

Deteksi Gestur Sistem Isyarat Bahasa Indonesia Menggunakan Metode *Deep learning* SSD MobileNet V2 FPNLite

Gesture Detection of Indonesian Sign Language Using Deep learning Method SSD MobileNet V2 FPNLite

Nurul Tazkiyah Adam^{1*}, Zahra Arwananing Tyas², Tikaridha Hardiani³

^{1,2,3}Program Studi Teknologi Informasi, Universitas Aisyiyah Yogyakarta

*corr_author: tazyatazkiyah@gmail.com

ABSTRAK

Penyandang disabilitas tunarungu memiliki keterbatasan pada indra pendengaran yang mengakibatkan hambatan dalam perkembangan bahasa sehingga memerlukan bimbingan dan pendidikan khusus. Penyandang disabilitas dengan usia 15 tahun ke atas sebagian besar baru mampu menyelesaikan pendidikan dengan tamatan SD berdasarkan data dari Badan Pusat Statistik. Tantangan lainnya yakni masih terdapat kesenjangan komunikasi antara tunarungu dan non-tunarungu karena tidak sedikit yang masih tidak mengerti bahasa isyarat. Seiring berkembangnya waktu, teknologi diadopsi untuk menyelesaikan permasalahan ini, salah satunya sistem *deep learning* yang mampu mendeteksi gestur SIBI (Sistem Isyarat Bahasa Indonesia). Penelitian ini menggunakan Arsitektur Model *Pre-Trained SSD MobileNet V2 FPNLite* dan *Framework TensorFlow Object Detection* untuk mendeteksi 15 gestur SIBI (Huruf Vokal A, I, U, E, O dan 10 kata yakni Baik, Bangku, Bel, Dia, Meja, Pramuka, Sakit, Saya, Teman, dan Tugas). Penelitian ini melibatkan *dataset* berjumlah 15.000 gambar yang dibagi menjadi 80% data pelatihan, 10% data validasi, dan 10% data pengujian. Kumpulan data mencakup variasi sudut, kamera, warna, latar belakang, jenis kelamin, maupun usia. Model dilatih selama 40.000 *steps* dengan *batch size* 16. Hasil evaluasi menunjukkan bahwa model berhasil mendeteksi 15 gestur SIBI tersebut dengan nilai akurasi model mAP @ 0.5:0.05:0.95 sebesar 89,93% dan nilai *loss* sebesar 0.1434.

Kata-kata kunci: SIBI, *deep learning*, SSD MobileNet V2 FPNLite, TensorFlow Object Detection, Mean Average Precision

ABSTRACT

Hearing-impaired individuals face limitations in auditory perception, leading to challenges in language development and necessitating special guidance and education. According to data from the Central Statistics Agency, most people with hearing disabilities over the age of 15 have only completed elementary school education. Another challenge is the communication gap between the hearing-impaired and non-hearing-impaired communities, as many are still unfamiliar with sign language. With the advancement of technology, deep learning systems have been adopted to address this issue. This research utilized the Pre-Trained Model Architecture SSD MobileNet V2 FPNLite and the TensorFlow Object Detection Framework to detect 15 SIBI (Indonesian Sign Language System) gestures, including vowel letters (A, I, U, E, O) and 10 words (Baik, Bangku, Bel, Dia, Meja, Pramuka, Sakit, Saya, Teman, Tugas). The study involved a dataset of 15,000 images, divided into 80% training data, 10% validation data, and 10% testing data. The dataset includes variations in angles, cameras, colors, backgrounds, gender, and age. The

model was trained for 40,000 steps with a batch size of 16. Evaluation results show that the model successfully detected the 15 SIBI gestures with a mean Average Precision (mAP) score @ 0.5:0.05:0.95 of 89.93% and a loss value of 0.1434.

Keywords: *SIBI, deep learning, SSD MobileNet V2 FPNLite, TensorFlow Object Detection, Mean Average Precision*

PENDAHULUAN

Pengertian penyandang disabilitas menurut Undang-Undang Republik Indonesia Pasal 1 ayat 1 Nomor 8 Tahun 2016 adalah setiap individu yang mengalami keterbatasan fisik, intelektual, mental dan/atau sensorik dalam jangka waktu yang lama, penyandang disabilitas dapat mengalami hambatan dan kesulitan untuk berpartisipasi secara penuh dan efektif dengan warga negara lainnya berdasarkan kesamaan hak (Indonesia, 2016). Pengertian tunarungu secara medis yakni kurang atau hilangnya kemampuan dalam mendengar yang disebabkan oleh rusak atau tidak berfungsinya sebagian atau seluruh alat-alat pendengaran. Kurang atau hilangnya pendengaran dapat mengakibatkan hambatan dalam perkembangan bahasa sehingga memerlukan bimbingan dan pendidikan khusus (Suryanah, 1996).

Penyandang disabilitas, termasuk tunarungu, memiliki hak untuk mendapatkan kesempatan pendidikan yang setara dengan non-disabilitas serta menerima proses pembelajaran yang sesuai dengan kebutuhan dan kemampuan mereka. Akan tetapi, masih terdapat ketimpangan dalam capaian pendidikan antara penyandang disabilitas dan non-disabilitas. Penyandang disabilitas dengan usia 15 tahun ke atas, sebagian besar, baru mampu menamatkan pendidikan Sekolah Dasar (SD) sedangkan 40 persen dari mereka belum mampu atau bahkan tidak pernah mengenyam pendidikan SD dan hanya 16,78 persen yang mampu menamatkan pendidikan sampai dengan SMA ke atas (Statistik, 2023).

Tantangan besar yang dihadapi oleh individu tunarungu, selain tantangan capaian pendidikan, yakni komunikasi, khususnya keterbatasan dalam berkomunikasi dengan orang yang tidak mengerti bahasa isyarat. Bahasa isyarat menjadi alat komunikasi yang digunakan oleh tunarungu, tetapi tidak sedikit orang yang tidak memahami bahasa tersebut. Hal ini menyebabkan kesenjangan komunikasi antara tunarungu dan individu non-tunarungu (Sari and Altirika, 2023).

