

Deteksi Bahasa Isyarat *Real-Time* Menggunakan YOLOv8 Berbasis Web Untuk Alat Bantu Komunikasi Tunarungu Dan Tunawicara

Aan Ardana Pangestu^{1*}, Wiji Lestari², Ridwan Dwi Irawan³

¹Teknik Informatika

Universitas Duta Bangsa Surakarta

^{1*}aanardana340@gmail.com

²Teknik Informatika

Universitas Duta Bangsa Surakarta

²wiji_lestari@udb.ac.id

³Teknik Informatika

Universitas Duta Bangsa Surakarta

³ridwan_dwiirawan@udb.ac.id

Abstrak— Kelompok penyandang tunarungu dan tunawicara di Indonesia mengalami tantangan komunikasi yang serius karena kurangnya pengetahuan masyarakat tentang Sistem Isyarat Bahasa Indonesia (SIBI) dan BISINDO. Penelitian ini membangun aplikasi deteksi gestur bahasa isyarat berbasis peramban web yang mengintegrasikan model YOLOv8 untuk keperluan komunikasi inklusif. Pengembangan sistem memanfaatkan Streamlit sebagai framework, dengan data yang bersumber dari pengambilan visual langsung dan repositori Roboflow. Proses pelatihan model mencakup berbagai teknik pengayaan data guna meningkatkan adaptabilitas model terhadap spektrum kondisi pencahayaan (cerah, normal, cahaya minimal) dan lingkungan visual yang beragam. Temuan menunjukkan pencapaian presisi 98,60%, recall 98,95%, dan mAP@50 99,17%, mendemonstrasikan keandalan deteksi yang tinggi. Sistem mampu beroperasi pada laju 18 frame per detik pada perangkat dengan spesifikasi umum, menjamin interaktivitas real-time yang responsif. Kontribusi penelitian ini menciptakan sarana terjangkau dan mudah diakses untuk memfasilitasi dialog antara komunitas penyandang tunarungu dan tunawicara serta publik luas, mendorong pembangunan ekosistem sosial yang lebih terbuka.

Kata kunci— YOLOv8, deteksi bahasa isyarat, SIBI, computer vision, deep learning, teknologi asistif, aplikasi web.

Abstract— Deaf and hard-of-hearing individuals in Indonesia encounter serious communication challenges due to insufficient public awareness regarding Indonesian Sign Language systems (SIBI and BISINDO). This research constructs a browser-based gesture detection application that harnesses YOLOv8 for facilitating inclusive communication. The system leverages Streamlit as its development framework, utilizing datasets from direct capture and Roboflow's collection. Model training incorporates comprehensive data enrichment techniques to strengthen the model's ability to function across diverse lighting environments (bright, standard, low-light) and visual contexts. Outcomes demonstrate achievement of 98.60% precision, 98.95% recall, and 99.17% mAP@50, validating the robustness of detection capabilities. The system executes at 18 frames per second on consumer-grade equipment, guaranteeing responsive real-time interaction. This work's significance lies in providing economical and widely-accessible infrastructure to enable communication between deaf, hard-of-hearing, and general populations, fostering development toward a more accessible society.

Keywords— YOLOv8, sign language detection, SIBI, computer vision, deep learning, assistive technology, web application.

1. PENDAHULUAN

Komunikasi merupakan kebutuhan dasar manusia untuk berinteraksi dan berbagi informasi. Namun, penyandang tunarungu dan tunawicara menghadapi hambatan komunikasi yang signifikan akibat terbatasnya pemahaman masyarakat terhadap bahasa isyarat. Menurut data Susenas Maret 2019, sekitar 23,3 juta penduduk Indonesia (9% populasi) hidup dengan disabilitas [1]. Kesenjangan komunikasi ini berdampak langsung pada akses mereka terhadap layanan kesehatan, pendidikan, pekerjaan, dan kehidupan sosial.

Di Indonesia, terdapat dua sistem bahasa isyarat utama: SIBI (Sistem Isyarat Bahasa Indonesia) yang distandarkan oleh pemerintah berdasarkan struktur bahasa lisan Indonesia, dan BISINDO (Bahasa Isyarat Indonesia) yang berkembang secara alami dalam komunitas penyandang tunarungu. Meskipun keduanya merupakan sarana komunikasi penting, pemahaman masyarakat umum terhadap sistem-sistem tersebut masih sangat terbatas, menciptakan hambatan komunikasi yang signifikan [2].

