

Implementasi Hand Gesture Recognition untuk Bahasa Isyarat Indonesia berbasis MobileNetV2

Radit Julian^{1*}, Hanindya Aisyah², Ardiansyah³

^{*1,2,3} Program Studi Teknologi Informasi, Fakultas Kesehatan dan Teknologi, Universitas Muhammadiyah Klaten, Klaten

Email: ¹raditjulian27@gmail.com, ²haninaisyah514@gmail.com, ³ardiansyah.dsn@gmail.com

ABSTRACT — *Communication is a fundamental human necessity; however, for the deaf community, barriers to interaction with the general public remain a significant challenge due to limited literacy in sign language. This research aims to implement a hand gesture recognition system capable of translating the alphabet of the Indonesian Sign Language System (SIBI) in real-time. The MobileNetV2 architecture was selected as the base model due to its efficiency in processing data on resource-constrained devices without significantly compromising accuracy. The methodology involves several crucial stages, beginning with image pre-processing—including resizing and image normalization—to the application of data augmentation strategies such as rotation, shifting, and brightness adjustment to enhance the model's generalization capabilities in real-world conditions. The dataset comprises SIBI alphabet classifications from A to Z, collected with high variability to minimize the risk of overfitting. The results demonstrate that the use of depthwise separable convolutions in MobileNetV2 allows the system to perform gesture detection with high responsiveness and low computational overhead. Through hyperparameter optimization, this model is expected to achieve optimal accuracy, providing a practical and inclusive communication tool for the deaf community within social environments and public services.*

KEYWORDS — *Hand Gesture Recognition, SIBI, MobileNetV2, Deep Learning, Computer Vision.*

INTISARI — *Komunikasi merupakan kebutuhan fundamental manusia, namun bagi penyandang tunarungu, hambatan interaksi dengan masyarakat umum masih menjadi tantangan besar akibat keterbatasan literasi terhadap bahasa isyarat. Penelitian ini bertujuan untuk mengimplementasikan sistem hand gesture recognition yang mampu menerjemahkan alfabet Sistem Isyarat Bahasa Indonesia (SIBI) secara real-time. Arsitektur MobileNetV2 dipilih sebagai model dasar karena efisiensinya dalam memproses data pada perangkat dengan sumber daya terbatas tanpa mengorbankan akurasi secara signifikan. Metode yang digunakan melibatkan beberapa tahapan krusial, dimulai dari image pre-processing berupa resizing dan normalisasi citra, hingga penerapan strategi data augmentation seperti rotasi, pergeseran (shifting), dan penyesuaian kecerahan untuk meningkatkan kemampuan generalisasi model terhadap kondisi dunia nyata. Dataset yang digunakan mencakup klasifikasi alfabet SIBI dari A hingga Z yang dikumpulkan secara variatif guna meminimalisir risiko overfitting. Hasil penelitian menunjukkan bahwa penggunaan mekanisme depthwise separable convolutions pada MobileNetV2 memungkinkan sistem melakukan deteksi gestur dengan responsivitas tinggi dan beban komputasi yang rendah. Melalui optimasi hyperparameter, model ini diharapkan dapat mencapai tingkat akurasi yang optimal, sehingga dapat diimplementasikan sebagai alat bantu komunikasi yang praktis dan inklusif bagi komunitas Tuli di lingkungan sosial maupun layanan publik.*

KATA KUNCI — *Pengenalan Gerakan Tangan, SIBI, MobileNetV2, Pembelajaran Mendalam, Visi Komputer.*

I. PENDAHULUAN

Komunikasi inklusif bagi komunitas tunarungu di Indonesia saat ini masih menghadapi tantangan besar berupa jurang pemisah antara pengguna Sistem Isyarat Bahasa Indonesia (SIBI) sebagai standar formal di sekolah luar biasa dengan masyarakat umum yang mayoritas sama sekali tidak memahami arti gerakan isyarat tersebut [1]. Akibat minimnya pemahaman ini, penyandang disabilitas rungu sering kali mengalami marginalisasi dan kesulitan besar saat harus mengakses layanan vital di sektor publik, seperti kesehatan, hukum, transportasi, maupun administrasi pemerintahan [2]. Hambatan interaksi sosial yang terus berulang ini pada akhirnya membatasi mobilitas ekonomi mereka dan memperlebar ketimpangan sosial di tengah masyarakat [3].

Untuk menjembatani kesenjangan tersebut, perkembangan teknologi *Computer Vision* berbasis *Artificial Intelligence* (AI)

hadir menawarkan solusi solutif melalui sistem pengenalan gestur tangan (*Hand Gesture Recognition*) otomatis menggunakan kamera ponsel. Pada awalnya, banyak peneliti terdahulu yang mencoba memecahkan masalah ini dengan menerapkan metode *Machine Learning* konvensional seperti *Support Vector Machine* (SVM) atau *K-Nearest Neighbor* (KNN) [4]. Namun, pendekatan jadul ini terbukti memiliki kelemahan besar karena sangat bergantung pada proses ekstraksi fitur gambar yang diatur secara manual, sehingga sistem mudah bingung dan salah tebak.

Kelemahan fatal dari metode tradisional tersebut biasanya langsung terlihat nyata ketika aplikasi diuji di luar area laboratorium, terutama saat dihadapkan pada variasi latar belakang yang ramai atau perubahan intensitas cahaya yang tidak menentu [5]. Guna mengatasi keterbatasan itu, tren riset global kini mulai bergeser memanfaatkan keandalan teknologi

Deep Learning melalui arsitektur *Convolutional Neural Network* (CNN). Metode modern ini dinilai jauh lebih cerdas karena komputer dapat melatih dirinya sendiri secara otomatis untuk mengenali detail lekukan geometri tangan dan artikulasi jari yang rumit tanpa intervensi manual dari manusia [6].

