

Implementasi Deteksi Objek Dengan Model YOLOV8 pada Pengenalan Bahasa Isyarat

Khurotul Aeni^{*1}, Asep Saeful Millah²

^{1,2})Program Studi Informatika, Fakultas Sains dan Teknologi, Universitas Peradaban

Email: ^{*}khaeni988@gmail.com, ²asepsaeful232@gmail.com

(Naskah masuk: 3 Agustus 2024, diterima untuk diterbitkan: 20 Januari 2025)

Abstrak: Bahasa isyarat adalah alat komunikasi yang diterapkan oleh penyandang tunarungu dan tunawicara untuk berkomunikasi dengan masyarakat umum. Gerakan tangan, mimik, dan gerakan tubuh berupa simbol yang mewakili huruf maupun kata digunakan penyandang tunarungu dan tunawicara untuk berinteraksi dengan orang lain. Namun, kesenjangan komunikasi antara mereka dan masyarakat umum masih ada sampai saat ini. Pemahaman yang mendalam dibutuhkan untuk komunikasi ini, khususnya untuk mereka yang tidak terbiasa menggunakan bahasa isyarat. Hal ini membuat masyarakat umum kesulitan berkomunikasi dengan orang yang tunarungu dan tunawicara. Sebab itu diperlukan sebuah sistem yang membantu masyarakat umum berinteraksi dengan para penyandang tunarungu dan tunawicara dengan menerapkan teknologi informasi yang saat ini berkembang seperti penggunaan deep learning. Maksud dari penelitian ini yaitu mengimplementasikan deteksi objek dengan metode YOLOV8 untuk mengenali bahasa isyarat dari data gambar dan dibuatnya sistem pengenalan bahasa isyarat ini diharapkan dapat mempermudah masyarakat dalam mengenali bahasa isyarat SIBI berupa abjad menggunakan gestur tangan. YOLOV8 banyak digunakan dalam deteksi objek karena tingkat akurasi yang dihasilkan sangat tinggi dan kecepatan yang seimbang, sehingga memungkinkan untuk mendeteksi objek secara cepat dan akurat. Model ini berhasil mendapatkan nilai akurasi terbesarnya yaitu 99,5%. Selain itu, kecepatan deteksi gestur SIBI pada model ini memiliki rata-rata 4046 ms.

Kata Kunci – Deteksi Objek; YOLOv8; Bahasa Isyarat; SIBI

Implementation of Object Detection with YOLOV8 Model in Sign Language Recognition

Abstract: Implementation of Object Detection with YOLOV8 Model on Sign Language Recognition
Sign language is a communication tool used by deaf and speech impaired to communicate in society. Hand gestures, facial expressions, and body movements in the form of symbols representing letters or words are used by deaf and speech impaired to interact with others. However, the communication gap between them and public still exists now. Deep understanding is needed for this communication, especially for those who are not using the sign language. This cause the society difficult to communicate with them. Therefore, a system is needed that helps society interact with the deaf and speech impaired by applying the latest information technology such as the use deep learning. The purpose of this study is to implement object detection with YOLOV8 method to recognize sign language from image data and by creating sign language recognition system is expected to facilitate public to recognize SIBI sign language in the form of alphabets using hand gestures. YOLOV8 is widely used in object detection because of the level of accuracy produced is very high with balance speed, it is possible to detect objects quickly and accurately. This model managed to get highest accuracy value of 99.5%. In addition, the speed of SIBI gesture detection has an average of 4046 ms..

Keywords – Object Detection, YOLOv8, Sign Language, SIBI

1. PENDAHULUAN

Beberapa tahun belakangan ini, teknologi deteksi objek dan pengenalan isyarat seperti pengenalan wajah dan pengenalan bahasa isyarat menjadi sangat penting dalam cabang bahasa isyarat. Gestur adalah berbagai gerakan yang digunakan pada proses komunikasi [1]. Bahasa isyarat sama pentingnya dengan bahasa lisan, dimana bahasa isyarat ini digunakan oleh para penyandang disabilitas tunarungu dan tunawicara. Bahasa isyarat dilakukan dengan gerakan tangan disertai

ekspresi wajah sesuai dengan kaidah tata bahasanya. Setiap bahasa isyarat memiliki ribuan tanda, masing-masing berbeda satu sama lain melalui perubahan kecil pada bentuk tangan, gerakan, posisi, dan fitur non-manual (gerakan komponen wajah seperti ekspresi alis) [2].