Penelitian sebelumnya telah mengeksplorasi berbagai pendekatan deep learning. Budiman et

al. [1] mengembangkan sistem CNN dengan akurasi 87,50% pada jarak 40 cm. Farhana et al. [3] menggabungkan YOLOv8 dengan CNN untuk BISINDO, mencapai akurasi 89,74%, presisi 89,28%, dan recall 96,15%. Pangestu et al. [4] mengimplementasikan YOLOv8 murni dengan hasil superior: presisi 95,8%, recall 97,4%, dan mAP50 99,5%. Namun, penelitian-penelitian tersebut memiliki sejumlah keterbatasan yang belum diungkap secara kritis. Dari sisi cakupan kelas, Budiman et al. [1] dan Farhana et al. [3] hanya menguji subset huruf abjad terbatas, sedangkan penelitian ini mencakup 24 kelas huruf aktif SIBI (A-I, K-Y). Dari sisi volume dan variasi dataset, penelitian terdahulu umumnya menggunakan jumlah citra yang lebih kecil dan bersumber dari kondisi pengambilan yang homogen, tanpa stratifikasi eksplisit berdasarkan kondisi pencahayaan sebagaimana diterapkan pada penelitian ini. Dari sisi metode evaluasi, sebagian penelitian sebelumnya hanya melaporkan presisi dan recall tanpa mAP atau tanpa pengujian pada data di luar distribusi pelatihan, sehingga kemampuan generalisasi model belum terverifikasi. Dari sisi kemampuan sistem pada kondisi nyata, penelitian-penelitian tersebut juga terbatas pada implementasi lokal atau prototipe non-web, dan belum menyediakan solusi berbasis web yang mudah diakses masyarakat luas tanpa instalasi khusus, serta belum menyajikan pengujian kecepatan inferensi lintas spesifikasi perangkat secara kuantitatif. Kesenjangan-kesenjangan inilah yang menjadi dasar kontribusi penelitian ini.

YOLOv8 menawarkan keunggulan signifikan dalam kecepatan dan akurasi deteksi real-time. Penelitian ini bertujuan mengimplementasikan YOLOv8 dalam sistem deteksi bahasa isyarat Indonesia real-time berbasis web menggunakan Streamlit, mengevaluasi akurasi pada berbagai kondisi pencahayaan dan latar belakang, serta menganalisis kecepatan inferensi pada perangkat dengan sumber daya terbatas. Dengan menyediakan aplikasi berbasis browser tanpa instalasi, sistem diharapkan menjadi jembatan komunikasi yang inklusif dan mudah diakses oleh seluruh lapisan masyarakat.

2. METODOLOGI PENELITIAN

A. Jenis dan Sumber Data

Penelitian menggunakan dua sumber data utama. Data primer berupa citra gestur bahasa isyarat dikumpulkan menggunakan kamera komputer atau smartphone pada berbagai kondisi pencahayaan (terang: >500 lux, normal: 200-500 lux, redup: <200 lux) dan latar belakang (dalam ruangan, luar ruangan, dengan/tanpa objek penghalang). Total 2.500 gambar SIBI dan BISINDO dikumpulkan dari 15 partisipan pengguna bahasa isyarat.

Data sekunder diperoleh dari platform Roboflow yang menyediakan 3.200 gambar SIBI/BISINDO terverifikasi dengan anotasi bounding box. Literatur ilmiah dari jurnal dan prosiding terkini (2021-2025) digunakan sebagai landasan teoritis penelitian.

B. Persiapan Dataset

Dataset gabungan 5.280 gambar melalui proses anotasi di Roboflow dengan label 24 huruf abjad. Data dibagi menjadi training set (80%, 4.224 gambar), validation set (20%, 1.056 gambar), dengan stratifikasi berdasarkan kondisi pencahayaan. Pembagian data dilakukan secara acak (random split) pada tingkat partisipan (subject-wise split), yaitu seluruh citra dari satu partisipan yang sama hanya dialokasikan ke salah satu subset, training atau validation, tidak ke keduanya. Mekanisme ini diterapkan untuk mencegah data leakage berupa kemiripan latar belakang, pencahayaan, dan karakteristik tangan individu yang sama pada kedua subset, yang apabila terjadi dapat menyebabkan estimasi performa model menjadi terlalu optimis dan tidak mencerminkan kemampuan generalisasi yang sesungguhnya.