Meski begitu, penggunaan model CNN konvensional generasi lama yang strukturnya sangat besar dan masif, seperti ResNet atau VGG16, justru memunculkan dilema teknis baru yang kurang bersahabat bagi pengguna lapangan [7]. Model-model raksasa tersebut membutuhkan spesifikasi komputer yang tinggi dengan dukungan kartu grafis (GPU) mahal, sehingga jika dipaksakan masuk ke dalam aplikasi ponsel akan membuat sistem menjadi sangat lambat. Efek buruk lainnya adalah aplikasi akan menguras kapasitas memori, membuat baterai ponsel pintar cepat boros, dan menimbulkan panas berlebih pada perangkat [8].

Oleh sebab itu, penelitian ini mengusulkan implementasi arsitektur MobileNetV2 yang sengaja dirancang super ringan dan lincah untuk mengatasi masalah keterbatasan sumber daya pada perangkat seluler. Melalui mekanisme *depthwise separable convolutions*, MobileNetV2 mampu memangkas beban hitungan matematika komputer ponsel secara drastis tanpa harus mengorbankan tingkat akurasi sistem secara keseluruhan [9]. Dengan mengombinasikan kelincahan model ini bersama strategi *image pre-processing* dan *data augmentation* yang matang, penelitian ini diharapkan dapat melahirkan aplikasi penerjemah SIBI yang tangguh, adaptif, dan bekerja secara *real-time* di lingkungan nyata [10].

II. TINJAUAN PUSTAKA

Dalam berbagai jurnal penelitian mengenai bahasa isyarat, para peneliti terdahulu banyak memilih arsitektur *Convolutional Neural Network* (CNN) karena kemampuannya dalam mengekstraksi fitur visual secara otomatis [11]. Di jurnal-jurnal tersebut, CNN utamanya digunakan untuk membedah citra tangan dari piksel mentah hingga menjadi pola geometri yang spesifik. Riset terdahulu membuktikan bahwa CNN sukses besar dalam mengenali perbedaan tipis pada gestur jari tanpa perlu lagi menggunakan hitungan manual yang rumit. Namun, catatan penting dari mayoritas jurnal adalah bahwa CNN standar membutuhkan waktu komputasi yang lama jika lapisannya terlalu dalam [12].

Untuk mengatasi masalah model CNN yang berat, banyak jurnal terbaru yang mulai beralih menguji arsitektur MobileNetV2 [13]. Berdasarkan laporan eksperimen di berbagai literatur, para peneliti memanfaatkan struktur *depthwise separable convolutions* milik MobileNetV2 untuk memotong ukuran berkas model agar bisa masuk ke aplikasi Android atau iOS. Hasil pengujian dari jurnal-jurnal tersebut menunjukkan bahwa MobileNetV2 berhasil memangkas waktu tunggu (*latency*) deteksi pada ponsel pintar secara drastis, sehingga pengenalan gestur isyarat bisa berjalan lancar secara *real-time* dengan akurasi yang tetap terjaga tinggi [14].

Pada bagian metodologi di berbagai jurnal pustaka, tahap *image preprocessing* selalu menjadi langkah wajib yang dilakukan peneliti sebelum memasukkan gambar ke dalam model AI [15]. Aktivitas utama yang mereka lakukan adalah mengubah ukuran resolusi gambar (*resizing*) menjadi seragam, biasanya berukuran 224 x 224 piksel agar pas dengan input

bawaan MobileNetV2 [16]. Selain itu, para peneliti di jurnal tersebut juga menerapkan normalisasi nilai piksel (mengubah rentang warna menjadi 0 hingga 1) untuk mempercepat proses pelatihan model dan mencegah eror akibat perbedaan pencahayaan kamera [17].

Mengenai ketersediaan data, jurnal-jurnal bertema sejenis menjabarkan bahwa mereka umumnya menggunakan dua jenis sumber data, yaitu dataset publik yang sudah tersedia di internet (seperti Kaggle atau Github) atau memproduksi dataset mandiri secara lokal [18]. Untuk konteks Indonesia, para peneliti menjabarkan bahwa dataset mereka berfokus pada 26 kelas alfabet Sistem Isyarat Bahasa Indonesia (SIBI) [19]. Dalam dokumen jurnal mereka, dataset tersebut dijabarkan harus diisi oleh variasi kontributor yang berbeda-beda—mulai dari bentuk tangan pria dan wanita hingga perbedaan warna kulit—agar model yang dilatih tidak bias [20].

Untuk mencegah terjadinya *overfitting* karena jumlah sampel yang terbatas, dokumen jurnal orang lain menjabarkan penggunaan strategi *Data Augmentation* secara masif [21]. Di dalam laporan penelitian mereka, teknik augmentasi yang diterapkan dijabarkan sebagai berikut:

- Rotasi Gambar: Peneliti memutar gambar tangan (misalnya antara 10 hingga 30 derajat) untuk meniru posisi tangan pengguna asli yang sering kali miring di depan kamera ponsel [22].
- Pergeseran Posisi (*Shifting*): Peneliti menggeser posisi objek secara horizontal dan vertikal agar AI terbiasa mendeteksi tangan yang posisinya tidak tepat berada di tengah bingkai layar [23].
- Manipulasi Kecerahan (*Brightness*): Gambar sengaja dibuat menjadi lebih gelap dan lebih terang di komputer guna menyimulasikan kondisi ruangan pengguna di dunia nyata yang pencahayaannya tidak menentu [24].
- Penskalaan Jarak (*Zooming*): Peneliti memperbesar atau memperkecil skala gambar untuk melatih ketahanan model terhadap variasi jarak antara tangan pengguna dengan kamera HP [25].

