

Penggunaan Metode *Deep Learning* untuk Pengembangan Sistem Komunikasi Cerdas bagi Penyandang Disabilitas

**NORMA NINGSIH¹, AFIFAH DWI RAMADHANI², DJOKO SANTOSO³,
BINTANG DESTA RAMADHANI⁴, ILYASA AYASY EL GHOFIQI⁵**

^{1,2,4,5}Program Studi Teknologi Rekayasa Internet, PENS, Surabaya, Indonesia

³Program Studi Teknik Telekomunikasi, PENS, Surabaya, Indonesia

Email: norma@pens.ac.id

Received 10 Oktober 2024 | Revised 30 November 2024 | Accepted 24 Desember 2024

ABSTRAK

Komunikasi merupakan kebutuhan mendasar bagi makhluk hidup agar dapat berinteraksi dengan lingkungan sekitar. Di dunia, orang-orang disabilitas khususnya tuna rungu dan sulit mendengar sebagian besar berkomunikasi menggunakan bahasa isyarat. Penelitian ini mengembangkan sistem klasifikasi Bahasa Isyarat Indonesia (SIBI) menggunakan model *Convolutional Neural Network (CNN)* VGG16 dan VGG19 yang diintegrasikan dengan aplikasi berbasis web. Sistem ini dirancang untuk membantu komunikasi dengan penyandang disabilitas melalui klasifikasi gerakan tangan secara real-time menggunakan gambar atau webcam. Hasil pengujian menunjukkan bahwa model mencapai akurasi sebesar 96,4% dengan nilai loss 0,1055, menunjukkan performa yang stabil dan generalisasi yang baik tanpa indikasi overfitting. Evaluasi menggunakan confusion matrix menunjukkan distribusi prediksi yang akurat pada 24 kelas isyarat tangan, dengan *precision*, *recall*, dan *f1-score* yang tinggi untuk setiap kelas. Sistem ini diharapkan dapat menjadi alat bantu komunikasi yang efektif bagi penyandang disabilitas dalam kehidupan sehari-hari.

Kata kunci: *Deep Learning*, VGG16, Klasifikasi, SIBI, Disabilitas

ABSTRACT

Communication is a basic need for living things to interact with their environment. In the world, people with disabilities, especially deaf and hard of hearing, mostly communicate using sign language. This study develops an Indonesian Sign Language (SIBI) classification system using the VGG16 and VGG19 Convolutional Neural Network (CNN) models integrated with a web-based application. This system is designed to assist communication with people with disabilities through real-time hand gesture classification using images or webcams. The test results show that the model achieves an accuracy of 96.4% with a loss value of 0.1055, indicating stable performance and good generalization without any indication of overfitting. A confusion matrix evaluation shows an accurate prediction distribution across 24 hand gesture classes, with high precision, recall, and f1-score for each class. This system is expected to be an effective communication tool for people with disabilities in everyday life.

Keywords: *Deep Learning*, VGG16, Classification, SIBI, Disability

1. PENDAHULUAN

Berkomunikasi adalah salah satu keterampilan fundamental yang memungkinkan manusia untuk berbagi informasi, pemikiran, dan perasaan dengan orang lain. Kemampuan untuk berkomunikasi dengan efektif memegang peranan penting dalam berbagai aspek kehidupan, seperti hubungan sosial, pekerjaan, pendidikan, dan bahkan kesehatan mental. Melalui komunikasi, kita bisa mengekspresikan ide, emosi, dan sudut pandang dengan cara yang bisa dipahami orang lain **(El-Din and El-Ghany, 2020)**. Menurut perkiraan dari Organisasi Kesehatan Dunia (WHO), sekitar 2,5 miliar individu diperkirakan akan mengalami gangguan pendengaran pada berbagai tingkat pada tahun 2050, dengan sekitar 700 juta orang membutuhkan layanan rehabilitasi pendengaran **(WHO, 2024)**. Di seluruh dunia, orang-orang tuna rungu dan sulit mendengar sebagian besar berkomunikasi menggunakan bahasa isyarat **(Izzah and Suciati, 2014)**.

Angka ini terus meningkat seiring dengan bertambahnya populasi dan harapan hidup, serta meningkatnya prevalensi penyakit kronis yang sering berujung pada disabilitas. Kondisi ini menyebabkan jutaan individu menghadapi tantangan dalam berkomunikasi, berinteraksi sosial, dan mengakses layanan yang diperlukan untuk menjalani kehidupan yang mandiri dan produktif. Dengan kemajuan teknologi, seperti pengembangan perangkat bantu, aplikasi berbasis AI, dan solusi berbasis komunikasi digital, banyak hambatan yang dapat diatasi **(Barto and Sutton, 2015)**. Teknologi tidak hanya membantu mengatasi keterbatasan fisik, tetapi juga memberikan peluang baru bagi penyandang disabilitas untuk berpartisipasi penuh dalam masyarakat. Kecerdasan buatan (*artificial intelligence* atau AI) memiliki peran besar dalam kemajuan teknologi informasi di berbagai bidang. AI digunakan untuk mendukung dan mempermudah berbagai aktivitas melalui teknologi digital **(Latif, dkk, 2023) (Zawacki-Richter, dkk, 2019)**

Deep learning dengan kemampuannya untuk memahami dan memproses data kompleks seperti suara dan gambar, dapat digunakan untuk mengembangkan sistem pengenalan bahasa isyarat yang lebih canggih dan akurat. Salah satu metode *deep learning* yang digunakan untuk mendeteksi, mengklasifikasi dan mengidentifikasi objek dalam gambar adalah CNN (*convolutional neural network*) **(Dwi Nurhayati, Eridani, and Hafiz Tsalavin, 2022)**. Pada **(Alrowais, dkk, 2023)** Mengembangkan sistem yang disebut Bayesian Optimization with *Deep Learning-Driven Hand Gesture Recognition Based Sign Language Communication* (BODL-HGRSLC) untuk penyandang disabilitas. Hal ini bertujuan untuk mengenali isyarat tangan dengan mengintegrasikan *computer vision* dan *deep learning*. Metode *resnet* digunakan untuk ekstraksi fitur. Hasil yang diperoleh yaitu akurasi maksimu sebesar 99.75%.