Aksesibilitas bagi penyandang disabilitas, terutama tunarungu dan tunawicara, sangat dipengaruhi oleh bahasa isyarat. Menurut UU No. 19 Thn. 2011 Pasal 9 (2), aksesibilitas bagi penyandang disabilitas dapat meliputi perantara dan pemberian bantuan langsung, seperti pembaca, pemandu, dan penerjemah profesional. Langkah-langkah ini dirancang untuk memudahkan akses mereka ke bangunan dan fasilitas umum lainnya [3].

SIBI adalah salah satu dari 2 jenis bahasa isyarat yang digunakan di Indonesia. SIBI adalah bahasa isyarat yang berlaku secara nasional. Bahasa isyarat SIBI ini diseragamkan dan merupakan bahasa isyarat baku. SIBI biasanya banyak digunakan pada acara formal seperti pada kegiatan di sekolah dan kegiatan resmi [4].

Kurangnya pemahaman masyarakat umum tentang bahasa isyarat disebabkan oleh keterbatasan pengetahuan mereka tentang bahasa tersebut. Hal ini terjadi karena bahasa isyarat terbentuk dan berkembang secara alami dalam komunitas penyandang tunarungu dan tunawicara [3]. Kondisi tersebut menyebabkan masyarakat kesulitan dalam melakukan komunikasi dengan para penyandang disabilitas tunarungu dan tunawicara. Oleh karena itu, masyarakat umum membutuhkan teknologi yang bisa membantu mereka untuk dapat berinteraksi dengan para penyandang tunawicara dan tunarungu.

Model deep learning saat ini banyak diterapkan di berbagai bidang computer vision, termasuk object detection, yang digunakan untuk mendeteksi objek yang terdapat pada gambar dan video, seperti wajah, pohon, dan berbagai objek lainnya. Deteksi objek adalah salah satu aplikasi dalam computer vision yang meniru cara penglihatan manusia, atau yang dikenal sebagai human vision. [5]. Teknologi object detection ini dapat digunakan sebagai alat komunikasi bagi individu tunarungu dan tunawicara, membantu menerjemahkan bahasa isyarat yang mereka gunakan sehingga mudah dipahami oleh masyarakat pada umumnya yang belum mengenal Bahasa isyarat..

Deteksi objek secara real time telah muncul sebagai komponen penting dalam berbagai aplikasi. Di antara berbagai algoritma pendeteksian objek, kerangka kerja YOLO (You Only Look Once) menonjol karena keseimbangan kecepatan dan akurasi yang luar biasa, memungkinkan identifikasi objek dalam gambar dengan cepat dan andal. YOLO telah berevolusi melalui beberapa iterasi, masing-masing dikembangkan berdasarkan versi sebelumnya untuk mengatasi keterbatasan dan meningkatkan kinerja. Versi terbaru dari YOLO adalah YOLOv8 yang dirilis pada 10 Januari 2023 oleh Ultralytics, yaitu perusahaan yang mengembangkan YOLOv5 [6].

YOLOv8 mempunyai tingkat akurasi yang tinggi pada pengujian resminya dengan menggunakan dataset dari Microsoft yaitu COCO dan Roboflow 100. YOLOv8 tersedia banyak fitur kenyamanan pengembang, mulai dari CLI yang mudah digunakan hingga paket Python yang terstruktur dengan baik. Berbeda dengan model lain yang tugasnya dibagi menjadi banyak file Python berbeda, YOLOv8 hadir dengan CLI yang membuat pelatihan model menjadi intuitif. Hal ini merupakan tambahan dari paket Python yang memberikan pengalaman pengkodean yang lebih lancar dibandingkan model sebelumnya [7].

Aplikasi pemanfaatan teknologi deteksi objek menggunakan YOLOv8 ini bertujuan untuk membantu masyarakat dalam mempelajari dan memahami bahasa isyarat SIBI. Sehingga aplikasi ini diharapkan dapat mengatasi keterbatasan komunikasi antara masyarakat umum dengan para penyandang disabilitas tunarungu dan tunawicara.

1.1. Pengelanaan Bahasa Isyarat

Terdapat beberapa penelitian yang bertujuan untuk mendeteksi bahasa isyarat. Pada penelitian yang dilakukan pada tahun 2022 menerapkan metode Convolutional Neural Network [8]. Penelitian tersebut menggunakan 416 citra dan memperoleh nilai akurasi sebesar 80,76%.