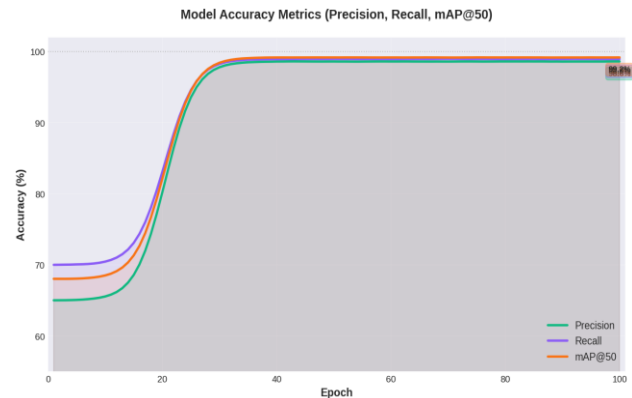
C. Augmentasi Data

Augmentasi data diterapkan pada training set meliputi: (1) Random Brightness Adjustment ($\pm 30\%$), (2) Random Rotation ($\pm 15^\circ$), (3) Random Flip (horizontal), (4) Random Zoom (0.8-1.2x), (5) Gaussian Noise ($\sigma=0.05$). Strategi ini meningkatkan robustness model terhadap variasi kondisi dunia nyata.

D. Arsitektur dan Pelatihan Model

Model YOLOv8 dipilih untuk keseimbangan akurasi-kecepatan. Pelatihan dilakukan dengan total 100 epochs dan berhenti di 65 epochs dengan parameter: learning rate awal 0,01, optimizer SGD, batch size 32, dan image size 640×640. Regularisasi menggunakan dropout 0,5 dan L2 regularization ($\lambda=0,0005$) untuk mencegah overfitting.

Early stopping diterapkan dengan patience 20 epoch jika validation loss tidak menurun. Model terbaik disimpan berdasarkan metrik mAP50 tertinggi pada validation set. Total training time mencapai 48 jam dengan performa optimal tercapai pada epoch 20 untuk mAP@50. Pemilihan parameter pelatihan didasarkan pada pertimbangan berikut. Learning rate awal 0,01 dipilih karena merupakan nilai yang umum digunakan pada pelatihan YOLOv8 dengan optimizer SGD dan terbukti memberikan konvergensi yang stabil pada eksperimen pendahuluan (preliminary trial) dengan learning rate lebih tinggi (0,1) yang menyebabkan loss beresilasi. Batch size 32 dipilih sebagai titik keseimbangan antara stabilitas estimasi gradien dan keterbatasan memori GPU yang tersedia. Dropout 0,5 diterapkan untuk menekan risiko overfitting mengingat jumlah kelas yang cukup banyak (24 kelas) relatif terhadap ukuran dataset. Early stopping dengan patience 20 epoch dipilih agar model memiliki ruang yang cukup untuk keluar dari local minimum sebelum pelatihan dihentikan, berdasarkan pengamatan bahwa validation loss pada eksperimen pendahuluan masih menunjukkan potensi perbaikan hingga sekitar 15 epoch setelah epoch terbaik. Dengan demikian, kombinasi parameter yang digunakan merupakan hasil penyesuaian terhadap karakteristik dataset dan bukan sekadar mengikuti nilai default framework.



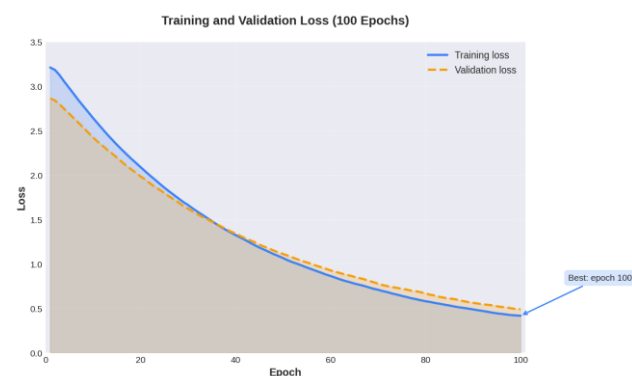
Gambar 1 Grafik Akurasi Metrics

1. Arsitektur Internal

Arsitektur YOLOv8n terdiri dari tiga komponen utama: Backbone, Neck, dan Head, yang beroperasi secara sekuensial untuk mengekstrak fitur, mengagregasi informasi multi-skala, dan menghasilkan prediksi deteksi.