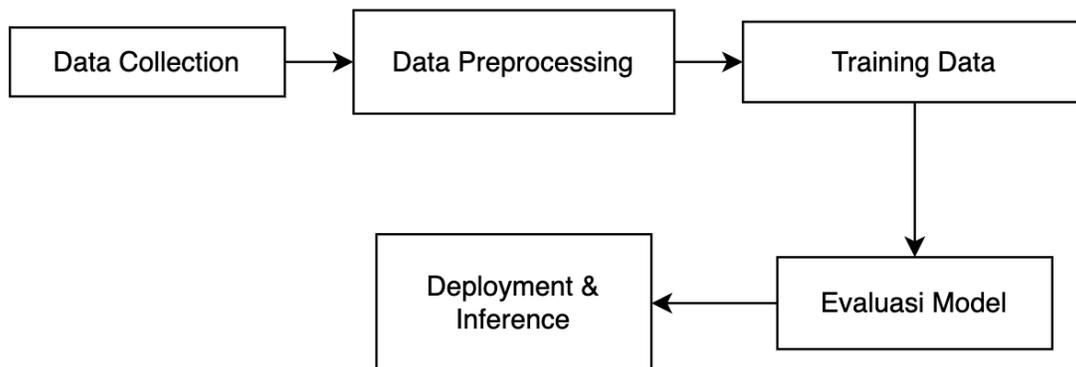
Penelitian lain mengenai *Hand Gesture Recognition System* menggunakan webcam untuk tracking gerakan tangan memanfaatkan metode *Region of Interest (ROI)* dan algoritma *Kernelized Correlation Filters (KCF)*, dengan pengaturan ukuran gambar sebesar 100x120 piksel. Sistem pengenalan gesture tangan ini bekerja secara efektif dengan menggabungkan tiga komponen utama, yaitu deteksi tangan, tracking tangan, dan pembacaan gerakan tangan. Selain itu, terdapat metode berbasis *deep learning* menggunakan *Convolutional Neural Network (CNN)* dengan arsitektur AlexNet dan VGGNet yang telah dimodifikasi. Pendekatan ini mampu mencapai akurasi pelatihan sebesar 99,90% dan akurasi pengujian sebesar 95,61% **(Chung, Chung, and Tsai, 2019)**

Penelitian terkait lainnya dilakukan oleh **(Obi, dkk, 2022)**. Pada penelitian ini membuat aplikasi desktop yang dapat mengenali bahasa isyarat dan mengubahnya menjadi teks secara real time.

Penelitian ini menggunakan *dataset American Sign Language (ASL)* dan *Convolutional* Sistem klasifikasi Jaringan Syaraf Tiruan (CNN). Dalam klasifikasi, gambar tangan terlebih dahulu melewati filter dan setelah filter diterapkan, tangan melewati pengklasifikasi yang memprediksi kelas gerakan tangan. Pada penelitian ini menghasilkan akurasi sebesar 96,3% untuk 26 huruf alfabet. Terdapat pula penelitian SIBI dengan memanfaatkan metode machine learning dan deep learning untuk melakukan klasifikasi citra (Naufal & Kusuma, 2023).

penelitian ini bertujuan untuk mengembangkan sistem komunikasi yang interaktif dan efisien guna membantu penyandang disabilitas berkomunikasi secara lebih mudah dengan lingkungan mereka. Dengan memanfaatkan arsitektur *Convolutional Neural Network (CNN)* seperti VGG16 dan platform *streamlit*, sistem ini diharapkan dapat memproses input data dengan akurasi yang tinggi. Kontribusi spesifik dari penelitian ini adalah untuk menghadirkan solusi berbasis teknologi yang dapat diakses dan digunakan oleh masyarakat luas, sehingga dapat meningkatkan kualitas hidup individu dengan keterbatasan komunikasi.

2. METODOLOGI



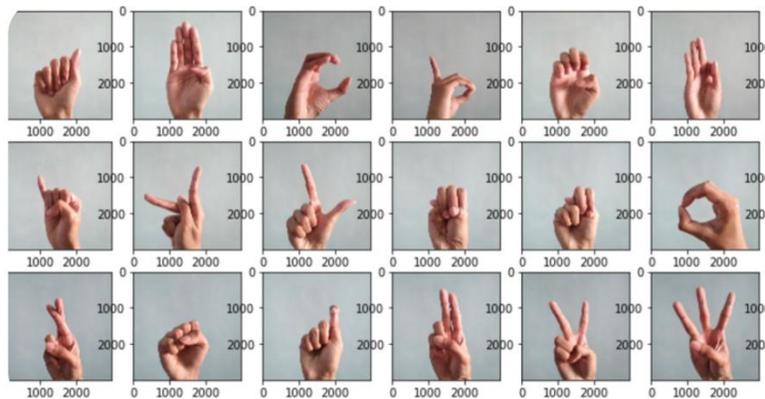
Gambar 1. Blok Diagram Sistem

Pada Gambar 1. menunjukkan blok diagram yang menggambarkan alur keseluruhan dari sistem klasifikasi bahasa isyarat untuk membantu berkomunikasi dengan penyandang disabilitas. Diagram ini terdiri dari beberapa tahapan yaitu tahapan pertama adalah *data collection*, di mana data bahasa isyarat SIBI dikumpulkan. Data ini berupa data gambar yang menunjukkan berbagai isyarat tangan yang merepresentasikan huruf, kata, atau kalimat dalam bahasa SIBI. Tahap kedua adalah *data preprocessing*, yaitu Setelah data dikumpulkan, data tersebut diproses terlebih dahulu sebelum digunakan dalam model. Tahap ketiga adalah *Training data*, Setelah data siap digunakan, data tersebut dibagi menjadi data pelatihan dan data validasi. Data pelatihan digunakan untuk melatih model CNN (*Convolutional Neural Network*). Pada tahap ini, model belajar untuk mengenali pola-pola spesifik dari data isyarat SIBI yang diberikan. Tahap keempat adalah evaluasi model, Model yang telah dilatih dievaluasi menggunakan data validasi atau data uji. Tahap evaluasi ini dilakukan untuk menilai kinerja model dalam hal akurasi, presisi, *recall*, dan metrik evaluasi lainnya. Hasil evaluasi ini digunakan untuk menentukan apakah model sudah siap untuk digunakan atau perlu dilakukan penyempurnaan lebih lanjut. Tahap terakhir adalah *deployment* dan *inference*, model diimplementasikan pada sistem nyata untuk melakukan inferensi secara *real-time*. Pengguna dapat memasukkan input bahasa isyarat melalui upload

gambar atau kamera secara *realtime*, dan model akan mengklasifikasikan isyarat tersebut menjadi teks berupa huruf.