Sekolah Luar Biasa (SLB) merupakan sekolah khusus untuk para penyandang disabilitas mulai jenjang Sekolah Dasar hingga Sekolah Menengah Atas. Di Indonesia terdapat 2 jenis bahasa isyarat yakni Bahasa Isyarat Indonesia (BISINDO) dan Sistem Bahasa Isyarat Indonesia (SIBI). Keputusan Menteri Pendidikan dan Kebudayaan Republik Indonesia Nomor 0161/U/2994, pemerintah membakukan Kamus Sistem Isyarat Bahasa Indonesia sehingga SIBI menjadi bahasa isyarat resmi yang dipakai seluruh SLB di bawah naungan Kementerian Pendidikan dan Kebudayaan (Kemendikbud) (Bagus, Bakti and Pranoto, 2019).

Seiring berkembangnya waktu, teknologi diadopsi untuk upaya dalam mengatasi kesenjangan dalam komunikasi dan pendidikan antara tunarungu dan non-tunarungu, serta teknologi dapat memperkaya kreatifitas dalam pembelajaran SIBI seperti sistem *Deep learning* yang dapat mendeteksi gerakan atau gestur SIBI (Maulida, 2023). *Deep learning* merupakan salah satu metode dalam kecerdasan buatan yang memungkinkan komputer untuk memproses data dengan cara yang meniru fungsi otak manusia. Model-model *Deep learning* mampu mengenali pola-pola kompleks dalam berbagai jenis data seperti gambar, teks, dan suara, sehingga dapat menghasilkan wawasan dan prediksi yang akurat (Kahlil, 2023).

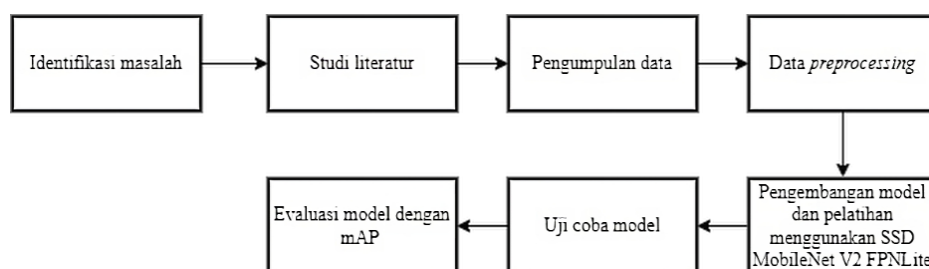
Penelitian tentang deteksi SIBI yang menerapkan *Deep learning* sebelumnya sudah pernah dilakukan oleh beberapa peneliti, diantaranya penelitian yang mendeteksi 6 gestur SIBI (Saya, Dia, Kamu, Cinta, Maaf, dan Sedih) menggunakan *Convolutional Neural Network* (CNN) dan *dataset* sebanyak 660 gambar. Hasilnya menunjukkan masih terdapat nilai *loss* yang signifikan, yaitu 25% untuk pelatihan dan 23% untuk validasi, dan menunjukkan kesalahan prediksi yang cukup besar pada beberapa gestur (Putra, 2021). Tahun selanjutnya, penggunaan arsitektur Model *Pre-Trained Single Shot Detector* (SSD) *MobileNet V2* untuk mendeteksi 6 gestur SIBI yang sama dengan jumlah *dataset* sebanyak 600 gambar. Hasil terbaik menunjukkan nilai *loss* 0,1387 dan nilai akurasi 86,6% (Hidayatullah, 2022).

Penelitian yang membandingkan performa arsitektur Model *Pre-Trained SSD MobileNet V2* dan Model *Pre-Trained SSD MobileNet V2 Feature Pyramid Network Lite* (FPN_{Lite}) menggunakan *Framework TensorFlow Object Detection* untuk kasus deteksi helm pengendara dengan *dataset* 604 gambar menunjukkan SSD *MobileNet V2 FPN_{Lite}* mempunyai akurasi yang baik dengan tingkat kepercayaan 99% dan memiliki nilai *loss* yang lebih kecil dengan waktu *training* model lebih cepat dibandingkan SSD *MobileNet V2* (Dionisius, Riti and Christian, 2024).

Penelitian ini mendeteksi gestur SIBI menggunakan metode *Deep learning* dengan memanfaatkan *Framework TensorFlow Object Detection* dan arsitektur Model *Pre-Trained SSD MobileNet V2 FPN_{Lite}*. Penelitian ini menggunakan sejumlah *dataset* yang terbagi ke dalam 15 gestur SIBI (Huruf Vokal A, I, U, E, O dan 10 kata yakni Baik, Bangku, Bel, Dia, Meja, Pramuka, Sakit, Saya, Teman, dan Tugas) dengan berbagai variasi sudut, kamera, warna, latar belakang, jenis kelamin, maupun usia. Diharapkan penelitian ini menghasilkan model deteksi yang memiliki ketepatan dalam mendeteksi dan nilai akurasi yang lebih baik dari penelitian sebelumnya. Selain itu, penelitian ini diharapkan dapat berkontribusi dalam pengetahuan sistem deteksi serta dapat membantu melestarikan dan mempopulerkan SIBI sebagai bahasa isyarat nasional Indonesia.

METODE PENELITIAN

Metode penelitian dilakukan dengan beberapa tahapan kerja untuk mencapai tujuan penelitian, yaitu model dapat mendeteksi 15 (lima belas) kelas gestur SIBI dan mengetahui berapa nilai akurasinya. Gambaran tahapan-tahapan kerja tersebut dapat dilihat pada diagram alur penelitian Gambar 1.



