

BAB I

PENDAHULUAN

1.1 Latar Belakang

Bahasa isyarat merupakan serangkaian gerakan tangan yang teratur, melibatkan aspek visual dan menggunakan berbagai bagian tubuh seperti jari, tangan, lengan, kepala, tubuh, dan ekspresi wajah untuk menyampaikan pesan dalam proses berkomunikasi. Bagi individu tunarungu, bahasa isyarat menjadi alat yang sangat penting dalam interaksi sehari-hari (Hossen et al., 2018). Di Indonesia terdapat dua sistem bahasa isyarat yaitu Sistem Bahasa Isyarat Indonesia (SIBI) dan Bahasa Isyarat Indonesia (BISINDO) (Mursita, 2015).

Bagi individu yang mengalami tunarungu, berkomunikasi melalui bahasa lisan seringkali menjadi tantangan yang sulit atau bahkan tidak mungkin dilakukan. Oleh karena itu, bahasa isyarat menjadi alat utama dalam interaksi dan komunikasi sehari-hari bagi mereka. Namun, bahasa isyarat tidak umum di kalangan masyarakat (Farooq et al., 2021), dan hal ini menjadi sebuah masalah ketika individu tunarungu berkomunikasi dengan masyarakat di sekitarnya yang tidak memahami bahasa isyarat. Oleh karena itu, masyarakat perlu memahami bahasa isyarat agar komunikasi antara individu tunarungu dengan masyarakat bisa berjalan baik. Para peneliti telah melakukan penelitian yang mengembangkan model deteksi bahasa isyarat sebagai salah satu solusi dari permasalahan tersebut. Contohnya pada penelitian (Sholawati et al., 2022) mengimplementasikan metode *CNN* untuk

pengenalan abjad jari SIBI dengan hasil *sensitivity*, *specificity*, *precision*, *recall* dan *accuracy* sebesar 80,76%.

Selanjutnya pada penelitian (Naufal & Kusuma, 2023) melakukan perbandingan antara metode *KNN*, *SVM*, dan *CNN* dengan arsitektur *transfer learning* untuk klasifikasi citra abjad jari SIBI. Hasil dari perbandingan tersebut adalah *CNN* dengan arsitektur Xception memiliki performa akurasi yang terbaik yaitu 99,57%, sedangkan VGG16 memiliki performa terburuk, yaitu dengan akurasi 42,13%. Waktu komputasi total proses *training* dan *testing* di algoritma *machine learning* jauh lebih cepat jika dibandingkan dengan *transfer learning*. *KNN* untuk semua nilai K memiliki waktu komputasi sebesar 0,03 detik. *SVM* membutuhkan waktu komputasi paling cepat 11,42 detik yaitu pada kernel linear dan paling lama 14,81 detik yaitu pada kernel RBF. *Transfer learning* membutuhkan waktu komputasi paling cepat 1332,94 detik yaitu pada arsitektur VGG16 dan paling lama 1396,11 detik yaitu pada arsitektur ResNet50 dengan jumlah epoch 100. (Naufal & Kusuma, 2023) menyarankan untuk menggunakan pemilihan fitur selain warna dan mengkombinasikan metode ekstraksi fitur tersebut dengan *classifier machine learning* untuk meningkatkan akurasi sehingga diharapkan bisa mendekati akurasi dari *transfer learning* dengan waktu komputasi yang rendah.

Selanjutnya, pada penelitian (Saputro & Sumantri, 2022) yang melakukan klasifikasi citra abjad jari SIBI menggunakan algoritma *KNN* memperoleh akurasi sebesar 88,8%. Terdapat beberapa huruf yang belum konsisten dikenali seperti huruf C, E, L, U dan V yang cenderung memiliki nilai kemiripan yang dekat.

(Saputro & Sumantri, 2022) menyarankan untuk mengembangkan metode klasifikasi atau ekstraksi fitur lain yang relevan untuk meningkatkan akurasi.

Pada beberapa penelitian tersebut belum optimalnya performa model disebabkan oleh adanya kemiripan pada beberapa huruf sehingga dapat meningkatkan potensi kesalahan deteksi yang membuat model sulit untuk membedakan antara huruf-huruf yang mirip secara visual. Berdasarkan saran dari beberapa penelitian sebelumnya, untuk melakukan pengembangan model dibutuhkan metode ekstraksi fitur lain dengan pemilihan fitur selain warna dengan harapan bisa meningkatkan performa dari model.