2.1. Pengumpulan Data

Dataset yang digunakan pada penelitian ini adalah dataset SIBI yang dapat dicari secara bebas di internet khususnya pada kaggle. Dataset SIBI berisi 5280 data gambar yang terdiri dari 24 kelas yang merepresentasikan huruf A hingga Y. masing-masing kelas terdiri dari 220 data gambar.



Gambar 2. Dataset Gambar (Bintang, 2022)

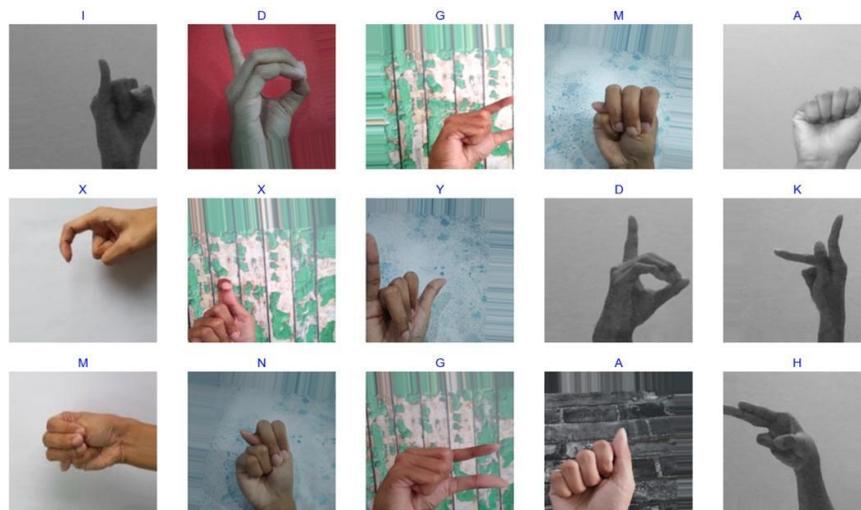
Bahasa Isyarat adalah metode komunikasi yang menggunakan gerakan tangan (gestur) dan dipersepsi melalui indra penglihatan untuk mengenali serta mendapatkan informasi. Bahasa ini umum digunakan oleh penyandang disabilitas tuli atau tunarungu sebagai sarana berkomunikasi. Bahasa SIBI (Sistem Isyarat Bahasa Indonesia) adalah sistem bahasa isyarat formal yang dirancang untuk mewakili bahasa Indonesia dalam bentuk visual. SIBI menggunakan gerakan tangan dan posisi jari untuk menyusun kata-kata dan kalimat, memungkinkan penyandang disabilitas tuli atau tunarungu untuk berkomunikasi dengan lebih mudah. Sistem ini diresmikan oleh Kementerian Pendidikan dan Kebudayaan pada tahun 1997 dan merupakan adaptasi dari *American Sign Language (ASL)*. SIBI memiliki aturan dan tata bahasa tersendiri yang mengacu pada bahasa Indonesia, termasuk penggunaan tangan kanan untuk menunjukkan huruf alfabet.

2.2. Persiapan Data

Tahap yang penting dalam pengembangan sistem klasifikasi berbasis *machine learning* yaitu mengubah data mentah menjadi bentuk yang dapat digunakan dan diinterpretasikan oleh model. langkah-langkah *preprocessing* yang digunakan pada penelitian ini adalah *resize* gambar yaitu semua gambar diubah ukurannya menjadi dimensi yang seragam yaitu 128x128 piksel dengan 3 kanal warna (RGB). Penskalaan ini diperlukan agar ukuran input konsisten dengan dimensi input yang diperlukan oleh model. Hal ini juga bertujuan untuk mengurangi beban komputasi tanpa kehilangan informasi penting dari gambar. selain itu dilakukan juga Pseudocoloring yaitu Teknik ini diterapkan untuk meningkatkan kontras pada gambar. Pseudocoloring dilakukan dengan mengubah gambar skala abu-abu (*grayscale*) menjadi gambar RGB dengan tujuan memberikan informasi tambahan dalam bentuk representasi warna untuk meningkatkan kemampuan model dalam membedakan pola isyarat.

Penggunaan Metode Deep Learning untuk Pengembangan Sistem Komunikasi Cerdas Bagi Penyandang Disabilitas

Tahap *preprocessing* data yang dilakukan selanjutnya adalah augmentasi data. Augmentasi diterapkan untuk memperluas jumlah data pelatihan tanpa perlu menambah data baru. Ini dilakukan dengan menerapkan beberapa transformasi pada data yang ada seperti *rescale* yaitu semua nilai piksel pada gambar dinormalisasi dengan membaginya dengan nilai 255. Hal ini dilakukan untuk mengubah rentang nilai piksel dari $[0, 255]$ menjadi $[0, 1]$, yang dapat membantu model konvergen lebih cepat dan stabil selama pelatihan. *Rotation range* yaitu gambar diputar secara acak dengan sudut hingga ± 10 derajat untuk mensimulasikan variasi sudut pandang. *Width shift dan height shift* merupakan gambar digeser secara acak secara horizontal (lebar) dan vertikal (tinggi) hingga 10% dari ukuran asli untuk membuat model lebih tahan terhadap variasi posisi isyarat pada gambar. *Zoom range* yaitu gambar di-*zoom in* atau *zoom out* hingga 20% untuk memberikan variasi skala isyarat. *Shear range* merupakan distorsi geser diterapkan hingga 20% untuk mengatasi perubahan perspektif. *Brightness range* merupakan kecerahan gambar diubah secara acak antara 80% hingga 120% dari kecerahan asli untuk mensimulasikan variasi kondisi pencahayaan. dan terakhir adalah *fill mode* yaitu pada saat augmentasi menghasilkan area kosong, mode pengisian 'nearest' digunakan untuk mengisi area kosong tersebut dengan nilai piksel tetangga terdekat. Gambar 2 menunjukkan dataset setelah dilakukan proses augmentasi.