Gambar 1. Diagram alur penelitian

1. Pengumpulan Data

Proses pengumpulan data pada penelitian ini dilakukan secara primer, gambar-gambar dikumpulkan langsung oleh peneliti menggunakan berbagai perangkat kamera mobile. *Dataset* yang digunakan terdiri dari gambar-gambar atau citra yang diambil dengan rasio aspek yang konsisten, yaitu 16:9, atau dalam mode potret, menggunakan berbagai merek kamera device yang berbeda-beda seperti Samsung Galaxy S22, Samsung Galaxy

S23 Ultra, Vivo Y12, Xiomi Redmi Note13, dan iPhone 12. Pemilihan berbagai merek kamera ini dilakukan untuk memastikan variasi kualitas gambar yang dihasilkan

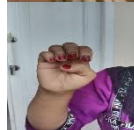
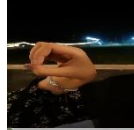
Latar belakang atau background pengambilan gambar dipilih dengan beragam situasi, baik di dalam ruangan maupun di luar ruangan. Hal ini dipertimbangkan untuk memperoleh variasi pencahayaan dan latar belakang yang lebih beragam, yang dapat mempengaruhi kualitas dan karakteristik visual dari gambar-gambar yang diambil.

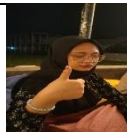
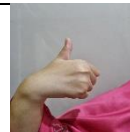
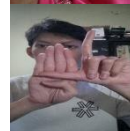
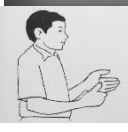
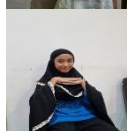
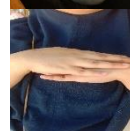
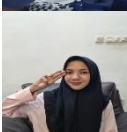
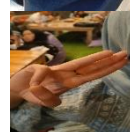
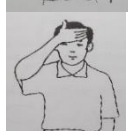
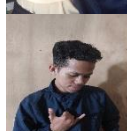
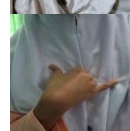
Subjek dalam gambar-gambar ini adalah manusia, yang mencakup berbagai perbedaan seperti jenis kelamin (pria dan wanita), yang memiliki variasi rambut pendek, rambut panjang, dan penggunaan jilbab pada subjek wanita. Selain itu, subjek juga dipilih dengan melibatkan variasi usia, mulai dari anak-anak, remaja, hingga dewasa. Hal ini dilakukan untuk memperluas variasi bentuk dan ukuran tangan yang mengilustrasikan gestur SIBI karena tangan antar kelompok usia berbeda.

Pengambilan gambar dilakukan dengan variasi sudut, yakni dari sisi depan, kanan, dan kiri, serta dengan fokus memperbesar bagian tangan yang membentuk gestur SIBI. Setiap subjek atau orang yang berpartisipasi dalam penelitian diminta untuk membentuk 15 jenis gestur SIBI yang berbeda. Gestur-gestur ini mencakup huruf vokal SIBI (A, I, U, E, dan O) dan sepuluh kata SIBI (Saya, Dia, Teman, Baik, Bel, Bangku, Meja, Sakit, Tugas, dan Pramuka). Pemilihan kata-kata ini didasarkan pada kosa kata yang akrab didengar dalam lingkungan sekolah. Panduan gestur mengacu pada Kamus SIBI yang dikeluarkan oleh Kementerian Pendidikan dan Kebudayaan (Kemendikbud).

Dataset yang dikumpulkan dalam penelitian ini tidak hanya mencakup variasi visual yang luas, tetapi juga memastikan bahwa gestur-gestur yang direkam sesuai dengan standar dan pedoman gestur SIBI resmi yang ada di Indonesia, sehingga dapat dijadikan *dataset* yang valid untuk penelitian sistem deteksi gerakan SIBI. Tabel 1 memberikan gambaran *dataset*. Total jumlah *dataset* awal yang dikumpulkan peneliti adalah 3.000 gambar dengan masing-masing gestur memiliki 200 gambar. *Dataset* nantinya akan bertambah setelah proses augmentasi menjadi 15.000 gambar.

Tabel 1. Dataset

No.	Bahasa Lisan	Kamus SIBI	Contoh Hasil Variasi Dataset Gestur SIBI			
1.	A					
2.	E					
3.	I					
4.	O					
5.	U					

No.	Bahasa Lisan	Kamus SIBI	Contoh Hasil Variasi <i>Dataset</i> Gestur SIBI			
6.	Baik					
7.	Bangku					
8.	Bel					
9.	Dia					
10.	Meja					
11.	Pramuka					
12.	Sakit					
13.	Saya					
14.	Teman					
15.	Tugas					

2. Data Preprocessing

Data *preprocessing* meliputi serangkaian teknik maupun prosedur yang digunakan untuk mempersiapkan dan mengubah data mentah menjadi format yang cocok dalam pelatihan model *Deep learning*. Pada tahap *preprocessing* awal bertujuan untuk mengubah ukuran gambar, memberikan label gestur SIBI pada gambar, dan melakukan augmentasi dengan *color shifting* untuk memperbanyak variasi *dataset*.

Semua *dataset* gambar yang berukuran potret dengan rasio 19:6 dari berbagai kamera diubah menjadi ukuran 320x320 piksel. Proses ubah ukuran gambar menggunakan *script* untuk memudahkan *resize* dalam jumlah banyak.

Tahap selanjutnya, setiap objek (gestur SIBI) dalam gambar dianotasi dengan memberikan *bounding box* dan label kelas secara manual dengan menggunakan program *open source* Labellmg (Community, 2015). Hasil anotasi tersimpan dalam format file .xml. Anotasi diperlukan untuk memberi tahu model lokasi dan kelas objek (gestur SIBI) dalam setiap gambar. File XML berisi informasi-informasi penting yang akan digunakan untuk proses pelatihan berupa label *class* gestur, *filename* gambar, *width* 320, *height* 320, dan titik koordinat *bounding box* label berupa *xmin*, *ymin*, *xmax*, dan *ymax*.