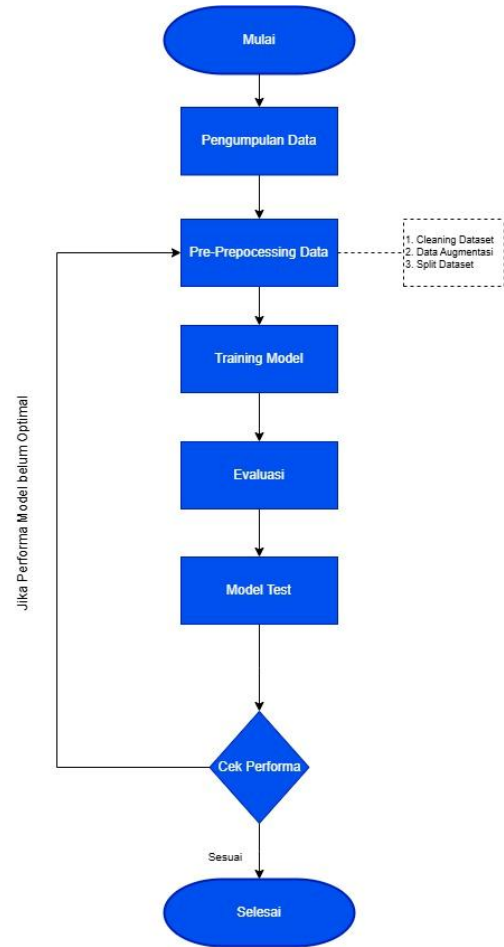
III. METODOLOGI

Penelitian ini merupakan penelitian eksperimental yang berfokus pada pengembangan sistem klasifikasi bahasa isyarat SIBI menggunakan pendekatan Deep Learning berbasis Convolutional Neural Network (CNN) dengan arsitektur MobileNetV2 [26]. Secara garis besar, penelitian ini terdiri dari serangkaian tahapan yang saling berkaitan, dimulai dari persiapan dataset citra bahasa isyarat SIBI, proses preprocessing dan augmentasi data, pembangunan serta pelatihan model MobileNetV2, hingga evaluasi performa dan implementasi sistem deteksi real-time berbasis webcam [27].

Setiap tahapan dirancang secara sistematis untuk memastikan bahwa model yang dihasilkan tidak hanya memiliki akurasi yang tinggi pada data uji, tetapi juga mampu beroperasi secara efisien dalam kondisi nyata. Pemilihan MobileNetV2 sebagai arsitektur utama didasarkan pada pertimbangan efisiensi komputasi dan kemampuan transfer learning dari dataset ImageNet, yang memungkinkan model mempelajari representasi fitur gesture tangan secara lebih cepat dan akurat meskipun dengan dataset berukuran sedang [9].

A. ALUR PENELITIAN

Berikut Gambar 1 merupakan tahapan penelitian yang dilakukan pada sistem SignBridge. Tahapan awal penelitian dimulai dengan pengumpulan dataset bahasa isyarat SIBI yang terdiri dari citra gesture tangan untuk setiap huruf alfabet statis [28]. Setelah dataset terkumpul, dilakukan tahap pre-processing data yang meliputi cleaning dataset, data augmentation, serta splitting dataset menjadi data training, validation, dan testing. Proses preprocessing dilakukan untuk meningkatkan kualitas data dan membantu model mengenali gesture tangan dengan lebih baik pada berbagai kondisi. Dataset yang telah melalui tahap preprocessing kemudian digunakan pada proses training model menggunakan arsitektur MobileNetV2 [21]. Setelah proses training selesai, model akan dievaluasi menggunakan beberapa parameter seperti *accuracy*, *loss*, *confusion matrix*, *precision*, dan *recall* untuk mengetahui performa model dalam mengenali gesture bahasa isyarat. Tahap berikutnya adalah pengujian model menggunakan data testing untuk memastikan model mampu melakukan generalisasi terhadap data baru yang belum pernah dilihat sebelumnya. Jika performa model belum optimal, maka proses akan kembali ke tahap preprocessing dan training dengan melakukan perbaikan seperti penambahan augmentasi data, penyesuaian jumlah epoch, learning rate, batch size, maupun optimizer. Namun apabila model telah mencapai performa yang sesuai, maka model akan diintegrasikan ke dalam sistem Python berbasis Computer Vision menggunakan OpenCV untuk melakukan deteksi gesture tangan secara real-time melalui webcam [29].



Gambar 1. Alur Penelitian

B. DATASET PENELITIAN

Dataset yang digunakan dalam penelitian ini adalah dataset bahasa isyarat SIBI yang terdiri dari citra gesture tangan dalam format gambar statis. Dataset mencakup 24 kelas huruf alfabet (A, B, C, D, E, F, G, H, I, K, L, M, N, O, P, Q, R, S, T, U, V, W, X, dan Y). Huruf J dan Z tidak disertakan karena merupakan gesture dinamis yang memerlukan analisis temporal, di luar cakupan penelitian ini [28]. Setiap kelas memiliki sekitar 220 gambar sehingga total dataset mencapai lebih dari 5.000 gambar. Pembagian dataset dilakukan sebagai berikut:

Tabel I. Splitting Dataset

Subset Data	Persentase	Jumlah Gambar
Training	70%	3514
Validation	20%	1001
Test	10%	500

C. PREPROCESSING DATA

Tahap preprocessing data dilakukan untuk mempersiapkan dataset sebelum digunakan pada proses pelatihan model Deep Learning. Proses ini bertujuan untuk meningkatkan kualitas data, mengurangi noise, serta membantu model dalam mengenali pola gesture tangan dengan lebih optimal [30]. Pada penelitian ini, preprocessing dilakukan menggunakan library TensorFlow dan Keras melalui fitur ImageDataGenerator.

Tahapan preprocessing yang dilakukan meliputi cleaning dataset, resize gambar, rescaling, data augmentation, serta pembagian dataset menjadi data training, validation, dan testing [31].

