

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Dalam kehidupan sehari-hari, kita akan bertemu, berinteraksi, dan berkomunikasi dengan orang lain, baik orang yang dekat seperti anggota keluarga, teman, dan/atau orang yang kita temui di publik. Dalam interaksi ini, kita dapat bertemu dengan orang tunawicara dan tunarungu. Tunawicara merupakan keadaan di mana penyandang mengalami kesulitan atau tidak dapat berkomunikasi secara verbal atau berbicara secara normal (Mutiarra *et al.*, 2023). Tunarungu merupakan keadaan di mana penyandang tidak dapat mendengar sebagian atau seluruh suara dan biasanya mengalami gangguan komunikasi verbal dikarenakan sulit dalam menyampaikan informasi yang didengar (Septiawati *et al.*, 2021).

Agar dapat berkomunikasi dengan orang tunawicara dan tunarungu lainnya, seseorang yang tunawicara dan tunarungu dapat menggunakan bahasa isyarat sebagai salah satu medianya. Di Indonesia terdapat dua jenis bahasa isyarat yang digunakan, yaitu Sistem Isyarat Bahasa Indonesia (SIBI) dan Bahasa Isyarat Indonesia (BISINDO). Standar bahasa isyarat formal yang diakui oleh pemerintah Indonesia sebagai bahasa untuk berkomunikasi adalah SIBI (Syulistyo *et al.*, 2020). SIBI menggunakan gestur gerakan tangan, jari, bibir, dan lain-lain untuk mengutarakan alfabet dan/atau kata dalam Bahasa Indonesia (Anggraini *et al.*, 2019).

Walaupun komunikasi antar sesama dapat berjalan dengan lancar, penggunaan bahasa isyarat dapat menjadi suatu permasalahan ketika seseorang yang tunawicara dan/atau tunarungu ingin berkomunikasi dengan orang normal. Hal ini dikarenakan banyak orang normal yang tidak tahu cara berkomunikasi dengan menggunakan bahasa isyarat. Selain itu, banyak juga orang normal yang tidak memiliki dorongan yang kuat untuk mempelajarinya dikarenakan jumlah orang tunawicara dan/atau tunarungu yang jauh lebih sedikit sehingga kemungkinan untuk berinteraksi dan berkomunikasi antar kedua pihak menjadi kecil. Saat ini, banyak cara yang dapat digunakan untuk menginterpretasikan bahasa isyarat. Salah satu cara agar bahasa

isyarat dapat secara langsung diinterpretasikan adalah dengan menerapkan *machine learning*.

Dalam pengimplementasian *machine learning* untuk pengenalan dan penginterpretasian bahasa isyarat, dibutuhkan sebuah model *machine learning* yang dilatih (*train*) menggunakan dataset bahasa isyarat BISINDO, SIBI, atau bahkan *American Sign Language* (ASL) dengan salah satu algoritma atau model *machine learning* yang sesuai. Beberapa penelitian sebelumnya sudah mencoba membuat suatu model *machine learning* yang berperan dalam gestur bahasa isyarat Indonesia. Dalam penelitian Syulistyo pada tahun 2020, peneliti berhasil membuat model *machine learning* untuk mendeteksi dan mentranslasikan masukan tiga kelas gestur tangan SIBI, berupa “A”, “Assalamualaikum”, dan “Hallo” dengan menggunakan model *Convolutional Neural Network* (CNN) *pre-trained* berupa *faster_rcnn_inception_v2_coco* (Syulistyo, *et al.*, 2020). Walaupun dalam penelitian tersebut hanya menerapkan kelas “A”, “Assalamualaikum”, dan “Hallo”, akurasi model dalam mentranslasikan masukan gambar ke dalam kelas yang diharapkan sebesar 100%. Dalam penelitian Arisandi pada tahun 2022, peneliti berhasil membuat sebuah model *machine learning* untuk mengklasifikasikan gestur tangan BISINDO dengan menggunakan algoritma CNN (Arisandi dan Satya, 2022). Dalam penelitian tersebut, peneliti menggunakan *dataset* gambar dengan ukuran 256x256 piksel dan hasil akurasi model dalam mengklasifikasi gestur tangan mencapai 99,82%.