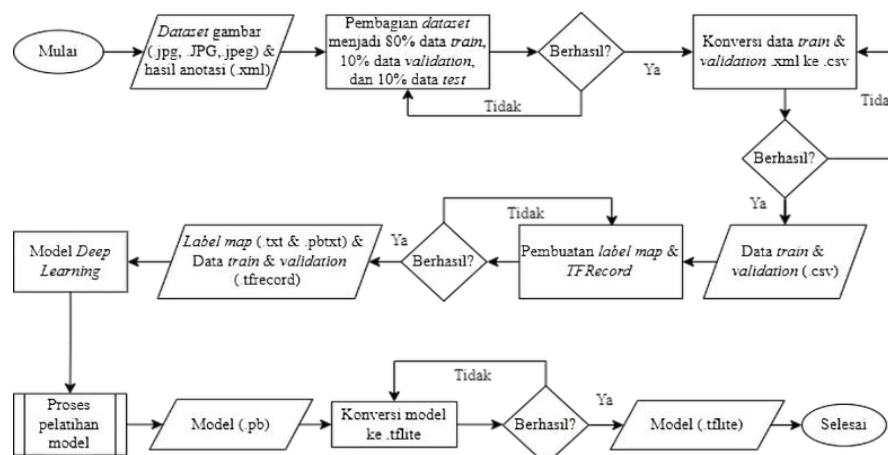
Tahap proses augmentasi menggunakan bantuan *script* untuk memudahkan proses augmentasi *color shifting* atau mengubah warna gambar tanpa harus melakukan proses *labeling* lagi, dengan menduplikat dan menyesuaikan hasil anotasi pada file .xml. Setiap gambar asli memiliki empat variasi tambahan dengan skema warna *greyscale*, *redscale*, *greenscale*, dan *bluescale*. Total *dataset* keseluruhan setelah proses augmentasi menjadi 15.000, dimana terdapat 3.000 gambar RGB, 3.000 gambar *greyscale*, 3.000 gambar *redscale*, 3.000 gambar *greenscale*, dan 3.000 gambar *bluescale*.

3. Menyiapkan Lingkungan Virtual

Deteksi objek menggunakan model arsitektur SSD *MobileNet V2 FPNLite* memerlukan persiapan lingkungan virtual seperti *Google Colaboratory* lalu menginstal *TensorFlow Object Detection API* beserta dependensinya. Proses ini melibatkan *cloning repository TensorFlow Models* dari *GitHub* resmi *TensorFlow*, yang berisi implementasi *TensorFlow Object Detection API* serta berbagai model dan konfigurasi. Peneliti juga menyiapkan dan menghubungkan *Google Colaboratory* dengan *Google Drive* untuk menyimpan *input* dan *output* selama proses pembuatan deteksi objek gestur SIBI, karena jika hanya mengandalkan penyimpanan lokal di *Google Colab*, data akan hilang ketika sesi *running* dihentikan.

4. Rancangan Proses Pelatihan dan Pemodelan

Penelitian ini menggunakan arsitektur Model *Pre-Trained SSD MobileNet V2 FPNLite* yang termasuk arsitektur model dalam Metode *Deep learning* dan menggunakan *Framework TensorFlow Object Detection API* untuk mendeteksi gestur SIBI. Diagram *Flowchart* Rancangan Proses Pembuatan Model ditunjukkan pada Gambar 2 yang memberikan gambaran proses dari data setelah data *preprocessing* hingga menjadi model akhir yang disiapkan untuk mendeteksi gestur SIBI.



Gambar 2. Flowchart rancangan proses pembuatan mode

Lingkungan *virtual* dan data *preprocessing* awal telah dilakukan maka selanjutnya melakukan beberapa tahap data *preprocessing* sebelum data diproses dalam pemodelan dan pelatihan, yakni *dataset* akan dibagi menjadi tiga bagian: 80% *train*, 10% *validation*, dan 10% *test*. Data pada *train* digunakan untuk melatih model. Pada setiap langkah pelatihan, sekumpulan gambar pada data *train* dimasukkan ke dalam jaringan saraf. Jaringan ini memprediksi kelas dan lokasi objek dalam gambar. Data pada *validation* digunakan oleh algoritma pelatihan untuk memeriksa kemajuan pelatihan. Berbeda dengan sekumpulan gambar pada data *train*, gambar pada data *validation* hanya digunakan secara periodik selama pelatihan. Sedangkan, data pada *test* adalah sekumpulan gambar yang tidak pernah dilihat oleh jaringan saraf selama pelatihan sehingga dapat digunakan untuk pengujian model dan memeriksa seberapa akurat model tersebut. Jika proses pembagian *dataset* tidak berhasil, perlu dilakukan pengecekan kembali pada *script Python* yang digunakan untuk membagi *dataset*.

File .xml pada folder data *train* dan folder data *validation* akan dikonversi ke format .csv (*Comma-Separated Values*) karena format ini lebih mudah diproses untuk tahap selanjutnya (*TFRecord*). Jika konversi tidak berhasil, *script Python* yang digunakan untuk konversi .xml ke .csv perlu diperiksa kembali. File .csv menyimpan informasi-informasi penting dari xml berupa *filename*, *width*, *height*, *class*, *xmin*, *ymin*, *xmax*, dan *ymax*.

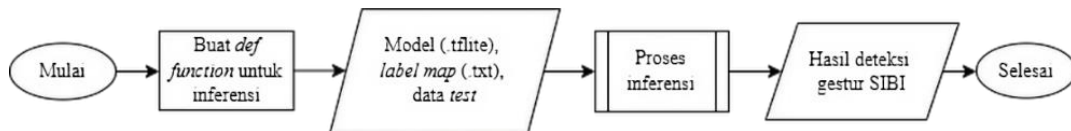
Label map dengan format .txt (*Text File*) dibuat dengan memasukkan nama-nama kelas gestur SIBI, dalam kasus ini ada 15 kelas. Kemudian, membuat format .tfrecord (*TensorFlow Record*) dari data *train* dan *validation* menjadi *train.tfrecord* dan *val.tfrecord*, selanjutnya membuat *label map* dengan format .pbtxt (*Protocol Buffer Text*). *TFRecord* berupa format file biner oleh *TensorFlow* untuk mengoptimalkan pembacaan dan penulisan data secara efisien dalam *Framework TensorFlow*. Dalam konteks *TensorFlow*, file PBTXT digunakan untuk menyimpan *label map* yang memetakan label numerik ke nama kelas. *Label map* dan *TFRecord* saling berkaitan karena *TFRecord* menyimpan data yang sudah dikonversi, sementara *label map* menyediakan informasi tentang kelas yang ada dalam *dataset*.