1. *Cleaning Dataset*

Tahap cleaning dataset dilakukan dengan memeriksa seluruh gambar pada dataset dan menghapus data yang tidak sesuai, seperti gambar rusak, gambar blur berlebihan, gambar duplikat, maupun gambar yang tidak menunjukkan gesture tangan dengan jelas. Proses ini dilakukan agar model hanya mempelajari data yang relevan dan berkualitas baik [32].

2. *Resize Image*

Seluruh gambar pada dataset diubah ukurannya menjadi 224 × 224 piksel agar sesuai dengan ukuran input yang dibutuhkan oleh arsitektur MobileNetV2. Proses resize dilakukan untuk menyeragamkan dimensi gambar sehingga proses training dapat berjalan lebih stabil dan efisien [33].

3. *Rescaling*

Nilai piksel gambar yang awalnya berada pada rentang 0–255 diubah menjadi 0–1 menggunakan proses normalisasi atau rescaling. Tahapan ini dilakukan agar model dapat melakukan proses pembelajaran dengan lebih cepat dan stabil [17].

4. *Data Augmentation*

Data augmentation dilakukan untuk meningkatkan variasi dataset dan mengurangi risiko overfitting pada model. Proses augmentasi dilakukan secara otomatis menggunakan ImageDataGenerator [34]. Beberapa teknik augmentasi yang digunakan meliputi:

- Rotation (rotasi gambar)
- Zoom (perbesaran gambar)
- Horizontal Flip (membalik gambar)
- Width Shift (pergeseran horizontal)
- Height Shift (pergeseran vertikal)

Dengan augmentasi data, model dapat mempelajari berbagai variasi posisi tangan, sudut pengambilan gambar, dan kondisi pencahayaan sehingga performa model menjadi lebih baik saat digunakan secara real-time [35].

5. *Splitting Dataset*

Dataset yang telah diproses kemudian dibagi menjadi tiga bagian utama, yaitu *training* data, *validation* data, dan *testing* data. Pembagian dataset dilakukan dengan persentase 70% untuk *training*, 20% untuk *validation*, dan 10% untuk *testing* [36].

D. *EVALUASI*

Evaluasi model dilakukan untuk mengetahui tingkat performa model MobileNetV2 dalam mengenali gesture bahasa isyarat SIBI [37]. Proses evaluasi dilakukan menggunakan beberapa parameter pengujian untuk mengukur kemampuan model dalam melakukan klasifikasi citra secara akurat. Pada penelitian ini, evaluasi model dilakukan menggunakan metrik *precision* (1), *recall* (2), dan *F1-score* (3) [38]. Persamaan (1)(2)(3) akan membentuk *confusion matrix* seperti pada Gambar 5, serta Tabel II.

Persamaan evaluasi yang digunakan pada penelitian ini antara lain:

$$Precision = \frac{TP}{TP+FP} \tag{1}$$

$$Recall = \frac{TP}{TP+FN} \tag{2}$$

$$F1-Score = 2\chi \frac{Precision \cdot Recall}{Precision+} \tag{3}$$

Keterangan:

- TP (*True Positive*) merupakan jumlah data yang berhasil diprediksi benar sesuai kelas sebenarnya.
- TN (*True Negative*) merupakan jumlah data yang berhasil dikenali bukan termasuk kelas tertentu.
- FP (*False Positive*) merupakan jumlah data yang diprediksi sebagai suatu kelas namun sebenarnya bukan kelas tersebut.
- FN (*False Negative*) merupakan jumlah data yang seharusnya termasuk suatu kelas tetapi diprediksi sebagai kelas lain.

IV. HASIL DAN PEMBAHASAN

A. *HASIL PENELITIAN*

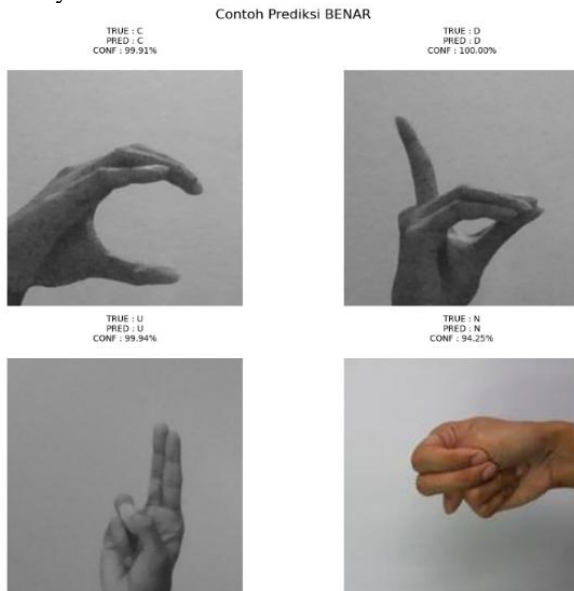
Penelitian ini berhasil membangun sistem deteksi bahasa isyarat SIBI berbasis Deep Learning menggunakan arsitektur MobileNetV2. Sistem dikembangkan untuk mengenali gesture alfabet bahasa isyarat secara *realtime* menggunakan webcam serta pengujian gambar statis. Proses pengembangan dan pelatihan model dilakukan menggunakan platform Google Colaboratory dengan bahasa pemrograman Python serta framework TensorFlow dan Keras. Dataset yang digunakan terdiri dari 24 kelas alfabet SIBI yaitu huruf A sampai Y tanpa huruf J dan Z. Seluruh dataset kemudian melalui tahapan preprocessing berupa *cleaning dataset*, *augmentasi data*, *resizing* gambar menjadi 224×224 piksel, serta pembagian *dataset* menjadi data *training*, *validation*, dan *testing*. Pembagian *dataset* dilakukan dengan rasio 70% data *training*, 20% data *validation*, dan 10% data *testing*. Pada proses pelatihan model digunakan beberapa parameter yaitu *optimizer* Adam, *learning rate* 0.0001, *batch size* 32, serta jumlah epoch sebanyak 100 epoch. Selain itu digunakan *callback ModelCheckpoint* untuk menyimpan model terbaik berdasarkan *validation accuracy* tertinggi selama proses *training* berlangsung.