Gambar 3. Dataset Gambar setelah proses Augmentasi

Tabel 1. Persiapan Data

No	Persiapan Data	Nilai
1	Pseudocoloring	-
2	Rezise	128x128x3
3	Augmentasi	<i>rescale=1./255,</i> <i>rotation_range=10</i> <i>width_shift_range=0.2</i> <i>height_shift_range=0.2</i> <i>zoom_range=0.2</i> <i>shear_range=0.2</i> <i>brightness_range=[0.8,1.2]</i> <i>fill_mode='nearest'</i>

2.3. Pelatihan Data

Sebelum dilakukan *training data*, Data ini dipisahkan menjadi tiga set yaitu data latih (*training data*) dimana digunakan untuk melatih model terdiri dari 80% dari total data. Data validasi digunakan untuk mengevaluasi model selama pelatihan dan untuk tuning parameter hyperparameter terdiri dari 10% dari keseluruhan data. Data uji digunakan untuk mengevaluasi kinerja model setelah pelatihan selesai dan tidak digunakan dalam proses pelatihan model terdiri dari 10% dari keseluruhan data. Dalam proses pelatihan data performa model dipengaruhi oleh nilai *hyperparameter* (Soekarta, Aras, and Ahmad Nur Aswad, 2023). *Hyperparameter* digunakan untuk mengontrol perilaku algoritma pembelajaran mesin dan dapat secara signifikan mempengaruhi kinerja dan hasil akhir dari model. Nilai *hyperparameter* yang digunakan dalam penelitian ini dapat dilihat pada Tabel 2.

Tabel 2. Persiapan Data

Parameter	Nilai
<i>Epoch</i>	50
<i>Learning rate</i>	0.001
<i>Batch size</i>	64
<i>Optimizer</i>	Adam
<i>Fungsi Loss</i>	<i>Cross Entropy</i>

Pada penelitian ini menggunakan *transfer learning* VGG16, *transfer learning* merupakan pendekatan ini memanfaatkan pengetahuan yang telah dipelajari oleh model pada dataset yang lebih besar atau lebih umum untuk diterapkan pada dataset baru yang lebih spesifik atau lebih kecil (Weiss, Khoshgoftar, and Wang, 2016). Model ini menggunakan pendekatan transfer learning dengan memanfaatkan arsitektur VGG16 yang telah dilatih sebelumnya (pre-trained) sebagai model dasar (*base model*). Selanjutnya menambahkan lapisan konvolusi, pooling, dropout, dan *dense* untuk menyesuaikan model dengan tugas klasifikasi gambar yang spesifik. Model ini dibuat untuk melakukan klasifikasi terhadap 24 kelas dengan menggunakan softmax pada lapisan output. Model dikompilasi dengan *optimizer Adam* dan fungsi kerugian *categorical cross-entropy* untuk memulai pelatihan.

2.4. Testing dan Evaluasi Model

Pada tahap ini, 10% dari dataset total digunakan untuk mengujia model klasifikasi CNN dengan arsitektur VGG16. untuk menghitung nilai performa dari model yang telah dibuat menggunakan *cross validation* dimana hal ini merupakan teknik evaluasi kinerja model dalam pembelajaran mesin yang digunakan untuk menilai seberapa baik model generalisasi terhadap data yang belum pernah dilihat selama pelatihan (Wijiyanto, dkk, 2024). Tujuannya adalah untuk menghindari masalah *overfitting* dan *underfitting* dengan memberikan gambaran yang lebih baik tentang bagaimana model akan berkinerja pada data baru yang tidak terlihat. *Metric* performa yang digunakan adalah *accuracy, precision, recall, F1 Score*, dan waktu komputasi total training dan testing. Rumus perhitungan *metric performa accuracy, precision, recall, dan F1 Score* untuk setiap algoritma klasifikasi dapat dilihat secara berturut turut pada Persamaan

(1), (2), (3), dan (4). TP adalah jumlah *True Positive*, TN adalah jumlah *True Negative*, FP adalah jumlah *False Positive*, dan FN adalah jumlah *False Negative*.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

$$F1\ Score = \frac{Precision \times Recall}{Precision+Recall} \quad (4)$$

2.5. Deployment Model dan Inference

Deployment Model adalah proses penempatan model pembelajaran mesin (*machine learning*) yang telah dilatih ke dalam lingkungan produksi agar dapat digunakan oleh aplikasi atau pengguna akhir untuk membuat prediksi atau inferensi. *Inference* adalah penggunaan model yang telah dilatih untuk menghasilkan prediksi berdasarkan data baru. Salah satu cara untuk membuat model ini dapat diakses secara interaktif oleh pengguna adalah dengan menggunakan *Streamlit*, sebuah *framework python* yang mudah digunakan untuk membuat aplikasi web interaktif khusus untuk aplikasi data dan pembelajaran mesin. Pada penelitian ini memanfaatkan *framework streamlit* untuk melakukan *deployment* model berbasis *website*.

3. HASIL DAN PEMBAHASAN

Dalam pengembangan sistem klasifikasi bahasa isyarat dilakukan menggunakan beberapa *software* yaitu untuk pengembangan model pembelajaran mesin dan analisis data, beberapa alat dan *software* yang digunakan meliputi *Python*, *Visual Studio Code* (VSCode), atau *Jupyter Notebook* sebagai lingkungan pengembangan. TensorFlow dan keras digunakan sebagai library utama untuk pembuatan dan pelatihan model *deep learning*. *NumPy* dan *pandas* merupakan *library fundamental* untuk analisis data. *Scikit-Learn* digunakan sebagai *library* yang menyediakan berbagai algoritma pembelajaran mesin serta fungsi untuk pemrosesan data seperti *split*, *scaling*, dan *cross-validation*. *Matplotlib* dan *seaborn* adalah *library* visualisasi data yang membantu dalam pemahaman pola dan distribusi data melalui grafik yang informatif.

