

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Bahasa isyarat merupakan bentuk komunikasi visual yang menggunakan gerakan tangan, ekspresi wajah, dan ataupun postur tubuh untuk menyampaikan sebuah pesan. Bahasa isyarat digunakan oleh komunitas Tuli, dan juga orang-orang yang memiliki kesulitan mendengar ataupun berbicara. Bahasa isyarat bukan merupakan versi visual dari bahasa lisan yang ada, tetapi memiliki tata bahasa dan struktur yang berbeda dengan bahasa lisan. Bahasa isyarat melibatkan pemahaman mengenai sintaksis dan leksikon khusus sesuai dengan jenis bahasa isyarat yang digunakan yang biasanya ditentukan oleh pemerintah dan masyarakat diwilayah tersebut.

Indonesia menggunakan Bisindo atau bahasa isyarat Indonesia dan SIBI atau sistem isyarat bahasa Indonesia. Bisindo merupakan bahasa isyarat yang muncul secara alami dan digunakan dalam kegiatan sehari-hari oleh komunitas tuli dan dipengaruhi oleh konteks sosial, budaya, serta interaksi antar pengguna bahasa. Kosa isyarat pada Bisindo bersifat simbolis terhadap bahasa lisan yang diwakilkan. Bisindo lebih menekankan ekspresi visual dan keunikan komunikasi antar individu. Sedangkan SIBI merupakan sistem bahasa isyarat yang diciptakan untuk melambangkan kosa kata bahasa Indonesia dalam bentuk bahasa isyarat sehingga memiliki struktur yang sama dengan tata bahasa lisan Indonesia, dimana terdapat awalan, akhiran, imbuhan, dan kata benda.

Secara linguistik, Bisindo memiliki karakteristik yang berbeda dari bahasa lisan, antara lain melalui penggunaan sinyal non-manual seperti ekspresi wajah, penerapan tata bahasa ruang (*spatial grammar*), serta penggunaan *classifier* yang berfungsi untuk merepresentasikan objek, gerakan, dan relasi secara visual. Oleh sebab itu, Bisindo tidak dapat dipandang sebagai representasi visual dari bahasa Indonesia, melainkan sebagai bahasa yang memiliki struktur gramatikal dan sistem kebahasaan yang mandiri.

Kurangnya pengajaran bahasa isyarat Indonesia baik secara formal maupun non-formal kepada masyarakat umum menjadi salah satu penyebab komunikasi antara komunitas Tuli dengan masyarakat umum sulit dilakukan. Masyarakat umum hampir tidak pernah mendapatkan pengajaran maupun pelatihan terkait dengan penggunaan bahasa isyarat, baik itu Bisindo maupun SIBI (Onggo, 2025). Selain itu, perbedaan signifikan antara bahasa isyarat dengan bahasa lisan menyebabkan sebagian besar masyarakat kurang memahami ataupun menguasai bahasa isyarat.

Berdasarkan hal tersebut, diperlukan sebuah sistem yang dapat menjembatani komunikasi bahasa isyarat dan bahasa lisan. Pemanfaatan teknologi visi komputer dapat membantu dilakukannya proses transliterasi atau translasi bahasa isyarat ke bahasa lisan atau latinnya dan sebaliknya. Proses translasi bahasa isyarat dapat dilakukan dengan mengenali tiap gerakan isyarat yang dilakukan dan mengidentifikasikan atau mengenali gerakan bahasa isyarat tersebut sesuai dengan bahasa lisan atau latinnya. Hasil identifikasi ini kemudian diolah kembali untuk membentuk kata dan kalimat dalam bentuk bahasa lisannya.

Salah satu penelitian pada bidang ini adalah penelitian yang dilakukan oleh (Zhang et al., 2021) yang melakukan proses rekognisi gerakan tangan dinamis dengan menggunakan metode CNN dan LSTM. Penelitian ini menggunakan video gerakan tangan yang selanjutnya dibagi menjadi beberapa bagian *frame* yang kemudian dipilih satu *frame* sehingga mendapatkan *frame* yang mewakili gerakan tangan pada video tersebut. *Frame* video ini juga mengalami proses *zoomed-out* untuk mengatasi masalah ketika gerakan tangan berada pada bagian terluar dari video. Selain itu, pada penelitian ini ditambahkan hasil *optical flow* yang digabungkan dengan grup *frame* yang telah diperoleh. CNN pada penelitian ini digunakan *layer* convolusinya untuk proses ekstraksi ciri yang kemudian digunakan sebagai input untuk LSTM. Hasil yang didapatkan pada 27 kelas gerakan tangan adalah sebesar 95,73%. Penelitian ini mengalami kendala pada dataset yang digunakan, gerakan tangan para subjek berada dibagian terluar video sehingga membutuhkan penambahan area diluar *frame* video. Selain itu, penambahan *optical flow* juga mempengaruhi biaya dari proses yang dilakukan.