Tabel II. Accuracy

Class	P	R	F1-Score
A	0.99	0.94	0.96
B	0.87	1.00	0.93
C	0.98	0.99	0.98
D	1.00	0.98	0.99
E	0.89	1.00	0.94
F	0.96	0.92	0.94
G	1.00	0.96	0.98
H	1.00	1.00	1.00
I	0.97	0.97	0.97
K	0.92	1.00	0.96
L	0.97	0.99	0.98
M	0.99	0.83	0.90
N	0.91	0.97	0.94
O	0.98	1.00	0.99
P	1.00	0.96	0.98
Q	1.00	1.00	1.00
R	0.97	0.97	0.97
S	0.94	0.96	0.95
T	0.99	0.95	0.97
U	0.98	0.98	0.98
V	0.95	0.95	0.95

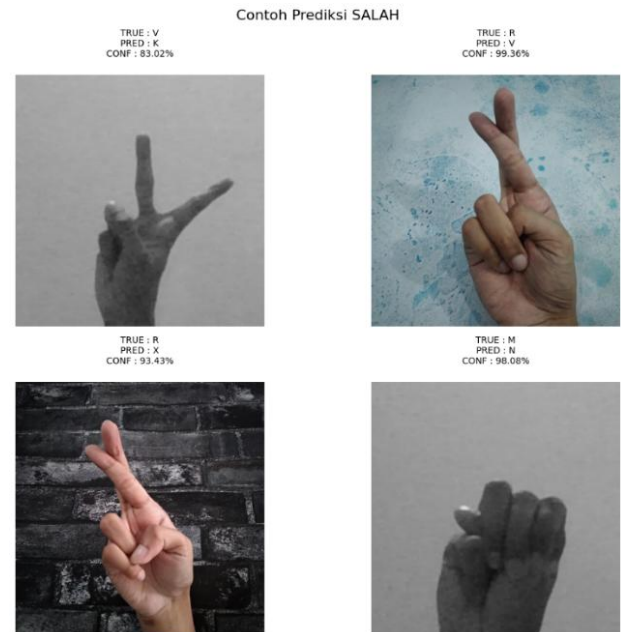
W	0.98	0.93	0.95
X	0.96	0.99	0.97
Y	1.00	0.97	0.99
Accuracy	0.97		

Hasil pelatihan menunjukkan bahwa model mampu mencapai tingkat akurasi yang baik dalam mengenali gesture bahasa isyarat. Model yang telah dilatih kemudian dievaluasi menggunakan data testing untuk mengetahui performa klasifikasi terhadap data baru yang belum pernah dikenali sebelumnya.



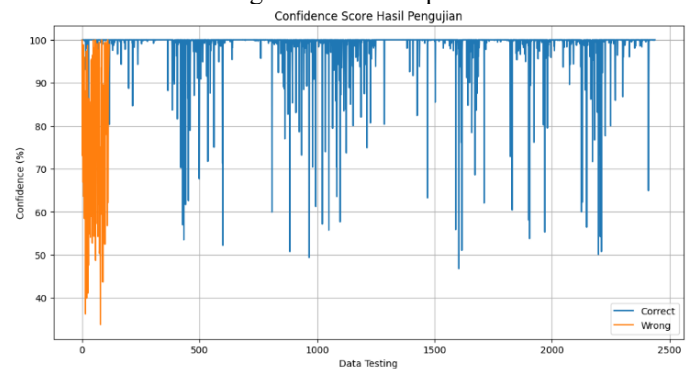
Gambar 2. Contoh Prediksi Benar

Berdasarkan hasil pengujian pada Gambar 3, model berhasil melakukan klasifikasi *gesture* bahasa isyarat dengan baik pada beberapa sampel data *testing*. Seluruh gambar yang diuji mampu diprediksi sesuai dengan label aslinya dengan nilai *confidence* yang tinggi. Hal tersebut menunjukkan bahwa model MobileNetV2 mampu mengekstraksi fitur penting pada bentuk tangan, posisi jari, serta pola *gesture* secara efektif. Selain itu, hasil prediksi yang konsisten menunjukkan bahwa model memiliki kemampuan generalisasi yang cukup baik terhadap data yang digunakan pada proses pengujian. Tingginya *confidence score* pada hasil prediksi benar juga menunjukkan bahwa model memiliki tingkat keyakinan yang tinggi terhadap kelas yang diprediksi.



Gambar 3. Contoh Prediksi Salah

Berdasarkan hasil pengujian pada Gambar 4, masih ditemukan beberapa kesalahan klasifikasi pada *gesture* tertentu. Kesalahan prediksi umumnya terjadi pada *gesture* yang memiliki kemiripan bentuk tangan maupun posisi jari antar kelas sehingga model mengalami kesulitan dalam membedakan karakteristik masing-masing *gesture*. Meskipun beberapa hasil prediksi salah memiliki *confidence score* yang cukup tinggi, kondisi tersebut menunjukkan bahwa model masih mengalami *overconfidence* terhadap pola tertentu yang dianggap mirip dengan data pelatihan. Selain itu, faktor lain seperti pencahayaan, sudut pengambilan gambar, *background*, dan kualitas citra juga memengaruhi performa model dalam melakukan klasifikasi *gesture* secara tepat.