Tahap selanjutnya yakni, menyiapkan model yang akan digunakan dan konfigurasi pelatihan (*training*). Model pre-trained SSD *MobileNet V2 FPNLite*, model arsitektur dengan metode *Deep learning*, digunakan sebagai dasar untuk deteksi objek pada penelitian ini. Model ini sebelumnya telah dilatih oleh *TensorFlow* menggunakan COCO 2017 *dataset*. Proses pemodelan ini melibatkan penyetelan ulang (*fine-tuning*). Kemudian, melakukan beberapa modifikasi pada parameter penting seperti jumlah langkah pelatihan (*steps*), jumlah gambar yang digunakan per langkah pelatihan (*batch size*), dan lain-lain. Proses pelatihan ditampilkan melalui *TensorBoard*, yang merupakan alat visualisasi untuk memantau proses pelatihan secara real-time dengan menampilkan beberapa grafik seperti grafik *classification_loss*, grafik *localization_loss*, grafik *regularization_loss*, grafik *total_loss*, dan grafik *learning_rate*. *TensorBoard* membantu dalam analisis performa model dan penyesuaian parameter. Setelah pelatihan selesai, model disimpan dalam format .pb (*Protocol Buffer*). File .pb ini berisi arsitektur model dan bobot yang telah dilatih. Bobot yang telah dilatih ini adalah parameter dalam jaringan *neural* yang telah belajar selama pelatihan untuk meminimalkan *loss* dan meningkatkan akurasi prediksi.

Model yang disimpan dalam format .pb akan dikonversi ke format *TensorFlow Lite* (.tflite) agar dapat digunakan pada perangkat dengan sumber daya terbatas, selain itu penggunaan model .tflite memungkinkan peneliti untuk melakukan *testing* pada *interface Google Colab*. Jika proses konversi tidak berhasil, perlu diperiksa kembali langkah-langkah konversinya. Model .tflite yang dihasilkan kemudian akan diuji dan dievaluasi menggunakan data *test* pada tahap selanjutnya.

5. Rancangan Uji Coba Model

Tahap ini akan menguji apakah model yang telah dibuat sebelumnya berhasil mendeteksi dan mengklasifikasi objek SIBI dan bagaimana gambaran hasil deteksi tersebut dengan menampilkan beberapa gambar dari data *test*. Diagram *Flowchart* Rancangan Uji Coba Model ditunjukkan pada Gambar 3.



Gambar 3. *Flowchart* rancangan uji coba model

Pertama membuat fungsi *def* untuk melakukan inferensi dengan model *TFLite* dan menampilkan hasilnya. Selanjutnya, menginput model *.tflite*, label map dengan file *.txt*, dan data *test* yang berisikan file gambar dan file *.xml*. Ketiga ini dibutuhkan untuk proses inferensi. Proses inferensi adalah tahap di mana model yang telah dilatih digunakan untuk membuat prediksi berdasarkan data baru yang tidak pernah dilihat selama pelatihan (data *test*). Dalam konteks ini, proses inferensi akan menampilkan hasil deteksi gestur SIBI dari beberapa gambar pada data *test* secara acak. Hasil deteksi ini akan menunjukkan *bounding box* yang berisi label gestur beserta nilai kepercayaan dari prediksi tersebut.

6. Rencana Evaluasi Model

Evaluasi model dilakukan untuk mengetahui nilai rata-rata akurasi menggunakan mAP (*mean Average Precision*), yang merupakan metrik umum dalam penilaian kinerja model deteksi objek seperti *Faster R-CNN*, *SSD*, dan lain-lain. mAP mengukur seberapa baik model dalam mendeteksi dan mengklasifikasikan objek dalam gambar. Dalam penelitian ini, peneliti menggunakan standar metrik COCO yakni menghitung mAP @ 0.5:0.05:0.95. Peneliti menggunakan alat kalkulator mAP yang bersifat *open source* dari (Cartucho, Ventura and Veloso, 2018). Untuk tahap evaluasi, gambar dan data anotasi dari data *test* akan diduplikat atau disalin ke tempat baru. Data ini akan digunakan sebagai "*ground truth data*", yaitu data referensi yang akan dibandingkan dengan hasil deteksi model. Hasil deteksi model akan melibatkan pemanggilan fungsi *def* yang telah dibuat pada tahap uji coba sebelumnya dan digunakan untuk mendeteksi seluruh gambar yang ada pada data *test*. Hasil dari evaluasi ini berupa nilai akurasi mAP pada tiap *threshold* IoU dalam rentang 0.5 hingga 0.95 dengan kenaikan 0.05, kemudian nilai akurasi mAP masing-masing gestur SIBI (15 gestur), dan nilai akurasi mAP @ 0.5:0.05:0.95 sebagai nilai akurasi model secara keseluruhan.

HASIL DAN PEMBAHASAN

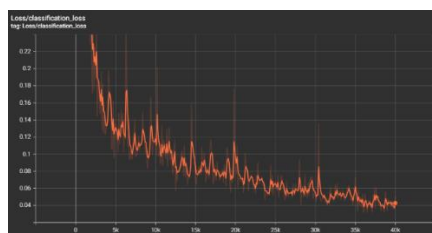
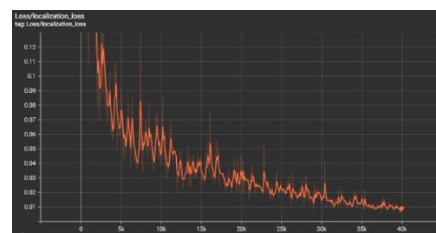
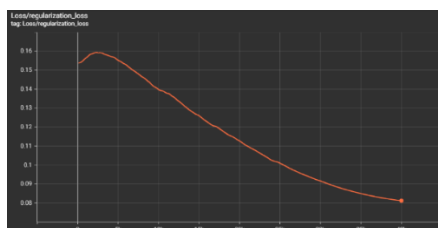
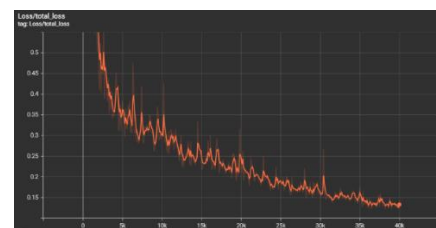
1. Hasil Pelatihan dan Pemodelan