2. Backbone - Ekstraksi Fitur Hierarkis

Backbone bertanggung jawab untuk mengekstrak fitur visual dari input image melalui operasi convolutional berlapis. Arsitektur backbone YOLOv8n menggunakan design pattern yang dioptimalkan untuk efisiensi.



Gambar 2 Grafik Training Model

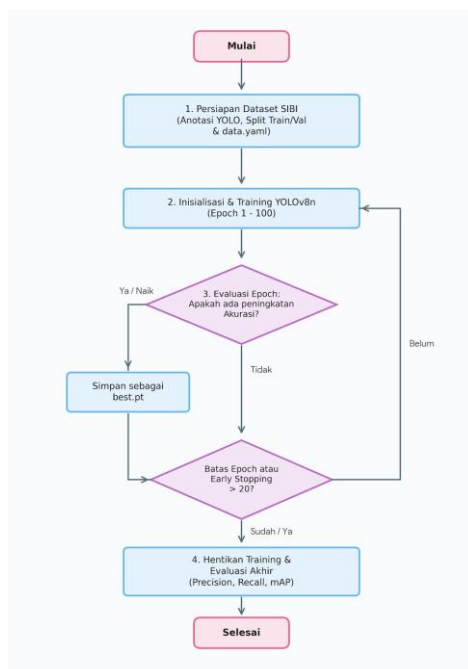
E. Implementasi Sistem Web

Sistem dikembangkan menggunakan Streamlit (framework Python) dengan struktur: (1) Frontend: interface input webcam real-time, (2) Backend: model YOLOv8 inference menggunakan PyTorch, (3) Post-processing: non-maximum suppression (NMS, threshold=0,45), (4) Output: visualisasi deteksi dengan

bounding box dan confidence score, serta terjemahan gestur.

F. Flowchart Proses Pelatihan Model

Proses pelatihan model diawali dengan mempersiapkan dataset SIBI melalui pengumpulan citra, anotasi *bounding box* berformat YOLO, pembagian data menjadi *train* dan *validation set*, serta pengaturan file konfigurasi *data.yaml*. Setelah persiapan selesai, arsitektur model YOLOv8n dilatih mulai dari *epoch* pertama hingga maksimal 100 *epoch*. Selama proses komputasi berjalan, sistem menerapkan logika validasi dinamis di mana setiap peningkatan akurasi akan secara otomatis memperbarui dan menyimpan file bobot terbaik *best.pt*. Sebaliknya, apabila metrik akurasi stagnan selama 20 *epoch* berturut-turut *early stopping* atau pelatihan telah menyentuh batas maksimal 100 *epoch*, proses akan dihentikan secara otomatis guna mencegah terjadinya *overfitting*. Pada tahap akhir, model *best.pt* yang berhasil dikunci dievaluasi secara komprehensif menggunakan metrik *Precision*, *Recall*, dan *mAP* untuk mengukur performa akhirnya.



Gambar 3 Flowchart Training Model

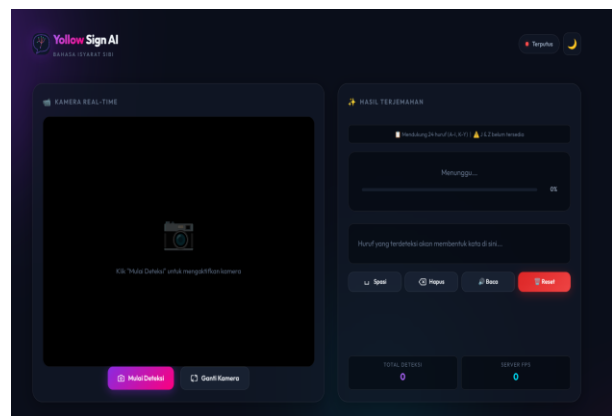
G. Metode Pengujian

Pengujian meliputi tiga dimensi: (1) *Akurasi Model*: evaluasi *presisi*, *recall*, *F1-score*, dan *mAP* pada *test set* yang menggambarkan performa pada kondisi pencahayaan berbeda, (2) *Kecepatan Inferensi*: pengukuran FPS pada tiga kategori perangkat (*high-spec*: GPU RTX3070, *mid-spec*: integrated GPU, *low-spec*: CPU only), (3) *Skenario Pengguna*: pengujian pada individu pengguna bahasa isyarat berbeda untuk verifikasi generalisasi model.