Dalam pengembangan sistem *machine learning*, tidak jarang ditemukan penggunaan lebih dari satu algoritma atau model *machine learning* dalam *training* model. Salah satu tujuan dalam perbandingan antara lebih dari satu algoritma *machine learning* adalah untuk mencari performa terbaik dari salah satu algoritma. Dalam penelitian Harahap pada tahun 2024, peneliti membandingkan performa dari segi akurasi *training* serta validasi dengan aspek *full retraining* (FR) dan *fine tuning* (FT) antara model FT-SqueezeNet, FR-SqueezeNet, FT-ResNet, dan FR-ResNet (Harahap *et al.*, 2024). Penelitian memberikan hasil perbandingan model di mana model FR-ResNet memberikan akurasi tertinggi pada akurasi *training* dan validasi dengan nilai akurasi sebesar 97,77% dan 97,22% secara berurut.

Tujuan dari penelitian ini adalah untuk menyediakan pengetahuan dari perbandingan performa antar arsitektur model ResNet dalam *training* sistem *machine learning* untuk pengenalan gestur tangan bahasa isyarat SIBI. *Dataset* bahasa isyarat yang tepat untuk digunakan pada saat *training* model akan dicari dan dikumpulkan. Hasil *training* dari masing-masing model akan dilakukan pengukuran performa dengan menggunakan metrik yang sesuai. Dengan demikian, penelitian ini akan memberikan pengetahuan dalam perbandingan performa antar arsitektur model ResNet dalam pengenalan gestur tangan SIBI.

Penelitian sebelumnya yang dilakukan oleh Brianorman pada tahun 2024 dengan judul “*Comparative Analysis of CNN Architectures for SIBI Image Classification*”, peneliti membandingkan performa di antara beberapa model CNN dalam klasifikasi gambar gestur tangan SIBI (Brianorman dan Utami, 2024). Penelitian menggunakan model CNN yang terdiri dari VGG16, Inception, MobileNetV2, ResNet50, dan Xception serta menggunakan dataset gambar gestur tangan huruf alfabet dan angka SIBI dengan total gambar sebanyak 2515. Penelitian ini menunjukkan bahwa VGG16 dan ResNet50 menghasilkan performa terbaik, di mana VGG16 menghasilkan nilai akurasi sebesar 99,60% dengan performa yang mirip pada *precision*, *recall*, dan *F1-Score* serta ResNet50 yang menghasilkan nilai akurasi sebesar 98,80% dengan nilai *precision*, *recall*, dan *F1-Score* yang setara.

Penelitian kedua yang dilakukan oleh Setyono pada tahun 2019 dengan judul “*Recognizing word gesture in sign system for Indonesian language (SIBI) Sentences using DeepCNN and BiLSTM*”, peneliti membandingkan performa di antara MobileNetV2 dan ResNet50 dengan menerapkan BiLSTM dalam klasifikasi gestur kata dan kalimat SIBI (Setyono dan Rakun, 2019). Dataset penelitian menggunakan 2275 total video gestur kata SIBI yang terbagi menjadi 28 kalimat. Penelitian menunjukkan bahwa ResNet50 merupakan model dengan performa terbaik dari segi akurasi *test*, sebesar 99,879% serta dari segi *precision*, *recall*, dan *F1-Score*, tetapi MobileNetV2 menghasilkan performa terbaik dari segi durasi *training* dengan durasi tercepat dibandingkan dengan durasi ResNet50.

Penelitian ketiga yang dilakukan oleh Shania pada tahun 2022 dengan judul “*Translator of Indonesian Sign Language Video using Convolutional Neural*

Network with Transfer Learning”, peneliti membandingkan performa dari di antara beberapa model CNN dalam klasifikasi bahasa isyarat BISINDO dan model dengan performa terbaik akan diimplementasikan ke dalam sistem translasi bahasa isyarat (Shania *et al.*, 2022). Peneliti menggunakan model MobileNetV2, ResNet50V2, dan Xception dengan dataset gambar sebanyak 220 gambar yang dibagi menjadi 20 kelas. Penelitian ini menunjukkan bahwa model ResNet50V2 menghasilkan nilai yang memuaskan, walaupun bukan model terbaik. Pengujian pada dataset ekspresi wajah, ResNet50V2 menghasilkan nilai akurasi *validation* sebesar 96,57%, nilai *validation loss* sebesar 0,251, nilai akurasi *test* sebesar 91,36%, nilai *test loss* sebesar 0.468, dan *f1-score* sebesar 91,4%. Pengujian pada dataset gestur tangan, ResNet50V2 menghasilkan nilai akurasi *validation* sebesar 99,38%, nilai *validation loss* sebesar 0.09, nilai akurasi *test* sebesar 95%, nilai *test loss* sebesar 0.186, dan nilai *f1-score* sebesar 95%.