Penelitian lain yang dilakukan pada tahun 2022 [9], menggunakan local directional pattern dan Klasifikasi K-Nearest Neighbour untuk pengenalan angka bahasa isyarat. Data yang digunakan

pada penelitian tersebut adalah data Turkey Ankara Ayranci Anadolu High Schools Sign Language Digits Dataset yang berjumlah 2062 data. Penelitian tersebut berhasil mendapatkan nilai akurasi sebesar 88.45%.

Berbeda pada penelitian-penelitian yang lain, penelitian pada tahun 2021 [10] dengan teknologi Augmented Reality untuk mengenalkan bahasa isyarat. Jenis BISINDO. Namun, pengujian yang dilakukan menggunakan kuesioner System Usability Scale (SUS) memperoleh nilai kurang dari 74 pada pengujian kepuasan pengguna, yang artinya pengguna kurang puas dengan aplikasi yang dibangun.

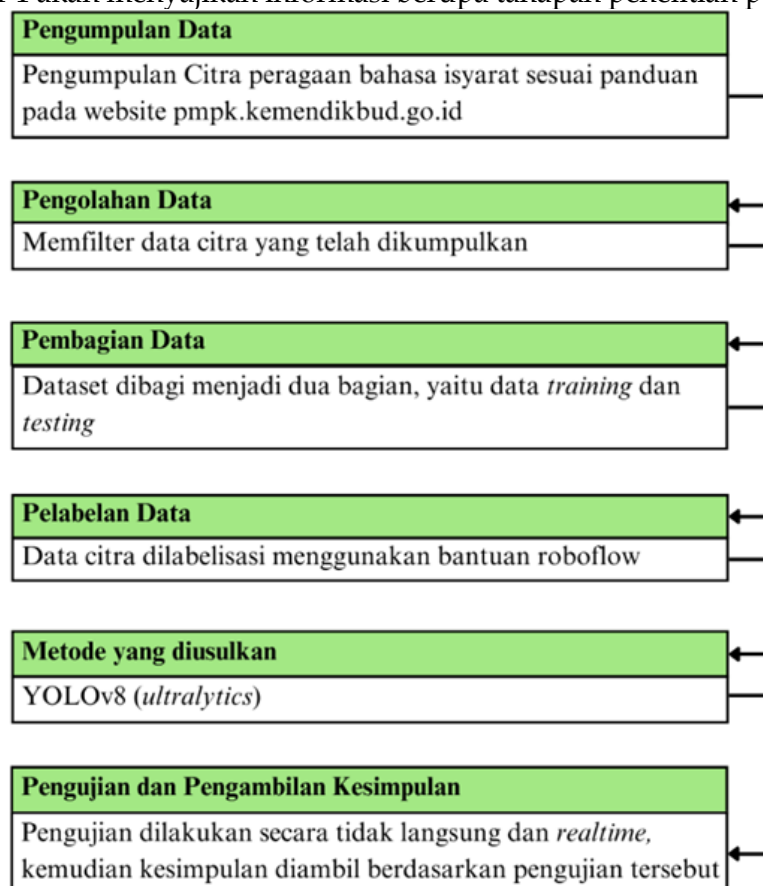
Pengenalan bahasa isyarat dapat dilakukan dengan berbasis game edukasi. Seperti pada penelitian tahun 2022 [11]. Penelitian tersebut menggunakan MYO ARMBAND pada Arsitektur CLIENT-SERVER. Pengujian secara real-time dengan waktu rerata 0,1107333 detik/ request pada latency jaringan rerata 36,7%.

Penggunaan metode lainnya seperti Deep Gated Recurrent Unit (GRU) juga digunakan untuk mendeteksi tangan pada video percakapan bahasa isyarat. Penelitian tersebut dilakukan pada tahun 2022 [12]. Pada pengujian yang menggunakan video percakapan bahasa isyarat berjenis BISINDO, didapatkan nilai akurasi sebesar 88%..

Pendahuluan menguraikan latar belakang permasalahan yang diselesaikan, isu-isu yang terkait dengan masalah yang diselesaikan, ulasan penelitian yang pernah dilakukan sebelumnya oleh peneliti lain yang relevan dengan penelitian yang dilakukan, kemudian dalam latar belakang juga harus berisi terkait state of the art dari sebuah penelitian.

2. METODE PENELITIAN

Proses membangun model deteksi objek bahasa isyarat SIBI menggunakan YOLOv8 ini melewati beberapa tahapan. Tahapan tersebut mencakup pengumpulan data, pengolahan data, pembagian data, pelabelan data, metode yang diusulkan, serta pengujian dan pengambilan keputusan. Gambar 1 akan menyajikan informasi berupa tahapan penelitian penggunaan YOLOv8.