Sebagai perbandingan untuk mengetahui algoritma *machine learning* dan metode ekstraksi fitur yang paling optimal terdapat beberapa penelitian yang mengembangkan model pada beberapa bahasa isyarat yang berbeda. Pada penelitian (Larasati et al., 2021) telah dilakukan pengenalan *American Sign Language (ASL)* dengan membandingkan algoritma klasifikasi *K-NN* dan *random forest*, hasil dari penelitian tersebut *random forest* memiliki nilai akurasi yang lebih tinggi dari *K-NN* dengan nilai 99,10% sedangkan *K-NN* hanya memperoleh akurasi sebesar 72,43%. Adapun penelitian dari (Yuan et al., 2020) yang melakukan pengenalan abjad *Chinese Sign Language (CSL)*, penelitian tersebut membandingkan tiga algoritma yaitu, *Artificial Neural Network (ANN)*, *Support Vector Machine (SVM)*, dan *random forest*. Dari perbandingan tersebut didapat akurasi yang terbaik adalah *random forest* dengan hasil akurasi sebesar 95,48%, sedangkan untuk *SVM* dan *ANN* masing-masing menghasilkan akurasi sebesar 94,68% dan 92,56%. Dari hasil beberapa penelitian tersebut terlihat pada beberapa

bahasa isyarat yang berbeda *random forest* lebih unggul dibanding dengan metode klasifikasi yang lainnya.

Mediapipe hands berpotensi untuk meningkatkan akurasi karena penggunaannya dapat melacak 21 *keypoint landmark* pada tangan termasuk ujung jari, buku jari, dan pergelangan tangan dengan tingkat akurasi yang tinggi (Zhang et al., 2020). Hal tersebut terbukti pada penelitian (Yasumuro & Jin'no, 2022) yang melakukan deteksi Bahasa Isyarat pada Bahasa Jepang menggunakan *mediapipe hands* menghasilkan akurasi 99% sampai 100%. Adapun penerapan lainnya pada Bahasa Asing yaitu *American Sign Language* memperoleh akurasi 94,4% (Alvin et al., 2021).

Oleh karena itu berdasarkan beberapa penelitian sebelumnya, maka pada penelitian ini dilakukan pengembangan model abjad jari bahasa isyarat SIBI menggunakan *mediapipe hands* sebagai metode ekstraksi fitur dan *random forest* sebagai algoritma klasifikasi, dengan harapan akan meningkatkan performa dari model pada penelitian sebelumnya.

1.2 Rumusan Masalah

Berdasarkan uraian pada latar belakang, rumusan masalah yang akan dibahas pada penelitian ini sebagai berikut:

- a. Bagaimana implementasi ekstraksi fitur *mediapipe hands* dan algoritma klasifikasi *random forest* untuk meningkatkan performa model deteksi abjad jari bahasa isyarat SIBI?

- b. Berapa nilai akurasi, presisi, *recall*, *f1-score*, dan waktu komputasi yang didapat pada proses klasifikasi?

1.3 Tujuan Penelitian

Dari rumusan masalah maka didapat tujuan dari penelitian ini adalah sebagai berikut:

- a. Mengimplementasikan ekstraksi fitur *mediapipe hands* dan algoritma klasifikasi *random forest* untuk meningkatkan performa model deteksi abjad jari bahasa isyarat SIBI.
- b. Menghitung nilai akurasi, presisi, *recall*, *f1-score*, dan waktu komputasi dari proses klasifikasi.

1.4 Manfaat Penelitian

Manfaat yang diharapkan dari penelitian ini adalah sebagai berikut:

- a. Mengembangkan model dengan cara meningkatkan performa dari model deteksi abjad jari bahasa isyarat SIBI.

1.5 Batasan Masalah

Batasan masalah yang terdapat pada penelitian ini adalah sebagai berikut:

- a. Klasifikasi dilakukan pada abjad jari bahasa isyarat SIBI.
- b. Dataset yang digunakan bersumber dari kaggle, dataset ini bernama “Datasets SIBI Sign Language Alphabets” (<https://www.kaggle.com/datasets/mlanangafkaar/datasets-lemlitbang-sibi-alphabets>).

- c. Bahasa pemrograman yang digunakan adalah Python 3.9 selain itu menggunakan *framework* mediapipe hands dan algoritma *klasifikasi random forest* yang disediakan oleh *scikit-learn*.