Penelitian lainnya dilakukan oleh (C. A. Sari dkk., 2024) yang melakukan penelitian deteksi bahasa isyarat untuk SIBI secara langsung dengan menggunakan MediaPipe dan LSTM. Penelitian ini menggunakan enam kelas bahasa isyarat dimana digunakan 30 *frame* dari setiap video gerakan bahasa isyarat. MediaPipe digunakan sebagai proses ekstraksi ciri dari gerakan tangan yang ada pada *frame* video. Ciri yang diekstrak adalah *keypoint* dari tangan, wajah, dan pose tubuh. LSTM selanjutnya digunakan untuk proses pelatihan data latih dan prediksi data uji. Hasil pengujian yang dilakukan menghasilkan nilai akurasi sebesar 92%.

Kekurangan pada penelitian ini adalah jumlah kelas yang digunakan sedikit dan belum mencerminkan bahasa isyarat SIBI yang digunakan pada penelitian tersebut. Selain itu, video yang digunakan memiliki *frame* sejumlah 30 untuk tiap gerakan belum dapat mewakili masalah dunia nyata.

Penelitian terkait Bisindo telah dilakukan oleh (Kaffah & Ramdhani, 2023) yang melakukan penelitian dengan melakukan komparasi lima metode/algoritma klasifikasi, yaitu *random forest*, *decision tree*, *naive bayes*, *support vector machine*, dan *k-nearest neighbour*. Dataset yang digunakan merupakan dataset primer yang terdiri dari 10 kosa isyarat. Fitur yang digunakan pada penelitian ini diekstraksi dengan menggunakan MediaPipe. Berdasarkan hasil penelitian yang dilakukan, metode *random forest* memiliki tingkat akurasi tertinggi jika dibandingkan dengan metode lainnya, yaitu sebesar 99,80%. Walaupun memperoleh hasil akurasi yang tinggi, pada penelitian terdapat kekurangan yaitu jumlah kosa isyarat yang masih terbatas. Selain itu, pada penelitian ini tidak dijelaskan *landmark* mana saja yang digunakan sebagai fitur dari gerakan bahasa isyarat.

Beberapa penelitian yang dilakukan telah dapat melakukan proses pengenalan gerakan bahasa isyarat dengan baik. Kendala yang dapat dialami pada proses pengenalan bahasa isyarat adalah perbedaan variasi gerakan masing-masing individu dan juga lama gerakan dilakukan. Gerakan untuk kosa kata bahasa isyarat bervariasi tidak hanya pada jenis atau bentuk gerakan, namun juga pada durasi gerakan itu dilakukan. Variasi pada gerakan ini menjadi kendala ketika proses pengenalan dilakukan dengan pengambilan *frame* berdasarkan jumlah tertentu ataupun pengambilan *frame* yang tidak mewakili gerakan secara penuh. Selain itu,

dalam kondisi dunia nyata, pencahayaan dan posisi subjek mempengaruhi proses identifikasi gerakan bahasa isyarat yang dilakukan.

Permasalahan lainnya adalah keterbatasan *dataset* yang digunakan. Domain masalah rekognisi kata pada bahasa isyarat memerlukan penelitian pada *open vocabulary* atau tidak terbatas dengan kelas yang digunakan. Penelitian-penelitian yang telah dilakukan terbatas pada jumlah data atau kosa isyarat dan terbatas pada masalah klasifikasi dengan percobaan dan/atau pengujian pada metode tertentu. Selain itu, beberapa penelitian pada Bisindo seperti penelitian oleh (Arisandi & Satya, 2022) masih terbatas pada rekognisi gerakan pada tingkat huruf (A-Z). Sedangkan, dalam komunikasi sehari-hari, bahasa isyarat yang digunakan berada pada tingkat kata ataupun kalimat.

Berdasarkan permasalahan ini, diusulkan penelitian dengan judul “*Rekognisi Kata pada Gerakan Bahasa Isyarat BISINDO dengan Metode Long Short-Term Memory*” dimana digunakan fitur *landmark* atau sendi yang diekstrak atau diambil dengan menggunakan *framework* MediaPipe dan metode *long short-term memory* sebagai metode prediksi gerakan. Dataset yang digunakan dibuat dari video untuk kosa isyarat bahasa isyarat Indonesia. Proses pelatihan untuk model rekognisi dimulai dari tahapan pengambilan *frame* yang dapat mewakili keseluruhan gerakan. Proses pengambilan *frame* dari video gerakan akan dilakukan dengan mengambil beberapa *sampling frame* dengan jumlah yang telah ditentukan berdasarkan urutan waktu sehingga diharapkan dapat mewakili gerakan secara utuh. *Frame* yang telah didapatkan kemudian dilakukan proses ekstraksi ciri dengan menggunakan *framework* MediaPipe sehingga mendapatkan nilai-nilai koordinat dari *landmark*

tubuh. Proses pelatihan dilanjutkan dengan menggunakan metode LSTM untuk membentuk model rekognisi atau prediksi gerakan. Model ini selanjutnya diuji dengan data uji yang telah disiapkan. Dengan dilakukannya penelitian ini, diharapkan hasil yang diperoleh adalah *label* dari kosa isyarat bahasa isyarat yang tepat.