Pada tahap *deployment* dan interaksi pengguna, *Streamlit* digunakan untuk membangun aplikasi web yang ringan dan interaktif. Ini memungkinkan integrasi model pembelajaran mesin dengan antarmuka yang mudah digunakan oleh pengguna non-teknis. Dalam hal pemrosesan gambar, *OpenCV* (cv2) dan *PIL* (*Python Imaging Library*) digunakan untuk membaca, memanipulasi, dan memproses gambar. *BytesIO* adalah modul yang memungkinkan manipulasi data biner dalam memori, sangat berguna untuk operasi *input/output file*. Sementara itu, *MediaPipe* digunakan sebagai *framework* yang dirancang oleh *google* untuk pemrosesan sinyal media (termasuk deteksi gestur dan pelacakan wajah), yang digunakan dalam aplikasi yang membutuhkan analisis gerak dan fitur visual.

Penggunaan Metode Deep Learning untuk Pengembangan Sistem Komunikasi Cerdas Bagi Penyandang Disabilitas

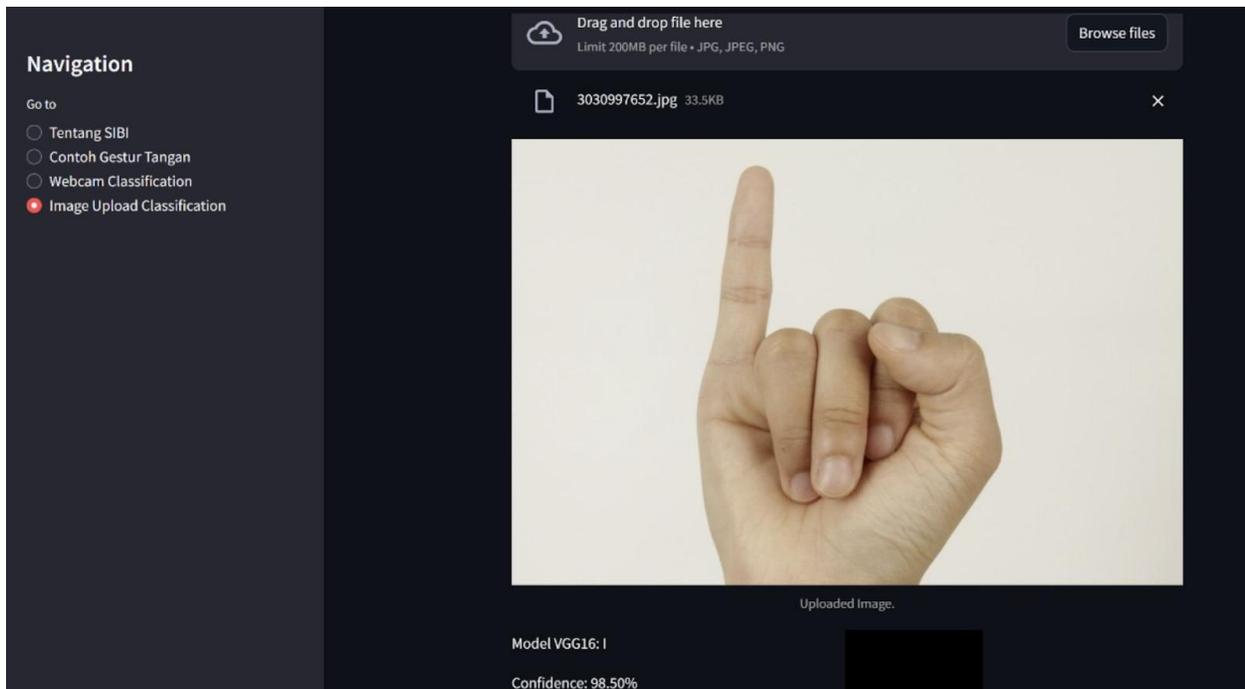


Gambar 4. Tampilan Aplikasi Klasifikasi Bahasa Isyarat

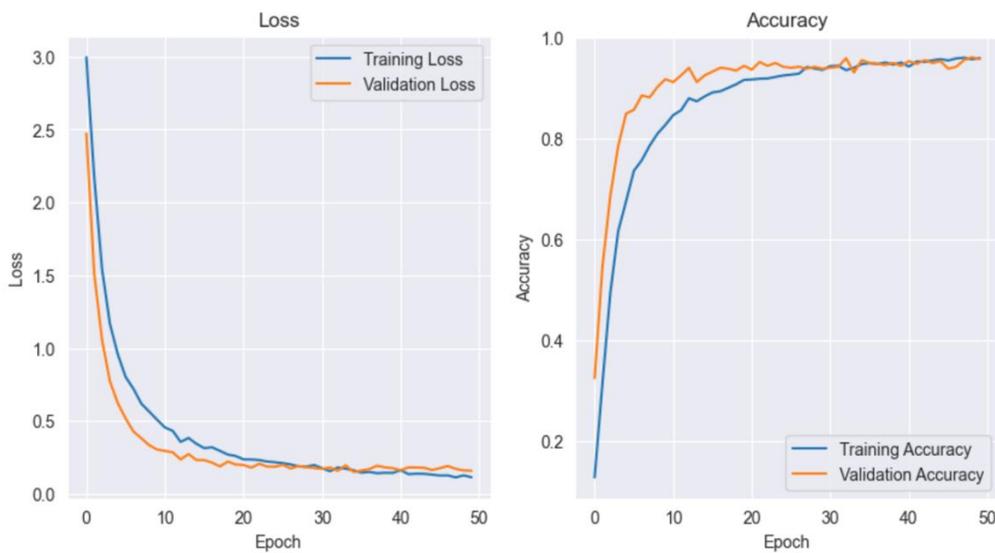
Pada Gambar 4 menunjukkan tampilan antarmuka aplikasi klasifikasi Sistem Isyarat Bahasa Indonesia (SIBI) yang digunakan untuk membantu komunikasi bagi penyandang disabilitas, khususnya tuli atau tunarungu. Pada bagian kiri layar, terdapat navigasi dengan opsi menuju "Landing Page," "Webcam Classification," dan "Image Upload Classification". Di bagian utama layar, aplikasi memperkenalkan dataset yang digunakan, menjelaskan bahwa bahasa isyarat diproduksi melalui gerakan tangan (gestur) yang diidentifikasi melalui penglihatan. Gambar ini juga menampilkan contoh gestur tangan yang mewakili huruf alfabet dalam SIBI, yaitu huruf A, B, C, dan D, yang digambarkan dengan foto tangan membentuk setiap huruf tersebut. Aplikasi ini dirancang untuk membantu pengguna memahami dan mengenali gestur bahasa isyarat dengan lebih mudah, baik melalui klasifikasi gambar yang diunggah maupun melalui penggunaan webcam secara langsung.