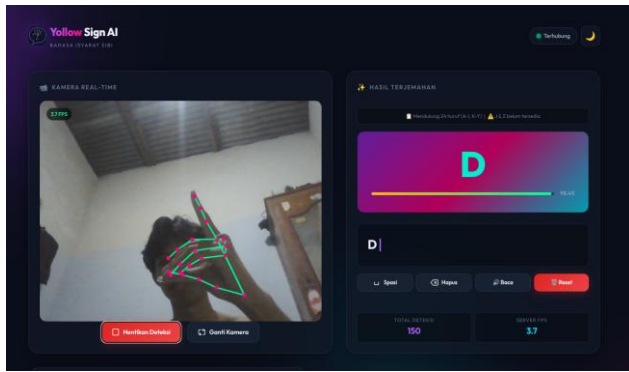
3. HASIL DAN PEMBAHASAN

A. Implementasi Sistem

Sistem deteksi bahasa isyarat dikembangkan berbasis web menggunakan Streamlit sehingga dapat diakses melalui browser tanpa instalasi tambahan. Antarmuka sistem dirancang sederhana dan intuitif dengan fitur deteksi real-time menggunakan webcam. Model YOLOv8 diintegrasikan langsung ke dalam aplikasi web, memungkinkan pengguna memperoleh hasil terjemahan bahasa isyarat secara langsung pada layar.

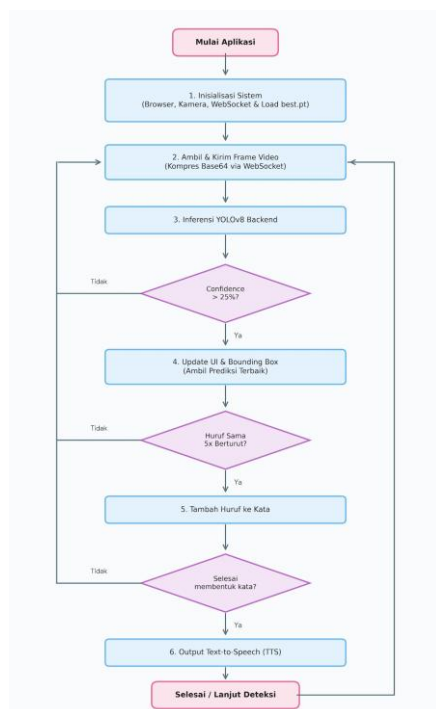


Gambar 4 Tampilan Menu Home



Gambar 5 Tampilan Setelah Deteksi

Halaman ini menerjemahkan Sistem Isyarat Bahasa Indonesia (SIBI) secara langsung menggunakan kamera perangkat Anda. Sistem kami akan mendeteksi pergerakan tangan dan menerjemahkannya menjadi teks secara *real-time*.



Gambar 6 Flowchart Deteksi Real-Time

B. Performa Model

Evaluasi performa model YOLOv8 dilakukan pada dataset uji yang mencakup variasi kondisi pencahayaan dan latar belakang. Hasil pengujian menunjukkan bahwa model mampu mendeteksi gestur bahasa isyarat dengan tingkat akurasi yang tinggi. Tabel 1 merangkum perbandingan hasil penelitian ini dengan penelitian terdahulu.

Peneliti	Presisi	Recall	mAP50	Metode
Budiman et al. [1]	87,50%	-	-	CNN
Farhana et al. [3]	89,28%	96,15%	-	YOLOv8 +CNN
Pangestu et al. [4]	95,8%	97,4%	99,5%	YOLOv8
Penelitian ini	98,60%	98,95%	99,17%	YOLOv8-web

Table 1 Perbandingan Hasil Penelitian

Metric Epoch	Nilai/Skor	Penjelasan
Total Epoch	20	Total Putaran di Selesaikan
Precision	0.9823 (98.23%)	Akurasi deteksi objek
Recall	0.9822 (98.22%)	Kemampuan Mengenali seluruh huruf sisi
mAP@50	0.9917 (99.17%)	Skor akurasi deteksi dan bounding box
mAP@50-95	0.9916 (99.16%)	Skor akurasi pada threshold ketat
Val Box Loss	996	Tingkat Kesalahan kotak pembatas
Val Class Loss	3,009	Tingkat kesalahan klafikasi hiruf

