dari jumlah itu, sebagian diantaranya sudah tuli sejak lahir atau usia dini, saat sedang dalam masa pertumbuhannya mengalami keterbatasan atas akses pendidikan formal. Penelitian menunjukkan bahwa kondisi ini berdampak langsung terhadap rendahnya literasi bahasa formal pada teman tuli karena untuk belajar membaca dan menulis sendiri membutuhkan keterampilan mendengar dan berbahasa untuk memadainya. Pada studi menunjukkan rata-rata penyandang tuli yang menggunakan bahasa isyarat sebagai bahasa utama mempunyai pemahaman bacaan yang setara dengan anak berusia 8-9 tahun, bahkan pada saat usia remaja/dewasa [2].

Dari sudut pandang produsen konten edukasi seperti editor video dan lembaga penyedia materi pembelajaran. Dari sudut pandang produsen konten edukasi, salah satunya JBI yang bekerja berdampingan dengan editor video bahasa isyarat yang terlibat langsung dalam produksi konten inklusif, keterbatasan JBI serta tingginya kebutuhan waktu dalam proses penerjemahan manual menjadi hambatan utama dalam produksi konten berbahasa isyarat [3]. Berdasarkan wawancara dengan bapak Stefanus Ming, seorang JBI yang bekerja berdampingan dengan editor video yang berpengalaman dalam produksi konten edukasi berbahasa isyarat, proses penerjemahan dan perekaman bahasa isyarat merupakan tahapan yang paling memakan waktu dalam keseluruhan alur produksi.

Keterbatasan sumber daya JBI dan tantangan pemahaman *subtitle* menciptakan bottleneck besar dalam produksi konten edukasi inklusif. Proses manual pembuatan video terjemahan bahasa isyarat yang melibatkan beberapa tahapan seperti penerjemahan, perekaman ulang adegan isyarat, sinkronisasi gerakan dan penyuntingan akhir. Tahap penerjemahan dan perekaman JBI yang membutuhkan

kehadiran secara langsung adalah komponen yang paling memakan waktu. Berdasarkan wawancara informal dan pengamatan terhadap proses produksi konten inklusif, tahap penerjemahan dan perekaman bahasa isyarat diketahui membutuhkan waktu yang signifikan dibandingkan durasi video yang dihasilkan. Kondisi ini menjadikan proses manual sebagai bottleneck utama dalam produksi konten edukasi berbahasa isyarat. [4]

Meskipun pemerintah dan lembaga pendidikan telah memproduksi berbagai macam video pembelajaran, kesenjangan ini masih menjadi masalah krusial karena penyertaan JBI yang masih minim[5]. Lebih dari 70% teman tuli mengaku kesulitan memahami informasi jika hanya mengandalkan subtitle tanpa dukungan bahasa isyarat[6][7]. Hal ini menuntut adanya solusi teknologi yang dapat menjembatani kesenjangan tersebut dengan menjamin efisiensi.

Penelitian ini bertujuan untuk mengembangkan sebuah sistem yang bisa mengkonversi audio/teks ke dalam gestur BISINDO. Sistem ini difokuskan pada konteks non-real-time untuk konten video yang mempunyai transkrip teks dan dioptimalkan untuk meningkatkan efisiensi waktu produksi melalui sintesis gestur dengan teknik penggabungan template, serta mengimplementasikan mekanisme *fingerspelling* sebagai *fallback*. Sistem otomatis menghilangkan seluruh proses manual yang paling memakan waktu seperti perekaman, cutting dan editing awal video. Dari sudut pandang editor video di bidang ini, output sistem sama fungsionalnya dengan hasil rekaman JBI manual, namun waktu produksi JBI dan editor dieliminasi. Kemudahan ini memberikan potensi bantuan yang signifikan kepada editor, karena mereka hanya perlu membuang greenscreen yang ada pada output video dan menempelkannya ke video utama. Melalui sistem ini, diharapkan proses produksi konten dapat dipersingkat secara signifikan, mengurangi ketergantungan pada interpreter manual dan pada akhirnya memperluas akses informasi yang mudah dipahami bagi teman tuli.