1.2 Identifikasi Masalah

Berdasarkan latar belakang yang telah dijabarkan, maka dapat diidentifikasi beberapa masalah sebagai berikut:

1. Belum ada penelitian yang meneliti tentang perbandingan performa antar arsitektur ResNet dalam kasus pengenalan bahasa isyarat SIBI.
2. Bahasa isyarat yang formal dan resmi diakui oleh pemerintah Indonesia.
3. Permasalahan yang dihadapi dalam komunikasi antara orang tunawicara dan/atau tunarungu dengan orang normal.
4. Permasalahan yang dihadapi dalam menginterpretasikan bahasa isyarat di Indonesia.
5. Solusi untuk menginterpretasikan bahasa isyarat di Indonesia.

1.3 Pembatasan Masalah

Penelitian mengenai pengembangan aplikasi pendeteksi bahasa isyarat dengan *machine learning* memiliki sejumlah batasan yang perlu dipertimbangkan, yaitu:

1. Penelitian ini akan berfokus pada aspek perbandingan antar performa hasil pelatihan model *machine learning* menggunakan enam arsitektur model ResNet berbeda untuk deteksi gestur tangan SIBI.

2. Penelitian ini akan menggunakan rasio *dataset* antara *training*, *validation*, dan *testing* sebesar 64:16:20 (64% total data untuk *training*, 16% total data untuk *validation*, dan 20% total data untuk *testing*).
3. Penelitian ini akan menggunakan arsitektur model ResNet50, ResNet101, ResNet152, ResNet50V2, ResNet101V2, dan ResNet152V2.
4. Penelitian ini akan mengukur performa dari masing-masing model dengan menggunakan metrik berupa *confusion matrix*, *classification report*, akurasi, nilai *train-test loss*, dan durasi *training* model.
5. Penelitian ini akan menggunakan SIBI sebagai sistem bahasa isyarat dengan gestur tangan huruf alfabet statis.
6. Penelitian ini akan menggunakan dataset gambar gestur tangan yang memiliki kesamaan dengan SIBI dengan gestur tangan huruf alfabet statis dalam pelatihan tiap model *machine learning*.
7. Penelitian ini memiliki pembatasan terkait dengan hasil penelitian di mana hasil yang dipaparkan berupa proses *training* model, pengukuran performa model, dan perbandingan performa model.

1.4 Perumusan Masalah

Berdasarkan latar belakang dan identifikasi masalah di atas, perumusan masalah penelitian ini adalah sebagai berikut:

Bagaimana cara mengetahui dan membandingkan performa hasil *training* di antara arsitektur model ResNet dalam pengenalan gestur tangan SIBI?

1.5 Tujuan Penelitian

Tujuan penelitian ini adalah untuk:

1. Memberikan gambaran *training* dan pengembangan beberapa arsitektur model ResNet yang ditentukan untuk deteksi objek dan mengukur performanya.
2. Memberikan pemahaman mengenai penggunaan model ResNet dalam mendeteksi dan menginterpretasikan gestur tangan dalam bahasa isyarat SIBI.
3. Menyediakan hasil berupa performa hasil *training* dari model-model ResNet yang ditentukan.

4. Memberikan perbandingan antar performa hasil *training* dari model-model ResNet yang ditentukan dalam mendeteksi dan menginterpretasikan gestur tangan dalam bahasa isyarat SIBI.

1.6 Manfaat Penelitian

1. Meningkatkan pemahaman mengenai perbedaan performa antar model-model ResNet.
2. Meningkatkan pemahaman mengenai penggunaan model ResNet dalam mendeteksi dan menginterpretasikan gestur tangan dalam bahasa isyarat SIBI.
3. Memberikan referensi bagi penelitian atau pengembangan aplikasi atau sistem di bidang bahasa isyarat yang menerapkan *Convolutional Neural Network* (CNN) dengan memperhatikan faktor performa.
4. Memberikan kontribusi bagi pengetahuan akademis mengenai implementasi *machine learning*, terutama dengan model ResNet dalam pengenalan dan penginterpretasian gestur tangan SIBI.