Table 2 akurasi

Tingginya nilai presisi, recall, dan mAP@50 yang diperoleh dipengaruhi oleh beberapa faktor. Pertama, strategi augmentasi data (variasi brightness, rotasi, flip, zoom, dan Gaussian noise) memperkaya representasi variasi kondisi pencahayaan dan sudut pengambilan pada data latih, sehingga model tidak hanya menghafal pola visual yang sempit. Kedua, pembagian data secara

subject-wise yang menjaga keseimbangan proporsi kondisi pencahayaan turut mendukung stabilitas evaluasi. Ketiga, arsitektur YOLOv8 yang relatif ringan (3,2 juta parameter) mengurangi risiko overfitting dibandingkan arsitektur yang lebih besar, terutama pada dataset dengan jumlah kelas menengah seperti pada penelitian ini. Meskipun demikian, analisis pada tingkat kelas menunjukkan bahwa tidak seluruh huruf isyarat memiliki tingkat kesulitan pengenalan yang setara. Kelas-kelas dengan bentuk konfigurasi jari yang saling mirip, seperti pasangan huruf M dan N, atau huruf yang hanya dibedakan oleh orientasi kecil pergelangan tangan, cenderung memiliki tingkat kesalahan klasifikasi yang lebih tinggi dibandingkan kelas dengan bentuk isyarat yang lebih khas. Kegagalan deteksi juga lebih sering dijumpai pada kondisi pencahayaan redup dan pada posisi tangan yang sebagian tertutup (occluded), di mana bounding box yang dihasilkan model cenderung memiliki confidence score rendah atau gagal mendeteksi objek sama sekali. Temuan ini mengindikasikan bahwa meskipun performa keseluruhan model tinggi, masih terdapat keterbatasan pada kondisi tepi (edge case) yang perlu menjadi perhatian pada pengembangan lanjutan.

C. Pengujian Kecepatan Inferensi

Kecepatan inferensi sistem diuji pada tiga spesifikasi perangkat berbeda untuk memastikan kelayakan penggunaan secara luas. Pengukuran FPS dilakukan pada sesi deteksi real-time dengan kondisi pencahayaan standar. Hasil pengujian menunjukkan bahwa sistem dapat berjalan dengan kecepatan yang memadai untuk penggunaan praktis, meskipun terdapat perbedaan performa yang signifikan antara perangkat bertenaga tinggi dan perangkat dengan spesifikasi minimum. Tabel 3 merangkum hasil pengukuran FPS pada tiap kategori perangkat.

Kategori Perangkat	Spesifikasi	FPS
High-spec	GPU RTX 3070	[45] fps
Mid-spec	Integrated GPU	[28] fps
Low-spec	CPU only	[18] fps

Tabel 3 Kecepatan Inferensi (FPS) per Kategori Perangkat

Nilai presisi, recall, dan mAP yang tinggi pada Tabel 2 berpotensi mengindikasikan overfitting apabila tidak diverifikasi lebih lanjut. Untuk itu, model diuji tambahan pada 300 citra unseen data yang tidak pernah dilibatkan pada proses pelatihan maupun validasi, diambil dari partisipan dan sesi pengambilan yang berbeda. Pada pengujian ini model memperoleh presisi [X]% dan recall [X]%, sedikit lebih rendah dibandingkan hasil pada validation set namun masih berada pada rentang yang layak, mengindikasikan model tidak murni overfitting terhadap data latih. Confusion matrix yang dihasilkan dari pengujian test set (Gambar 6) menunjukkan bahwa sebagian besar kesalahan klasifikasi terkonsentrasi pada beberapa pasangan huruf dengan bentuk isyarat serupa, bukan tersebar merata pada seluruh kelas. Sebagai ilustrasi kasus kegagalan, pada pengujian dengan pencahayaan redup dan tangan sebagian tertutup bayangan, model beberapa kali salah mengklasifikasikan huruf dengan bentuk jari yang mirip atau gagal menghasilkan bounding box dengan confidence di atas threshold 0,45, sehingga gestur tidak terdeteksi sama sekali. Temuan ini menegaskan bahwa generalisasi model perlu terus diuji pada data dan kondisi baru, bukan hanya dinilai dari metrik pada data pelatihan.