Pada Gambar 5 menampilkan antarmuka aplikasi klasifikasi bahasa isyarat yang sedang dalam mode "Image Upload Classification." Pada bagian kiri layar, terdapat menu navigasi dengan beberapa opsi, yaitu "Tentang SIBI," "Contoh Gestur Tangan," "Webcam Classification," dan "Image Upload Classification," yang sedang aktif. Pada bagian kanan, terlihat area untuk mengunggah gambar dengan batas ukuran hingga 200MB dalam format JPG, JPEG, atau PNG. Gambar yang ditampilkan adalah hasil unggahan foto gestur tangan yang mewakili huruf "I" dalam bahasa isyarat. Sistem kemudian mengklasifikasikan gestur ini menggunakan model VGG16 dengan tingkat kepercayaan (*confidence*) sebesar 98.50%. Hal ini menunjukkan bahwa model telah berhasil mengenali dan mengidentifikasi gestur yang diunggah dengan akurasi yang sangat tinggi, hal ini membantu pengguna untuk memastikan bahwa gestur tersebut sesuai dengan huruf yang dimaksud dalam bahasa isyarat SIBI.

Penggunaan Metode Deep Learning untuk Pengembangan Sistem Komunikasi Cerdas Bagi Penyandang Disabilitas



Gambar 5. Tampilan Aplikasi Hasil Klasifikais Bahasa Isyarat



Gambar 6. Hasil Akurasi

Pada Gambar 6 menunjukkan grafik yang ditampilkan terdiri dari dua bagian utama: grafik *loss* di sebelah kiri dan grafik akurasi di sebelah kanan, yang masing-masing menggambarkan kinerja model selama proses pelatihan. Grafik *loss* memperlihatkan penurunan nilai *loss* pada sumbu Y yang mewakili kesalahan prediksi model, sedangkan sumbu X menunjukkan jumlah *epoch* atau berapa kali seluruh data pelatihan digunakan untuk melatih model. Garis biru menunjukkan nilai *loss* pada data pelatihan yang konsisten menurun seiring bertambahnya *epoch*, menandakan bahwa model semakin mampu meminimalkan kesalahan prediksi pada data pelatihan. Garis

oranye menunjukkan nilai *loss* pada data validasi yang juga menurun secara stabil, menunjukkan bahwa model mampu melakukan generalisasi dengan baik terhadap data yang belum pernah dilihat. Penurunan *training loss* dan *validation loss* yang konsisten tanpa divergensi menunjukkan bahwa model tidak mengalami *overfitting* atau *underfitting* yang signifikan, dengan nilai *loss* akhir mendekati 0.1 yang menunjukkan tingkat kesalahan rendah baik pada data pelatihan maupun validasi.

Grafik akurasi di sebelah kanan menggambarkan peningkatan akurasi model pada sumbu Y, yang mengindikasikan persentase prediksi yang benar, dan sumbu X menunjukkan jumlah epoch yang sama dengan grafik *loss*. Garis biru menunjukkan akurasi model pada data pelatihan yang awalnya rendah namun meningkat dengan cepat dan stabil seiring bertambahnya epoch. Garis oranye menunjukkan akurasi pada data validasi yang juga mencapai nilai tinggi sejak awal pelatihan dan tetap stabil, menunjukkan kemampuan model untuk melakukan generalisasi yang baik terhadap data baru. Grafik ini menunjukkan bahwa model mencapai akurasi tinggi pada data pelatihan dan validasi, sekitar 96.4%, tanpa tanda *overfitting* yang signifikan. Akurasi pada data validasi konsisten dengan akurasi pelatihan, yang menunjukkan bahwa model tidak hanya mengingat data pelatihan, tetapi juga memahami pola yang relevan. Secara keseluruhan, model berhasil mencapai akurasi tinggi sekitar 96.4% dengan nilai *loss* yang rendah sekitar 0.1055 pada data pengujian. Kurva *loss* dan akurasi menunjukkan proses pelatihan yang stabil tanpa tanda-tanda *overfitting* atau *underfitting*, dan kinerja generalisasi model sangat baik karena akurasi validasi stabil dan konsisten dengan akurasi pelatihan. Kesimpulannya, model ini menunjukkan performa yang sangat baik dalam menyelesaikan tugas klasifikasi dengan akurasi tinggi pada dataset validasi dan pengujian. Model berhasil menemukan keseimbangan yang baik antara kompleksitas dan akurasi tanpa mengalami masalah *overfitting*, yang merupakan tantangan umum dalam pelatihan model *deep learning*.

Berdasarkan Gambar 7 menunjukkan performa model, evaluasi dilakukan menggunakan metrik seperti *precision*, *recall*, dan *f1-score* untuk masing-masing kelas serta metrik keseluruhan. *Precision* mengukur ketepatan prediksi model terhadap kelas tertentu, di mana nilai *precision* yang tinggi menunjukkan bahwa prediksi model sangat akurat untuk kelas yang diprediksi. Misalnya, ketika model memprediksi suatu kelas, hasilnya cenderung benar dengan tingkat kesalahan yang minim. *Recall*, di sisi lain, mengukur kemampuan model dalam mengidentifikasi semua *instance* dari kelas tertentu. Nilai *recall* 1.00 untuk banyak kelas, seperti A, B, dan C, menunjukkan bahwa model tidak melewatkan *instance* apapun dari kelas tersebut, sehingga mampu mendeteksi seluruh *instance* yang seharusnya terklasifikasi ke dalam kelas tersebut. *F1-score*, yang merupakan rata-rata harmonis dari *precision* dan *recall*, memberikan gambaran yang seimbang antara ketepatan prediksi dan sensitivitas model untuk setiap kelas, menggambarkan performa model secara keseluruhan dengan lebih adil.