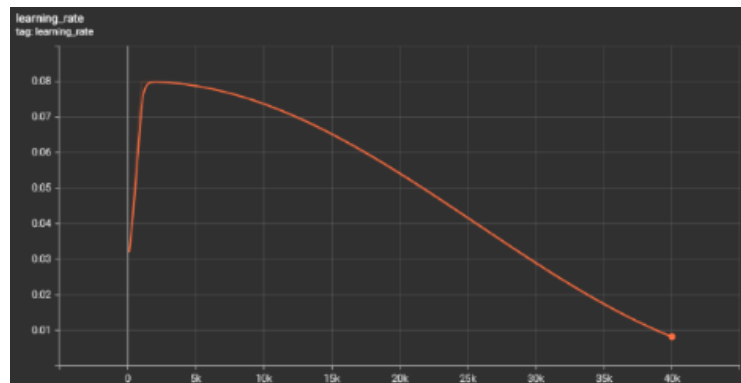
Penelitian menggunakan arsitektur Model *Pre-Trained SSD MobileNet V2 FPNLite* dan menggunakan *Framework TensorFlow Object Detection API* untuk mendeteksi gestur SIBI dengan jumlah *dataset* 15.000 gambar dan 15.000 file *xml*-nya. *Dataset* dibagi menjadi 80% data *train* (12.000), 10% data *validation* (1.500), dan 10% data *test* (1.500). Detail pembagian *dataset* dapat dilihat pada Tabel 2.

Tabel 2. Detail Pembagian Dataset

No.	Gestur SIBI	Data Train	Data Validation	Data Test
1.	A	792	103	105
2.	E	805	93	102
3.	I	807	92	101
4.	O	801	94	105
5.	U	810	102	88
6.	Baik	785	104	111
7.	Bangku	809	104	87
8.	Bel	780	109	111
9.	Dia	833	90	77
10.	Meja	804	95	101
11.	Pramuka	793	104	103
12.	Sakit	781	108	111
13.	Saya	815	95	90
14.	Teman	793	101	106
15.	Tugas	792	106	102
Total		12.000	1.500	1.500

Proses pelatihan menggunakan Model *Pre-Trained SSD MobileNet V2 FPNLite* versi 320x320 dengan jumlah langkah pelatihan 40.000 *steps* dan jumlah langkah gambar yang digunakan per langkah pelatihan (*batch size*) adalah 16. Hasil proses pelatihan model tervisualisasi dalam grafik secara *realtime* melalui TensorBoard, yakni *classification_loss* (Gambar 4), mengukur kesalahan model dalam mengklasifikasikan objek pada gambar, dengan nilai sebesar 0.05082879. *Localization_loss* (Gambar 5), menunjukkan kesalahan dalam penempatan *bounding box* atau posisi objek dalam gambar, tercatat sebesar 0.011469895. *Regularization_loss* (Gambar 6), menunjukkan kesalahan model dalam menggeneralisasi pola untuk setiap objek, memiliki nilai sebesar 0.081098825. *Total_loss* (Gambar 7), gabungan dari semua jenis loss, mencerminkan performa keseluruhan model, dengan nilai 0.14339751. Sementara itu, *learning_rate* (Gambar 8), yang mengukur kecepatan pembaruan parameter selama pelatihan, berada pada angka 0.007943453.

Gambar 4. *classification_loss*Gambar 5. *localization_loss*Gambar 6. *regularization_loss*Gambar 7. *total_loss*



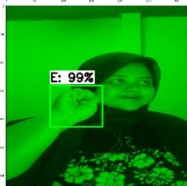
Gambar 8. *learning_rate*

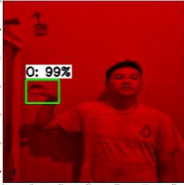
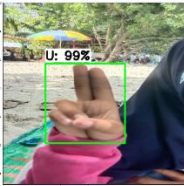
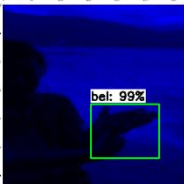
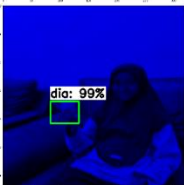
Proses pelatihan model berlangsung sekitar 2 jam 38 menit dan hasil model tersimpan dalam format .pb (*saved_model.pb*) selanjutnya dikonversi ke format *TensorFlow Lite*. File hasil konversi, yaitu *detect.tflite*, tersimpan dalam *Google Drive* dan akan digunakan untuk proses uji coba model.

2. Hasil Uji Coba Model

Fungsi *def tflite_detect_images* dibuat untuk proses inferensi menggunakan model *TFLite* dengan menginput model *.tflite*, *label map* dengan file *.txt*, dan data *test*. Proses inferensi pada tahap uji coba dilakukan dengan menampilkan hasil deteksi gestur SIBI dari beberapa gambar pada data *test* secara acak, dalam kasus ini 100 gambar. Hasil dari 100 gambar secara acak ini sudah memuat deteksi untuk setiap 15 gestur SIBI. Hasil uji coba model gestur SIBI dengan skor kepercayaan tertinggi dapat dilihat pada Tabel 3. Hasil uji coba menunjukkan model yang dibuat oleh peneliti sudah mampu mendeteksi 15 Gestur SIBI dengan berbagai kondisi tertentu yakni, model mampu mendeteksi pada pria maupun wanita, model mampu mendeteksi pada anak-anak, remaja, dan dewasa, model mampu mendeteksi pada sudut depan, kiri, dan kanan, model mampu mendeteksi dalam kondisi *indoor* maupun *outdoor*, model mampu mendeteksi warna dengan intensitas gelap maupun terang ditandai dengan model mampu mendeteksi pada kondisi *bluescale*, *redscale*, *greenscale*, *greyscale*, dan *RGB*, model mampu mendeteksi gestur pada berbagai latar belakang yang memiliki banyak atribut atau bervariasi maupun polos, dan model mampu mendeteksi gambar yang diambil dari kamera *device* yang beragam.