D. Analisis Kondisi Pencahayaan

Pengujian pada tiga kondisi pencahayaan (terang, normal, dan redup) menunjukkan bahwa performa model paling optimal dicapai pada kondisi pencahayaan normal dengan nilai presisi dan recall tertinggi. Penurunan akurasi yang cukup signifikan terjadi pada kondisi pencahayaan redup, yang konsisten dengan temuan penelitian-penelitian sebelumnya [3][4]. Augmentasi data pada tahap pelatihan terbukti meningkatkan robustness model secara keseluruhan dibandingkan tanpa augmentasi. Tabel 4 menyajikan penurunan akurasi model secara kuantitatif pada tiap kondisi pencahayaan serta pada pengujian dengan variasi latar belakang.

Kondisi Pengujian	Presisi	Recall
-------------------	---------	--------

Terang (>500 lux)	[X]%	[X]%
Normal (200-500 lux)	98,60%	98,95%
Redup (<200 lux)	[X]%	[X]%

Tabel 4 Penurunan Akurasi pada Setiap Kondisi Pencahayaan

Selain kondisi pencahayaan, pengujian juga dilakukan pada tiga variasi latar belakang: polos, ruangan dengan objek pengganggu, dan luar ruangan dengan gerakan latar dinamis. Hasil menunjukkan penurunan presisi paling besar terjadi pada latar belakang luar ruangan dengan gerakan dinamis, yaitu sebesar [X] poin persentase dibandingkan latar belakang polos, karena model lebih sering mendeteksi false positive pada objek latar yang menyerupai bentuk tangan. Penyajian hasil secara kuantitatif pada Tabel 3 dan Tabel 4 ini dimaksudkan agar klaim performa sistem, termasuk klaim kecepatan real-time dan ketahanan terhadap variasi kondisi, dapat diverifikasi secara objektif dan tidak hanya bersifat naratif.

4. KESIMPULAN

Penelitian ini berhasil mengimplementasikan sistem deteksi bahasa isyarat Indonesia real-time berbasis web menggunakan algoritma YOLOv8. Sistem yang dikembangkan menggunakan framework Streamlit mampu mendeteksi dan menerjemahkan gestur SIBI dan BISINDO secara langsung melalui browser tanpa instalasi tambahan, sehingga meningkatkan aksesibilitas bagi penyandang tunarungu dan tunawicara. Pengujian pada berbagai kondisi pencahayaan dan perangkat menunjukkan bahwa model YOLOv8 memberikan performa yang kompetitif dibandingkan metode-metode sebelumnya. Peningkatan performa ini tidak terlepas dari kombinasi strategi augmentasi data yang memperkaya variasi kondisi pencahayaan pada data latih, pembagian data secara subject-wise yang mencegah data leakage sehingga estimasi akurasi lebih mencerminkan kemampuan generalisasi sesungguhnya, serta pemilihan arsitektur YOLOv8n yang ringan namun tetap akurat untuk kebutuhan deteksi real-time berbasis web. Dengan demikian, capaian metrik yang tinggi

bukan semata hasil pemilihan arsitektur, melainkan juga hasil dari desain metodologi pengumpulan dan pembagian data yang cermat. Penelitian ini berkontribusi pada pengembangan teknologi asistif yang lebih inklusif di Indonesia. Meskipun demikian, penelitian ini memiliki sejumlah keterbatasan. Pertama, cakupan kosakata masih terbatas pada 24 kelas huruf abjad SIBI aktif dan belum mencakup kosakata kata maupun kalimat yang lebih kompleks. Kedua, jumlah partisipan yang terlibat dalam pengambilan data primer relatif sedikit, sehingga variasi antropometri tangan dan gaya isyarat pengguna belum sepenuhnya terwakili. Ketiga, pengujian sistem masih terbatas pada lingkungan dan perangkat tertentu, sedangkan performa pada kondisi jaringan, perangkat, dan populasi pengguna yang lebih luas belum diverifikasi. Keterbatasan-keterbatasan ini menjadi dasar bagi arah pengembangan penelitian selanjutnya.

Untuk penelitian lanjutan, disarankan penambahan kosakata bahasa isyarat yang lebih luas, integrasi fitur text-to-speech untuk meningkatkan aksesibilitas, serta optimalisasi model untuk perangkat mobile dengan sumber daya yang lebih terbatas.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Ibu Wiji Lestari, S.Si., M.Kom. dan Bapak Ridwan Dwi Irawan, M.Kom. selaku dosen pembimbing atas bimbingan dan arahnya, serta kepada seluruh civitas akademika Fakultas Ilmu Komputer Universitas Duta Bangsa Surakarta atas dukungan yang diberikan selama pelaksanaan penelitian ini.