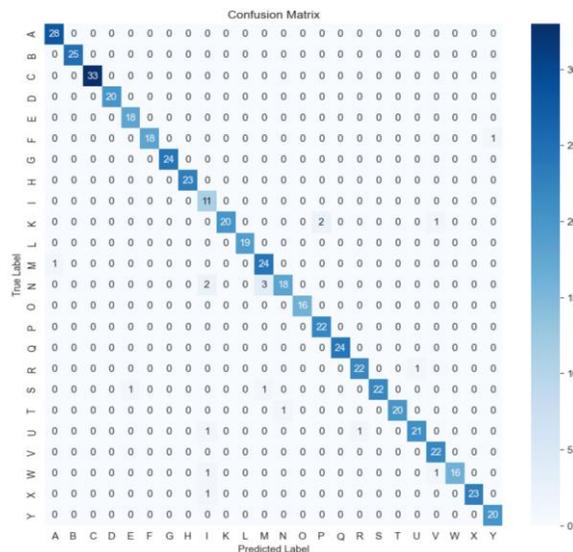
Secara keseluruhan, model mencapai akurasi sebesar 96%, yang berarti model mampu mengklasifikasikan dengan benar sebanyak 96% dari total *instance* yang ada. *Macro average* untuk *precision*, *recall*, dan *f1-score* masing-masing bernilai 0.96, yang menunjukkan bahwa rata-rata performa setiap kelas sangat baik, tanpa ada kelas yang memiliki performa yang jauh lebih buruk dibandingkan yang lain. Selain itu, *weighted average* untuk *precision*, *recall*, dan *f1-score* masing-masing sebesar 0.97, 0.96, dan 0.96, mengindikasikan bahwa model tidak hanya berkinerja baik pada rata-rata per kelas, tetapi juga memperhitungkan distribusi *instance* dalam penghitungan rata-rata performa, menjadikan evaluasi ini relevan terutama ketika kelas memiliki

Penggunaan Metode Deep Learning untuk Pengembangan Sistem Komunikasi Cerdas Bagi Penyandang Disabilitas

jumlah instance yang berbeda. Dengan demikian, model menunjukkan performa yang seimbang dan akurat dalam mengklasifikasikan data dengan distribusi yang beragam, menandakan kemampuannya dalam mempelajari dan mengenali pola dengan baik pada setiap kelas.

	precision	recall	f1-score	support
A	0.97	1.00	0.98	28
B	1.00	1.00	1.00	25
C	1.00	1.00	1.00	33
D	1.00	1.00	1.00	20
E	0.95	1.00	0.97	18
F	1.00	0.95	0.97	19
G	1.00	1.00	1.00	24
H	1.00	1.00	1.00	23
I	0.69	1.00	0.81	11
K	1.00	0.87	0.93	23
L	1.00	1.00	1.00	19
M	0.86	0.96	0.91	25
N	0.95	0.78	0.86	23
O	1.00	1.00	1.00	16
P	0.92	1.00	0.96	22
Q	1.00	1.00	1.00	24
R	0.96	0.96	0.96	23
S	1.00	0.92	0.96	24
T	1.00	0.95	0.98	21
U	0.95	0.91	0.93	23
V	0.92	1.00	0.96	22
W	1.00	0.89	0.94	18
X	1.00	0.96	0.98	24
...				
accuracy			0.96	528
macro avg	0.96	0.96	0.96	528
weighted avg	0.97	0.96	0.96	528

Gambar 7. Hasil *Confusion Matrix*



Gambar 8. Tabel *Confusion Matrix*

Pada Gambar 8 menunjukkan hasil confusion matrix yang menggambarkan distribusi prediksi model terhadap label sebenarnya dari 24 kelas yang berbeda, mulai dari A hingga X. Dalam matriks ini, setiap baris mewakili label sebenarnya, sedangkan setiap kolom menunjukkan label prediksi yang dihasilkan oleh model. Elemen pada diagonal *confusion matrix*, yang ditandai dengan *highlight* biru, menunjukkan jumlah *instance* yang diprediksi dengan benar untuk setiap

kelas. Misalnya, kelas A memiliki 28 *instance* yang diprediksi dengan benar, yang mencerminkan kesesuaian total antara prediksi dan label sebenarnya untuk kelas tersebut. Hal ini menunjukkan bahwa model bekerja dengan sangat baik dalam mengenali dan memisahkan kelas-kelas secara akurat.

Selain itu, tidak adanya angka selain nol pada elemen-elemen di luar diagonal menunjukkan bahwa model tidak membuat kesalahan dalam memprediksi antara kelas yang berbeda. Dengan kata lain, tidak ada *instance* dari suatu kelas yang salah diprediksi sebagai kelas lain, yang merupakan indikasi kuat dari kemampuan model untuk menghindari *misclassifications*. Skala warna pada matriks memberikan visualisasi intensitas prediksi yang benar, dengan warna biru yang lebih gelap menunjukkan jumlah prediksi benar yang lebih tinggi. Ini mempertegas bahwa model memiliki performa yang sangat baik, mampu mengklasifikasikan *instance* dari semua kelas secara akurat dan konsisten, tanpa kebingungan antara kelas yang berbeda.

```
# Add additional layers on top of the base model
x = base_model.output
x = Conv2D(filters=32, kernel_size=(3, 3), padding='same', activation='relu')(x)
x = MaxPooling2D(pool_size=(2, 2), strides=2)(x)
x = Dropout(rate=0.5)(x)

x = Flatten()(x)
x = Dense(64, activation='relu')(x)
outputs = Dense(24, activation='softmax')(x)

# Create the model
model = Model(inputs=inputs, outputs=outputs)

# Compile the model
model.compile(optimizer=tf.keras.optimizers.Adam(learning_rate=0.001),
              loss='categorical_crossentropy',
              metrics=['accuracy'])
```

Gambar 9. Kode CNN VGG16

Pada Gambar 9 menunjukkan gambar implementasi kode model jaringan saraf tiruan menggunakan Keras dengan TensorFlow untuk klasifikasi multi-kelas. Model ini memanfaatkan *transfer learning* dengan menggunakan output dari *base model* yang sudah dilatih sebelumnya, lalu menambahkan lapisan konvolusi (Conv2D) dengan 32 filter, lapisan pooling (MaxPooling2D) untuk mengurangi dimensi, serta dropout untuk mencegah overfitting. Setelah data di-*flatten*, dua lapisan dense diterapkan. Satu dengan 64 neuron dan aktivasi ReLU, dan lapisan keluaran dengan 24 neuron serta aktivasi softmax untuk klasifikasi multi-kelas. Model dikompilasi menggunakan optimizer Adam dengan learning rate 0.001, loss categorical crossentropy, dan akurasi sebagai metrik evaluasi.