Tabel 3. Hasil Uji Coba Model Gestur SIBI dengan Skor Kepercayaan Tertinggi

No.	Hasil Uji Coba Deteksi	Ket. Gambar	Ket. Deteksi
1.		Gestur A Wanita remaja Sudut depan <i>Indoor</i> <i>Redscale</i>	Gestur A terdeteksi dengan benar dengan skor kepercayaan 99%
2.		Gestur E Wanita remaja Sudut ke kiri <i>Indoor</i> <i>Greenscale</i>	Gestur E terdeteksi dengan benar dengan skor kepercayaan 99%

No.	Hasil Uji Coba Deteksi	Ket. Gambar	Ket. Deteksi
3.		Gestur I Wanita dewasa Sudut depan Indoor RGB	Gestur I terdeteksi dengan benar dengan skor kepercayaan 99%
4.		Gestur O Pria remaja Sudut ke kanan Indoor Redscale	Gestur O terdeteksi dengan benar dengan skor kepercayaan 99%
5.		Gestur U Anak perempuan Sudut depan Outdoor RGB	Gestur U terdeteksi dengan benar dengan skor kepercayaan 99%
6.		Gestur Baik Wanita dewasa Sudut depan Indoor Greyscale	Gestur Baik terdeteksi dengan benar dengan skor kepercayaan 99%
7.		Gestur Bangku Wanita remaja Sudut depan Indoor RGB	Gestur Bangku terdeteksi dengan benar dengan skor kepercayaan 99%
8.		Gestur Bel Anak perempuan Sudut depan Outdoor Bluescale	Gestur Bel terdeteksi dengan benar dengan skor kepercayaan 99%
9.		Gestur Dia Wanita Remaja Sudut ke kanan Indoor Bluescale	Gestur Dia terdeteksi dengan benar dengan skor kepercayaan 99%
10.		Gestur Meja Pria remaja Sudut depan Indoor Greyscale	Gestur Meja terdeteksi dengan benar dengan skor kepercayaan 97%

No.	Hasil Uji Coba Deteksi	Ket. Gambar	Ket. Deteksi
11.		Gestur Pramuka Anak perempuan Sudut depan <i>Outdoor</i> <i>Greenscale</i>	Gestur Pramuka terdeteksi dengan benar dengan skor kepercayaan 99%
12.		Gestur Sakit Anak laki-laki Sudut ke kiri <i>Indoor</i> <i>Greyscale</i>	Gestur Sakit terdeteksi dengan benar dengan skor kepercayaan 99%
13.		Gestur Saya Pria remaja Sudut depan <i>Indoor</i> <i>RGB</i>	Gestur Saya terdeteksi dengan benar dengan skor kepercayaan 99%
14.		Gestur Teman Anak perempuan Sudut ke kiri <i>Indoor</i> <i>RGB</i>	Gestur Teman terdeteksi dengan benar dengan skor kepercayaan 99%
15.		Gestur Tugas Wanita dewasa Sudut ke kanan <i>Outdoor</i> <i>Redscale</i>	Gestur Tugas terdeteksi dengan benar dengan skor kepercayaan 99%

3. Hasil Evaluasi Model

Tahap evaluasi model melibatkan deteksi seluruh gambar yang ada pada data *test*, 1.500 gambar, dengan memanggil fungsi *def tf_lite_detect_images* dengan tidak menampilkan gambar inferensi pada *Google Colaboratory*. Hal ini disebabkan *buffer* memiliki batasan dalam menampilkan *output* gambar pada *interface Google Colab*

Hasil evaluasi menunjukkan nilai $mAP_{IoU0.50}$ 99.93%, $mAP_{IoU0.55}$ 99.93%, $mAP_{IoU0.60}$ 99.93%, $mAP_{IoU0.65}$ 99.93%, $mAP_{IoU0.70}$ 99.93%, $mAP_{IoU0.75}$ 99.66%, $mAP_{IoU0.80}$ 99.35%, $mAP_{IoU0.85}$ 97.60%, $mAP_{IoU0.90}$ 81.07%, $mAP_{IoU0.95}$ 21.95% dan nilai mAP kelas dari total AP kelas tertentu dalam rentang IoU 0.5:0.05:0.9.5 yakni, mAP_A 88.76%, mAP_E 89.90%, mAP_I 91.23%, mAP_O 90.22%, mAP_U 87.97%, mAP_{Baik} 90.65%, mAP_{Bangku} 95.42%, mAP_{Bel} 93.31%, mAP_{Dia} 89.53%, mAP_{Meja} 83.72%, $mAP_{Pramuka}$ 88.59%, mAP_{Sakit} 90.10%, mAP_{Saya} 88.90%, mAP_{Teman} 91.00%, mAP_{Tugas} 89.63%, dimana nilai mAP_{Bangku} menunjukkan nilai tertinggi yakni 95.42% dan mAP_{Meja} menunjukkan nilai terendah yakni 83.72%.

Hasil keseluruhan model $mAP @ 0.5:0.05:0.95$ (*overall*) menunjukkan nilai akurasi model deteksi gestur SIBI yakni 89,93%. Hasil evaluasi model dapat dilihat pada Tabel 4. Hasil nilai akurasi dalam penelitian ini mendapatkan nilai akurasi model yang lebih baik jika dibandingkan dengan nilai akurasi model deteksi SIBI yang dikembangkan oleh (Hidayatullah, 2022), sebelumnya telah dipaparkan pada pendahuluan. Penelitian tersebut

mendeteksi 6 Gestur SIBI (Saya, Dia, Kamu, Cinta, Maaf, dan Sedih) menggunakan Arsitektur Model *Pre-Trained SSD MobileNet V2* dengan jumlah *dataset* sebanyak 600 gambar dan dilatih menggunakan 20.000 *steps* yang menunjukkan hasil nilai akurasi 86,6%, sedangkan dalam penelitian penulis mendeteksi 15 Gestur SIBI (Huruf Vokal A, I, U, E, O dan 10 kata yakni Baik, Bangku, Bel, Dia, Meja, Pramuka, Sakit, Saya, Teman, dan Tugas) menggunakan Arsitektur Model *Pre-Trained SSD MobileNet V2 FPNLite* dengan jumlah *dataset* sebanyak 3.000 gambar dan dilatih menggunakan 40.000 *steps* yang menunjukkan hasil lebih baik yakni 89,93%. Hasil kontribusi penulis dalam penelitian diharapkan menambah pengetahuan sistem deteksi serta dapat membantu melestarikan dan mempopulerkan SIBI sebagai bahasa isyarat nasional Indonesia.