## **1.2 Identifikasi Masalah**

**1.2.1** Apakah sistem *Text-to-Gesture* yang diimplementasikan dapat menunjukkan efisiensi waktu produksi dibandingkan metode manual?

## **1.3 Batasan Penelitian**

Batasan dari penelitian ini sebagai berikut:

- 1.3.1. Penelitian ini berfokus pada pengembangan sistem *audio/text-to-gesture* untuk konteks *non-real-time*.
- 1.3.2. Pengembangan sistem ini terbatas pada kosakata BISINDO sebanyak 91 kata pada dataset. Oleh karena itu, sistem hanya bisa me-terjemahkan kata yang tersedia dan belum mendukung struktur kalimat yang kompleks.
- 1.3.3. Penelitian ini mencakup pengembangan model konversi dan pembuatan *prototype* animasi. Integrasi ke platform secara penuh belum termasuk dalam ruang lingkup penelitian.
- 1.3.4. Visualisasi gerakan terbatas pada representasi karakter 2D *mediapipe* yang disintesis tanpa melibatkan proses *Inverse Kinematics* atau konversi ke format animasi 3D seperti BVH.
- 1.3.5. Validasi sistem dilakukan melalui analisis kuantitatif efisiensi waktu *end-to-end* sistem dan uji kepuasan pengguna untuk mengukur keterpahaman dan sinkronisasi.
- 1.3.6. Penelitian ini tidak mencakup pemodelan ekspresi wajah seperti mimik yang punya peran linguistik dalam BISINDO.
- 1.3.7. Sistem yang dikembangkan tidak melakukan pemotongan frame atau segmentasi gestur berdasarkan timestamp audio. Penggabungan gestur dilakukan secara sekuensial berdasarkan pemetaan kata ke *template* gestur, tanpa mekanisme alignment temporal terhadap sinyal audio.

#### **1.4 Tujuan Penelitian**

- 1.4.1 Tujuan dari penelitian ini untuk mengukur efisiensi waktu produksi video terjemahan BISINDO yang dihasilkan oleh sistem secara otomatis dibandingkan dengan metode produksi manual.
- 1.4.2 Memvalidasi tingkat keterpahaman dan sinkronisasi gestur yang dihasilkan melalui kuesioner kualitatif.

#### **1.5 Manfaat Penelitian**

Penelitian ini memiliki beberapa manfaat sebagai berikut:

- 1.5.1 Memberikan akses konten yang lebih mudah dipahami oleh teman tuli dengan literasi bahasa formal rendah melalui visualisasi bahasa isyarat yang sinkron
- 1.5.2 Menyediakan arsitektur sistem yang efisien waktu dan biaya sebagai referensi bagi lembaga pendidikan atau media massa yang ingin meng-

otomatisasi produksi konten aksesibilitas.

## **1.6 Sistemtika Penulisan**

- **Bab I Pendahuluan**, berisi latar belakang masalah, identifikasi masalah, batasan penelitian, tujuan penelitian, manfaat penelitian.
- **Bab II Tinjauan Pustaka**, berisi teori-teori yang mendasari penelitian, kajian terhadap penelitian terdahulu, serta pembahasan mengenai teknologi dan metode yang digunakan.
- **Bab III Metode Penelitian**, berisi uraian mengenai rancangan sistem, metode pengembangan, serta tahapan pengujian model dan sistem yang dilakukan.
- **Bab IV Hasil dan Pembahasan**, memaparkan implementasi sistem, hasil pengujian, serta analisis terhadap kinerja model dan sistem yang dikembangkan.
- **Bab V Penutup**, menyajikan kesimpulan dari penelitian dan saran untuk pengembangan lebih lanjut.