4. KESIMPULAN

Berdasarkan analisis dari *confusion matrix* dan metrik evaluasi lainnya, kesimpulan yang dapat diberikan adalah bahwa model klasifikasi yang digunakan memiliki performa yang sangat baik dalam mengenali dan membedakan 24 kelas yang berbeda. Model menunjukkan akurasi keseluruhan sebesar 96%, yang menunjukkan kemampuan yang kuat dalam

mengklasifikasikan *instance* dengan benar. *Confusion matrix* menunjukkan bahwa hampir tidak ada kesalahan prediksi antara kelas yang berbeda. Hal ini dibuktikan dengan nol pada elemen-elemen di luar diagonal, yang berarti model secara konsisten dapat memprediksi label dengan benar tanpa membingungkan kelas satu dengan yang lain.

UCAPAN TERIMA KASIH

Kami mengucapkan terima kasih yang sebesar-besarnya kepada Pusat Penelitian dan Pengabdian kepada Masyarakat (PPM) Politeknik Elektronika Negeri Surabaya (PENS) yang telah memberikan dukungan finansial dalam pelaksanaan penelitian ini. Bantuan yang diberikan oleh PPM PENS sangat berperan penting dalam terselenggaranya penelitian ini, mulai dari tahap perencanaan hingga penyelesaian. Kami berharap hasil dari penelitian ini dapat memberikan kontribusi positif dan menjadi langkah awal yang baik untuk penelitian-penelitian selanjutnya.

DAFTAR RUJUKAN

- Alrowais, Fadwa, Radwa Marzouk, Fahd N. Al-Wesabi, and Anwer Mustafa Hilal. (2023). Hand Gesture Recognition for Disabled People Using Bayesian Optimization with Transfer Learning. *Intelligent Automation and Soft Computing*, 36(3), 3325–42. <https://doi.org/10.32604/iasc.2023.036354>.
- Barto, Andrew. G, and Richard.S Sutton. (2015). *Reinforcement Learning: An Introduction*. 2nd ed. London, England: The MIT Press.
- Bintang, ALvian. (2022). Sistem Isyarat Bahasa Indonesia (SIBI). <https://www.kaggle.com/datasets/alvinbintang/sibi-dataset>. 2022.
- Chung, Huang-Yuan, Yao-Liang Chung, and Wei-Feng Tsai. (2019). An Efficient Hand Gesture Recognition System Based on Deep CNN. In *IEEE International Conference on Industrial Technology (ICIT)*, (pp. 853–58). Melbourne, VIC, Australia.
- Dwi Nurhayati, Oky, Dania Eridani, and Muhammad Hafiz Tsalavin. (2022). Sistem Isyarat Bahasa Indonesia (Sibi) Metode Convolutional Neural Network Sequential Secara Real Time A Real-Time Indonesian Language Sign System Using The Convolution Neural Network Method. *Jurnal Teknologi Informasi Dan Ilmu Komputer (JTIK)*, 9(4), 819–28. <https://doi.org/10.25126/jtik.202294787>.
- El-Din, Salma A Essam, and Mohamed A. Abd El-Ghany. (2020). Sign Language Interpreter System: An Alternative System for Machine Learning. In *Proceedings of the 2nd Novel Intelligent and Leading Emerging Sciences Conference*, (pp. 332–37). Novel Intelligent and Leading Emerging Sciences Conference.

- Farid Naufal, Mohammad, and Selvia Ferdiana Kusuma. (2023). Analisis Perbandingan Algoritma Machine Learning Dan Deep Learning Untuk Klasifikasi Citra Sistem Isyarat Bahasa Indonesia (Sibi). *Jurnal Teknologi Informasi Dan Ilmu Komputer (JTIK)*, 10, 873–81. <https://doi.org/10.25126/jtiik.2023106828>.
- Izzah, Abidatul, and Nanik Suciati. (2014). Translation of Sign Language Using Generic Fourier Descriptor and Nearest Neighbour. *International Journal on Cybernetics & Informatics*, 3(1), 31–41. <https://doi.org/10.5121/ijci.2014.3104>.
- Latif, Abdul, Erna Apriani, Muhamad Syahwildan, and Pupung Purnamasari. (2023). Development of Artificial Intelligence (AI): A Bibliometric Analysis Approach. *Jurnal Info Sains*, 13(3), 1031–43. <http://ejournal.seaninstitute.or.id/index.php/InfoSains>.
- Obi, Yulius, Kent Samuel Claudio, Vetri Marvel Budiman, Said Achmad, and Aditya Kurniawan. (2022). Sign Language Recognition System for Communicating to People with Disabilities. In *Procedia Computer Science*, 216, 13–20. Elsevier B.V. <https://doi.org/10.1016/j.procs.2022.12.106>.
- Soekarta, Rendra, Suhardi Aras, and Ahmad Nur Aswad. (2023). "Hyperparameter Optimization of CNN Classifier for Music Genre Classification. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 7(5), 1205–10. <https://doi.org/10.29207/resti.v7i5.5319>.
- Weiss, Karl, Taghi M. Khoshgoftaar, and Ding Ding Wang. (2016). A Survey of Transfer Learning. *Journal of Big Data*, 3(1). <https://doi.org/10.1186/s40537-016-0043-6>.
- WHO. (2024). Deafness and Hearing Loss. <https://www.who.int/news-room/fact-sheets/detail/deafness-and-hearing-loss>. February 2, 2024.
- Wijiyanto, Wijiyanto, Afu Ichsana Pradana, Sopingi Sopingi, and Vihi Atina. (2024). Teknik K-Fold Cross Validation Untuk Mengevaluasi Kinerja Mahasiswa. *Jurnal Algoritma*, 21(1). <https://doi.org/10.33364/algoritma/v.21-1.1618>.
- Zawacki-Richter, Olaf, Victoria I. Marín, Melissa Bond, and Franziska Gouverneur. (2019). Systematic Review of Research on Artificial Intelligence Applications in Higher Education – Where Are the Educators?. *International Journal of Educational Technology in Higher Education*. Springer Netherlands. <https://doi.org/10.1186/s41239-019-0171-0>.