1.2 Identifikasi Masalah

Berdasarkan uraian latar belakang di atas, penulis mengidentifikasi beberapa masalah dalam penelitian ini sebagai berikut.

1. Komunikasi yang sulit dilakukan dengan komunitas tuli akibat kurangnya pemahaman masyarakat umum terkait bahasa isyarat. Oleh sebab itu, perlu dibuat model pengenalan kata berdasarkan gerakan bahasa isyarat.
2. Penelitian berkaitan dengan bahasa isyarat masih berupa klasifikasi dengan jumlah kata atau kelas yang terbatas.

1.3 Pembatasan Masalah

Agar penelitian yang dilakukan terarah dan tidak meluas, maka penulis membuat batasan masalah sebagai berikut.

1. Kosa isyarat yang digunakan adalah kosa isyarat dalam bentuk kata.
2. Dataset berupa video gerakan bahasa isyarat Indonesia yang diambil dari bagian pinggang ke atas dengan peraga menghadap ke depan.
3. Data latih yang digunakan adalah video gerakan bahasa isyarat Indonesia yang dilakukan oleh 3 orang peraga.

4. Data uji yang digunakan adalah video gerakan bahasa isyarat Indonesia yang dilakukan oleh 1 orang peraga.
5. Fitur yang digunakan adalah *landmark* dari tubuh dan jari kedua tangan dari peraga.
6. Proses rekognisi dilakukan pada gerakan bahasa isyarat berbasis kata.
7. Aplikasi rekognisi yang dihasilkan melakukan rekognisi untuk video dengan satu sekuens gerakan dan melakukan rekognisi dari gerakan menjadi kata dan tidak sebaliknya.

1.4 Rumusan Masalah

Berdasarkan latar belakang di atas, maka dapat dirumuskan permasalahan yang akan diangkat dalam penelitian ini sebagai berikut.

1. Bagaimana hasil pengenalan rekognisi kata pada gerakan bahasa isyarat Indonesia dengan metode *Long Short-Term Memory*?
2. Bagaimana rancangan dan implementasi prototipe sistem rekognisi kata pada gerakan bahasa isyarat Indonesia dengan metode *Long Short-Term Memory*?

1.5 Tujuan Penelitian

Adapun tujuan dilakukannya penelitian ini adalah sebagai berikut.

1. Mengetahui hasil pengenalan rekognisi kata pada gerakan bahasa isyarat Indonesia dengan metode *Long Short-Term Memory*.

2. Menciptakan rancangan dan implementasi prototipe sistem rekognisi kata pada gerakan bahasa isyarat Indonesia dengan metode *Long Short-Term Memory*

1.6 Manfaat Penelitian

Hasil dari penelitian ini diharapkan dapat memberikan manfaat baik secara teoritis maupun praktis sebagai berikut.

1. Manfaat teoritis

Manfaat teoritis dari penelitian ini adalah mengetahui dan mengevaluasi kinerja metode LSTM dalam memprediksi gerakan bahasa isyarat Indonesia. Hasil penelitian ini juga diharapkan dapat memberikan manfaat kepada pembaca ataupun sebagai acuan referensi bagi penelitian selanjutnya dalam pengembangan ilmu pengetahuan yang berkaitan dengan rekognisi ataupun klasifikasi gerakan bahasa isyarat.

2. Manfaat praktis

Manfaat praktis yang dihasilkan oleh penelitian ini diharapkan dapat memberikan manfaat bagi pihak-pihak sebagai berikut.

- a. Bagi masyarakat umum

Manfaat yang diperoleh bagi masyarakat umum yaitu hasil dari penelitian ini dapat digunakan untuk mempermudah komunikasi dengan komunitas tuli. Selain itu, penelitian ini juga diharapkan mampu memberikan pengetahuan tentang penerapan ilmu visi komputer pada rekognisi gerakan bahasa isyarat.

b. Bagi penulis

Manfaat bagi penulis yaitu dapat mengembangkan dan mengimplementasikan ilmu yang telah diperoleh melalui penelitian rekognisi gerakan bahasa isyarat dinamis dengan menggunakan metode LSTM.

c. Bagi peneliti lainnya

Penelitian ini diharapkan dapat digunakan sebagai rujukan atau referensi dalam mengembangkan penelitian sejenis ataupun penelitian lainnya yang berkaitan dengan bahasa isyarat Indonesia maupun metode LSTM.