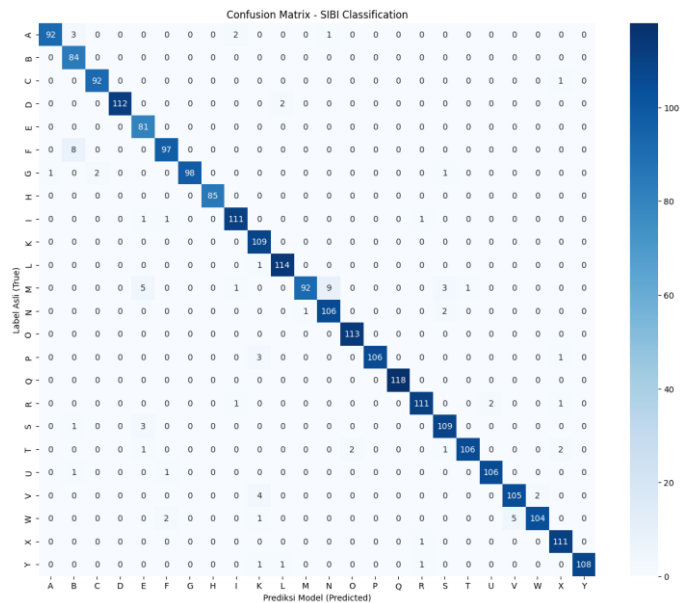


Gambar 4. Grafik Confidence Hasil Test

Berdasarkan grafik *confidence score* hasil pengujian, sebagian besar prediksi benar memiliki tingkat *confidence* tinggi mendekati 100%. Namun demikian, masih ditemukan beberapa prediksi dengan *confidence* rendah maupun prediksi salah dengan *confidence* tinggi. Kondisi tersebut menunjukkan bahwa model masih mengalami *overconfidence* pada beberapa *gesture* yang memiliki karakteristik visual hampir serupa.

Berdasarkan Gambar 5, hasil confusion matrix menunjukkan bahwa sebagian besar data berhasil diklasifikasikan dengan benar oleh model MobileNetV2. Hal ini terlihat dari dominasi nilai pada diagonal utama matriks yang menunjukkan kesesuaian antara label asli dan hasil prediksi model. Tingginya

jumlah prediksi pada diagonal utama mengindikasikan bahwa model mampu mengenali karakteristik *gesture* bahasa isyarat SIBI dengan baik pada sebagian besar kelas yang diuji. Beberapa kelas seperti H, Q, O, K, dan E menunjukkan tingkat klasifikasi yang sangat baik karena hampir seluruh data pada kelas tersebut berhasil diprediksi dengan benar. Kondisi ini menunjukkan bahwa *gesture* pada kelas-kelas tersebut memiliki karakteristik visual yang cukup berbeda sehingga lebih mudah dipelajari dan dikenali oleh model.



Gambar 5. Confusion Matrix

Meskipun demikian, masih ditemukan beberapa kesalahan klasifikasi pada beberapa kelas tertentu. Berdasarkan *confusion matrix*, kesalahan prediksi paling banyak terjadi pada kelas M yang beberapa kali diprediksi sebagai kelas N, serta pada kelas F yang sebagian data diprediksi sebagai kelas B. Selain itu, terdapat pula sejumlah kecil kesalahan klasifikasi pada kelas V, W, dan R yang diprediksi sebagai kelas lain yang memiliki kemiripan bentuk tangan. Kesalahan klasifikasi tersebut diduga terjadi karena adanya kemiripan posisi jari maupun bentuk *gesture* antar kelas. Selain itu, variasi pencahayaan, sudut pengambilan gambar, kualitas citra, dan posisi tangan saat proses pengambilan data juga dapat memengaruhi kemampuan model dalam membedakan fitur visual pada masing-masing *gesture*.

Secara keseluruhan, *confusion matrix* menunjukkan bahwa model MobileNetV2 memiliki kemampuan klasifikasi yang sangat baik dengan jumlah kesalahan prediksi yang relatif kecil dibandingkan total data pengujian. Hasil ini memperkuat temuan pada tahap evaluasi sebelumnya yang menunjukkan bahwa model mampu mencapai tingkat akurasi tinggi dalam mengenali *gesture* bahasa isyarat SIBI.

B. PEMBAHASAN

Berdasarkan hasil penelitian yang telah dilakukan, penggunaan arsitektur MobileNetV2 menunjukkan performa yang cukup baik dalam melakukan klasifikasi *gesture* bahasa isyarat SIBI baik pada pengujian gambar statis maupun *realtime*. MobileNetV2 mampu mengekstraksi fitur penting pada *gesture* tangan sehingga proses klasifikasi dapat dilakukan dengan cepat dan efisien. Hasil *confusion matrix* menunjukkan bahwa model

memiliki kemampuan klasifikasi yang baik karena sebagian besar prediksi berada pada diagonal utama.

Hal tersebut menunjukkan bahwa model berhasil mempelajari pola *gesture* pada masing-masing kelas alfabet. Akan tetapi, masih ditemukan beberapa kesalahan prediksi pada *gesture* tertentu yang memiliki kemiripan bentuk jari maupun posisi tangan. Pada pengujian gambar statis, model mampu menghasilkan *confidence score* tinggi pada sebagian besar data testing. Kondisi tersebut menunjukkan bahwa model memiliki tingkat keyakinan yang baik terhadap hasil prediksi yang diberikan. Namun, ditemukan pula beberapa prediksi salah dengan *confidence score* tinggi.

Hal tersebut menunjukkan bahwa model mengalami *overconfidence* terhadap beberapa *gesture* yang memiliki karakteristik visual hampir serupa. Pengujian *realtime* menunjukkan bahwa sistem mampu melakukan deteksi *gesture* secara langsung menggunakan webcam dengan performa yang cukup stabil. MobileNetV2 memiliki keunggulan dari sisi kecepatan inferensi sehingga cocok diterapkan pada sistem *realtime*.

Meskipun demikian, performa sistem masih dipengaruhi oleh faktor lingkungan seperti pencahayaan, kualitas kamera, *background*, serta posisi tangan pengguna. Secara keseluruhan, sistem deteksi bahasa isyarat yang dikembangkan telah mampu melakukan klasifikasi *gesture* SIBI dengan cukup baik. Penggunaan MobileNetV2 memberikan hasil yang ringan dan efisien sehingga berpotensi diterapkan pada aplikasi *realtime* berbasis desktop maupun perangkat mobile di masa mendatang.

V. KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, sistem deteksi bahasa isyarat SIBI menggunakan arsitektur MobileNetV2 berhasil diimplementasikan dengan baik untuk mengenali *gesture* alfabet secara *realtime* maupun melalui gambar statis. Penggunaan MobileNetV2 terbukti mampu memberikan performa klasifikasi yang cukup baik dengan keunggulan pada efisiensi komputasi dan kecepatan proses inferensi sehingga cocok diterapkan pada perangkat dengan sumber daya terbatas seperti laptop maupun perangkat mobile.

Tahapan *preprocessing* seperti *cleaning dataset*, *resizing*, normalisasi, dan data augmentation juga memberikan pengaruh penting terhadap peningkatan kemampuan model dalam mengenali berbagai variasi *gesture* tangan. Berdasarkan hasil pengujian, sebagian besar *gesture* berhasil dikenali dengan tingkat *confidence* yang tinggi. Hal ini menunjukkan bahwa model mampu mengekstraksi fitur penting pada bentuk tangan dan posisi jari secara efektif.

Meskipun demikian, masih ditemukan beberapa kesalahan klasifikasi pada *gesture* yang memiliki kemiripan bentuk tangan maupun posisi jari. Faktor eksternal seperti pencahayaan, *background*, kualitas kamera, dan sudut pengambilan gambar juga memengaruhi performa sistem dalam kondisi *realtime*. Namun secara keseluruhan, sistem yang dikembangkan telah mampu bekerja dengan cukup stabil dan memiliki potensi untuk diterapkan sebagai media bantu komunikasi bagi penyandang tunarungu dalam kehidupan sehari-hari.

Penelitian selanjutnya diharapkan dapat mengembangkan sistem dengan menambahkan pengenalan *gesture* dinamis seperti huruf J dan Z, meningkatkan jumlah *dataset*, serta mengimplementasikan model pada aplikasi *mobile* Android agar dapat digunakan secara lebih praktis dan luas oleh masyarakat.

REFERENSI

- [1] S. M. Ulfah and S. Ubaidah, "Penerapan bahasa isyarat dalam pembelajaran bagi anak berkebutuhan khusus tuna rungu," *Journal of Dissability Studies and Research (JDSR)*, vol. 2, no. 1, pp. 29–43, 2023.
- [2] F. Mutia, *Akses, informasi dan disabilitas*. Airlangga University Press, 2023.
- [3] R. HAFIZHA AR, "Segregasi Permukiman Sebagai Bentuk Representasi Kesenjangan Sosial Desa Pinggiran Kota (Studi Kasus Desa Natar, Kecamatan Natar, Lampung Selatan)," 2025.
- [4] R. Rustiyana et al., *Artificial Intelligence: Teori dan Penerapan AI di Berbagai Bidang*. PT. Sonpedia Publishing Indonesia, 2025.
- [5] K. Khairunnisa et al., *Image Processing*. PT. Green Pustaka Indonesia, 2025.
- [6] S. R. Dewi, "Deep Learning Object Detection Pada video menggunakan tensorflow dan convolutional neural network," 2018.
- [7] A. N. SYAUQI, "Komparasi Algoritma Convolutional Neural Network Dan Transfer Learning (Resnet 50, MobileNetV2, Dan VGG16) Untuk Identifikasi Ekspresi Wajah Potensi Kejahatan," 2024, *UNIVERSITAS PGRI SEMARANG*.
- [8] A. Khariri, "Aplikasi rekomendasi tablet android terbaik dengan metode weighted product berbasis Android: Studi kasus pada Toko Juragan Teknologi Surabaya," 2013, *Universitas Islam Negeri Maulana Malik Ibrahim*.
- [9] A. Bauravindah, "Deteksi Abnormalitas Toraks yang Efisien Menggunakan Model Deep Learning Mobile: Studi Komparatif MobileNet, ShuffleNet, dan EfficientNet dengan Grad-CAM," 2025, *Universitas Islam Indonesia*.
- [10] N. H. Amir, "Klasifikasi Pengenalan Huruf Sistem Isyarat Bahasa Indonesia (SIBI) Menggunakan Algoritma Attention dan VGG-16," 2025, *Universitas Islam Indonesia*.
- [11] A. Khairina, D. N. Triwibowo, and R. R. Al Hakim, "Klasifikasi Bahasa Isyarat Menggunakan Metode Convolutional Neural Network dengan Arsitektur Mobilenet," *Jurnal Ilmu Komputer dan Teknologi*, vol. 6, no. 2, pp. 65–73, 2025.
- [12] J. Pujoseno, "Implementasi Deep Learning Menggunakan Convolutional Neural Network Untuk Klasifikasi Alat Tulis," 2018.
- [13] W. Z. Ramadhan and H. Al Azies, "Evaluasi Kinerja Arsitektur CNN Berbasis Transfer Learning Xception Dan Mobilenetv2 Untuk Klasifikasi Citra Limbah," *Rabit: Jurnal Teknologi dan Sistem Informasi Univrab*, vol. 11, no. 1, pp. 1128–1139, 2026.
- [14] M. H. Fikri, "Sistem Deteksi Anthracnose Pada Tanaman Cabai Menggunakan Mobilenetv2 Berbasis Android," 2024, *Universitas Islam Sultan Agung Semarang*.
- [15] T. Nurhikmat, "Implementasi deep learning untuk image classification menggunakan algoritma Convolutional Neural Network (CNN) pada citra wayang golek," 2018.
- [16] A. SITI, "Perbandingan Kinerja Arsitektur Mobilenetv2 dan Resnet50 Pada Klasifikasi Tumbuhan Herbal Berbasis Citra Daun," 2025.
- [17] C. Cahyaningtyas et al., *Computer Vision untuk Pemula: Deteksi dan Analisis Ekspresi Wajah dengan CNN*. Uwais Inspirasi Indonesia, 2025.
- [18] R. IRAWAN, "Klasifikasi Bidang Ilmu pada Publikasi Terindeks Google Scholar Menggunakan Metode K-Nearest Neighbor," 2023, *UNIVERSITAS ISLAM SULTAN AGUNG*.
- [19] S. N. Budiman, S. Lestanti, and H. Yuana, *Klasifikasi Alfabet Sistem Isyarat Bahasa Indonesia (SIBI) Menggunakan Computer Vision dan Deep Learning*. Penerbit NEM, 2023.
- [20] A. Adriyendi, *Artificial Intelligence dengan Model Diskriminatif dan Model Generatif (Pendekatan Saintifik Populer pada Riset Kolaboratif)*. Penerbit Widina, 2025.
- [21] I. Rianto and I. P. I. Santosa, *Data Preparation untuk Machine Learning & Deep Learning*. Penerbit Andi, 2025.
- [22] S. PUTRI ARSYAD, "Hubungan Antara Durasi Paparan Layar Smartphone Dengan Forward Head Posture pada Siswa-Siswi SMA Negeri 2 Makassar The Relationship Between Smartphone Screen Time Duration and Forward Head Posture Among Students of SMA Negeri 2 Makassar," 2025, *UNIVERSITAS HASANUDDIN MAKASSAR*.
- [23] L. N. HILDA, "RANCANG BANGUN PENYANGGA SMARTPHONE DENGAN PENGATURAN POSISI OTOMATIS," 2022, *Universitas Andalas*.
- [24] Y. Fitrianto, "Dasar-Dasar Digital Imaging," *Penerbit Yayasan Prima Agus Teknik*, pp. 1–80, 2021.
- [25] A. G. H. Makkasau, "Peningkatan Akurasi Pengawasan Pegawai Video Processing Enhancing the Accuracy of Employee Surveillance through Video Processing," 2024, *Universitas Hasanuddin*.
- [26] R. Fahlevi and C. Rozikin, "Identifikasi isyarat tangan bisindo dengan algoritma cnn dan transfer learning menggunakan mobilenetv2," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 9, no. 4, pp. 6592–6597, 2025.
- [27] F. P. Al Fajri, "Perancangan Model Deep Learning untuk Penerjemah Bahasa Isyarat SIBI menggunakan Transfer Learning MobileNetV2," 2024, *Sekolah Tinggi Teknologi Terpadu Nurul Fikri*.
- [28] A. B. Firdaus, F. S. Atmojo, N. C. Ananda, V. G. Pramudito, and R. D. Irawan, "Deteksi Bahasa Isyarat Indonesia (SIBI) dengan MediaPipe dan Feedforward Neural Network," in *Prosiding Seminar Nasional Teknologi Informasi dan Bisnis*, 2025, pp. 86–93.
- [29] F. Al-Hafizh and A. Alamsyah, "Penerapan Model Deep-CNN Untuk Meningkatkan Akurasi Klasifikasi Bahasa Isyarat Alfabet Menggunakan Algoritma Convolutional Neural Network," *Indonesian Journal of Mathematics and Natural Sciences*, vol. 47, no. 1, pp. 1–11, 2024.
- [30] T. Pramuda and A. H. Mirza, "Implementasi Deep Learning CNN untuk Menerjemahkan Sistem Isyarat Bahasa Indonesia (SIBI) ke Teks," *Dinamik*, vol. 31, no. 1, pp. 304–312, 2026.