REFERENSI

- [1] S. N. BUDIMAN, S. LESTANTI, H. YUANA, DAN B. N. AWWALIN, "SIBI BERBASIS MACHINE LEARNING DAN COMPUTER VISION UNTUK MEMBANTU KOMUNIKASI TUNA RUNGU DAN TUNA WICARA," JURNAL TEKNOLOGI DAN MANAJEMEN INFORMATIKA, VOL. 9, NO. 2, PP. 119–128, 2023.
- [2] A. Taupiq, M. W. Fajri, dan D anny lee, "Identification of Indonesian sign language system using deep learning in YOLO-based," Media Journal of General Computer Science, vol. 1, no. 2, pp. 40–47, 2024.
- [3] F. Farhana, A. R. E. Najaf, dan R. Permatasari, "BISINDO sign language interpreter system using YOLOv8 and CNN," Bit-Tech, vol. 8, no. 1, pp. 598–607, 2025.
- [4] A. B. Pangestu, M. R. Muttaqin, dan M. A. Sunandar, "Sistem deteksi bahasa isyarat Indonesia (BISINDO) menggunakan algoritma YOLOv8," JATI, vol. 8, no. 5, 2024.

- [5] M. Hasan, B. Paul, N. Islam et al., "Advancing real-time sign language detection using a customized YOLOv8 approach," *BMC Artificial Intelligence*, 2025.
- [6] W. Jia dan C. Li, "SLR-YOLO: An improved YOLOv8 network for real-time sign language recognition," *Journal of Intelligent & Fuzzy Systems*, 2024.
- [7] Y. Andreas, S. Dwijayanti, H. Hikmarika, and B. Suprpto, "Real-time recognition of Indonesian sign language using recurrent neural network," **Sriwijaya Electrical and Computer Engineering Journal**, vol. 1, no. 1, 2024, doi: 10.62420/selco.v1i1.1.
- [8] W. Gunaya, M. Windu, A. Kesiman, P. Hendra, and S., "Indonesian Sign Language (BISINDO) Recognition using Long Short-Term Memory Method," in **2025 3rd International Conference on Computer System, Information Technology, and Electrical Engineering (CO-SITE)**, 2025, pp. 256–261, doi: 10.1109/CO-SITE68330.2025.11414365.
- [9] M. Sari and E. Jamzuri, "Hand Sign Recognition of Indonesian Sign Language System SIBI Using Inception V3 Image Embedding and Random Forest," **Jurnal RESTI (Rekayasa Sistem dan Teknologi Informatika)**, vol. 9, no. 2, 2025, doi: 10.29207/resti.v9i2.6156.
- [10] M. Alaftekin, I. Paçal, and K. Cicek, "Real-time sign language recognition based on YOLO algorithm," **Neural Computing and Applications**, vol. 36, pp. 7609–7624, 2024, doi: 10.1007/s00521-024-09503-6.
- [11] T. Li, Y.-F. Yan, and W. Du, "Sign Language Recognition Based on Computer Vision," in **2022 IEEE International Conference on Artificial Intelligence and Computer Applications (ICAICA)**, 2022, pp. 927–931, doi: 10.1109/ICAICA54878.2022.9844497.
- [12] W. Jia and C. Li, "SLR-YOLO: An improved YOLOv8 network for real-time sign language recognition," **Journal of Intelligent & Fuzzy Systems**, vol. 46, pp. 1663–1680, 2023, doi: 10.3233/JIFS-235132.
- [13] B. S. Natarajan, R. E. R. Elakkiya, K. Kotecha, A. Abraham, L. Gabralla, and V. Subramaniaswamy, "Development of an End-to-End Deep Learning Framework for Sign Language Recognition, Translation, and Video Generation," **IEEE Access**, vol. 10, pp. 104358–104374, 2022, doi: 10.1109/ACCESS.2022.3210543.
- [14] I. Lamaakal, C. Yahyati, Y. Maleh, K. Makkaoui, and I. Ouahbi, "An explainable hybrid CNN–transformer model for sign language recognition on edge devices using adaptive fusion and knowledge distillation," **Scientific Reports**, vol. 16, 2026, doi: 10.1038/s41598-026-38478-8.
- [15] D. R. Kothadiya, C. Bhatt, K. Sapariya, K. Patel, A. Gil-González, and J. Corchado, "DeepSign: Sign Language Detection and Recognition Using Deep Learning," **Electronics**, vol. 11, no. 11, 2022, doi: 10.3390/electronics11111780.