Tabel 4. Hasil Evaluasi Model

Gestur SIBI	Average Precision (AP) Pada Threshold Intersection over Union (IoU)										mAP Gestur
	0.50	0.55	0.60	0.65	0.70	0.75	0.80	0.85	0.90	0.95	
A	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	97.88%	95.49%	73.59%	21.81%	88.76%
E	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	97.58%	86.84%	14.62%	89.90%
I	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	98.90%	92.59%	20.80%	91.23%
O	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	97.50%	71.02%	11.17%	90.22%
U	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	84.24%	18.00%	87.97%
Baik	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	81.25%	25.28%	90.65%
Bangku	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	97.98%	56.26%	95.42%
Bel	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	88.39%	44.73%	93.31%
Dia	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	98.48%	98.48%	85.26%	13.06%	89.53%
Meja	99.01%	99.01%	99.01%	99.01%	99.01%	99.01%	96.74%	88.81%	51.29%	6.26%	83.72%
Pramuka	100.00%	100.00%	100.00%	100.00%	100.00%	97.09%	97.09%	97.09%	75.91%	18.75%	88.59%
Sakit	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	98.01%	83.94%	19.07%	90.10%
Saya	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	98.16%	76.37%	14.45%	88.90%
Teman	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	98.77%	85.82%	25.40%	91.00%
Tugas	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	95.20%	81.49%	19.65%	89.63%
mAP IoU	99.93%	99.93%	99.93%	99.93%	99.93%	99.66%	99.35%	97.60%	81.07%	21.95%	mAP (Overall) 89.93%

KESIMPULAN

Penelitian ini berhasil mengembangkan model deteksi gestur SIBI menggunakan metode *Deep learning* dengan *Framework TensorFlow Object Detection* dan model *pre-trained SSD MobileNet V2 FPNLite*. Dengan *dataset* berisi 15.000 gambar dengan variasi sudut, kamera, warna, latar belakang, jenis kelamin, dan usia, yang dibagi menjadi data *train* (80%), *validation* (10%), dan *test* (10%), serta dilatih menggunakan 40.000 *steps* dan *batch size* 16, menunjukkan hasil nilai *loss* 0.1434 dan hasil uji coba menunjukkan model mampu mendeteksi 15 gestur SIBI. Evaluasi mAP @ 0.5:0.05:0.95 menghasilkan akurasi sebesar 89,93%, dengan mAP tertinggi pada kelas "Bangku" (95.42%) dan terendah pada "Meja" (83.72%). Tindak lanjut penelitian ini bisa diarahkan pada distribusi *dataset* yang lebih proporsional untuk meningkatkan generalisasi model, implementasi *real-time* untuk menguji kinerja dalam kondisi nyata, serta pengembangan *dataset* dengan jumlah yang lebih banyak dengan variasi gestur dan kondisi yang lebih beragam.

DAFTAR PUSTAKA

Bagus, M., Bakti, S. and Pranoto, Y.M. (2019) *Pengenalan Angka Sistem Isyarat Bahasa Indonesia Dengan Menggunakan Metode Convolutional Neural Network*.

-
- Cartucho J, Ventura R and Veloso M (2018) *Robust Object Recognition Through Symbiotic Deep Learning In Mobile Robots*. 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). Available at: <https://github.com/Cartucho/mAP> (Accessed: 3 July 2024).
- Community, L. S. (2015) *Label Studio is a modern, multi-modal data annotation tool*. Available at: <https://github.com/HumanSignal/labelImg?tab=MIT-1-ov-file> (Accessed: 15 August 2024).
- Dionisius, R.A.S., Finsensia Riti, Y. and Christian, N. (2024) 'Perbandingan Performa Model SSD Mobilenet V2 dan FPNLite dalam Deteksi Helm Pengendara Sepeda Motor Performance Comparison of SSD Mobilenet V2 and FPNLite in Motorcycle Rider Helmet Detection', *Jurnal Sistem Komputer*, 13(1), p. 2020.
- Hidayatullah, A. M. (2022) *SISTEM DETEKSI SIMBOL PADA SIBI (SISTEM ISYARAT BAHASA INDONESIA) SECARA REALTIME MENGGUNAKAN MOBILENET-SSD*. Universitas Dinamika.
- Indonesia, P. R. (2016) *Pasal 1 Ayat 1 UU Nomor 8 Tahun 2016*.
- Kahlil (2023) *Computer Vison Berbasis Deep Learning untuk Aplikasi Pertanian: Teori dan Praktik*. Edited by Nana Diana. Syiah Kuala University Press.
- Maulida, S. (2023) *ANALISIS AKURASI PADA SIMBOL ABJAD SISTEM ISYARAT BAHASA INDONESIA (SIBI) DENGAN MENGGUNAKAN METODE CNN DAN YOLO (YOU ONLY LOOK ONCE)*.
- Putra, I.R.W. (2021) *SISTEM DETEKSI SIMBOL PADA SIBI (SISTEM ISYARAT BAHASA INDONESIA) MENGGUNAKAN CONVOLUTIONAL NEURAL NETWORK*.
- Sari, I. and Altiarika, E. (2023) *Sistem Pengembangan Bahasa Isyarat Untuk Berkomunikasi dengan Penyandang Disabilitas (Tunarungu)*, *Journal of Information Technology and society (JITS)*.
- Statistik, B. P. (2023) *buku-i-analisis-tematik-kependudukan-indonesia--fertilitas-remaja-kematian-maternal--kematian-bayi--dan-penyandang-disabilitas-*. Badan Pusat Statistik.
- Suryanah (1996) *Keperawatan Anak Untuk Siswa SPK*. Ester Diana A. Jakarta: Penerbit Buku Kedokteran EGC.
-