-
- [31] I. P. Agus, K. Hidjah, N. Sulistianingsih, G. Hendro, and S. Syahrir, "Implementasi Arsitektur Deep Convolutional Neural Network (CNN) dengan Transfer Learning untuk Klasifikasi Penyakit Kulit," *Jurnal Teknologi Informasi dan Multimedia*, vol. 7, no. 3, pp. 461–477, 2025.
- [32] J. Ken and T. Darmanto, "Rancang Bangun Sistem Virtual Mouse Dengan Hand Gesture Recognition Menggunakan Convolutional Neural Network," *INTEKSIS*, vol. 12, no. 2, pp. 51–60, 2025.
- [33] F. N. Firdaus, "Analisis Kinerja MobileNetV2 dalam Klasifikasi Penyakit Daun Cabai Berdasarkan Citra," in *Seminar Nasional Teknologi & Sains*, 2026, pp. 289–296.
- [34] D. L. Tyas, F. R. Rumambi, A. Patanduk, and R. C. J. Mailangkay, "Klasifikasi Jenis Tumor Otak Melalui Citra MRI dengan Menggunakan Convolutional Neural Network," *Informatik: Jurnal Ilmu Komputer*, vol. 21, no. 1, pp. 26–34, 2025.
- [35] H. Ardiansyah and T. Desyani, "Transfer Learning Menggunakan Model VGG16 untuk Klasifikasi Citra Hewan," *Jurnal Pustaka AI (Pusat Akses Kajian Teknologi Artificial Intelligence)*, vol. 5, no. 2, pp. 441–448, 2025.
- [36] R. Oktafiani, A. Hermawan, and D. Avianto, "Pengaruh komposisi split data terhadap performa klasifikasi penyakit kanker payudara menggunakan algoritma machine learning," *Jurnal Sains dan Informatika*, pp. 19–28, 2023.
- [37] R. R. Maulana, A. Kholiq, and E. Widodo, "Penerapan Convolutional Neural Network Pada Pengenalan Bahasa Isyarat Indonesia Secara Real-Time," *Jurnal Satya Informatika*, vol. 10, no. 2, pp. 114–120, 2025.
- [38] A. Ardiansyah, *Buku Monograf Pengenalan Objek Dalam Bidang Kesehatan*. PT. Sonpedia Publishing Indonesia, 2024.