Gambar 1. Tahapan Penelitian

2.1. Pengumpulan Data

Proses ini bertujuan untuk mengumpulkan dataset berisi foto gestur tangan yang memperagakan 10 huruf abjad pada Sistem Isyarat Bahasa Indonesia. Gestur tangan dilakukan sesuai dengan panduan yang tertera di website pmpk.kemendikbud.go.id, dapat dilihat pada Gambar 2.

2.2. Pengolahan Data

Dalam kegiatan pengolahan data, langkah pertama dalam proses ini adalah memilah dan memeriksa setiap foto untuk menentukan apakah foto tersebut sesuai dengan kriteria yang telah ditentukan. Foto-foto yang tidak memenuhi kriteria ini akan dihapus dari dataset. Setelah proses penyaringan ini selesai, jumlah foto yang tersisa dan akan digunakan dalam pemrosesan lebih lanjut adalah sebanyak 3000 foto.

2.3. Pembagian Data

Setelah pengolahan, data citra dibagi menjadi tiga kelompok: data pelatihan, validasi, dan pengujian. Pembagian dilakukan dengan rasio 80:5:15, di mana 80% dari data digunakan untuk pelatihan, 5% untuk validasi, dan 15% untuk pengujian. Dengan rasio ini, dataset yang digunakan perkelas berjumlah 300 citra, dengan masing kelas memiliki 240 data latih, 15 data validasi dan 45 data uji.

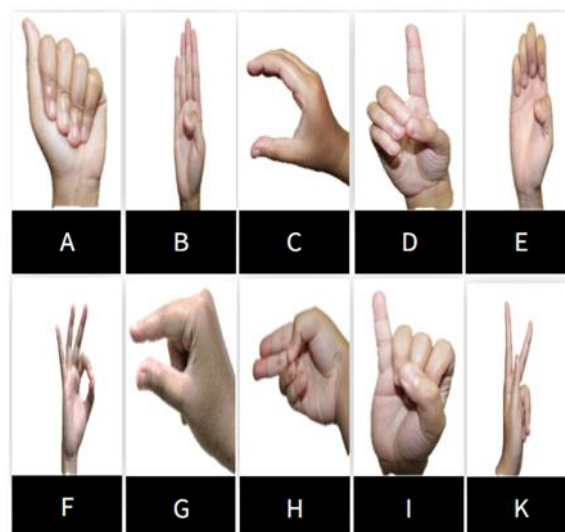
2.4. Pelabelan Data

Proses pelabelan dataset dengan Roboflow melibatkan penandaan kotak pembatas di sekitar objek dalam citra dan memberikan label untuk setiap objek yang ditandai. Ini menciptakan dataset yang diperkaya dengan informasi tentang lokasi dan jenis objek dalam setiap gambar. Pelabelan gambar dapat dilihat pada Gambar 4.

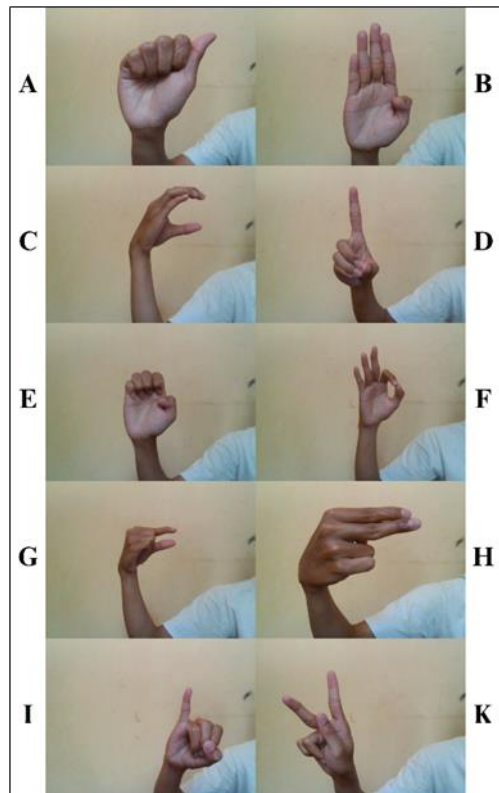
2.5. Metode Yang Diusulkan

Dalam penelitian ini, metode YOLOv8s diterapkan dari tahap persiapan data hingga pengujian. Model akan dibentuk dari data yang telah diolah. Dimulai dengan persiapan data, langkah-langkah metode ini kemudian meliputi pelatihan

3. HASIL DAN PEMBAHASAN



Gambar 2. Bentuk Isyarat huruf pada kamus SIBI



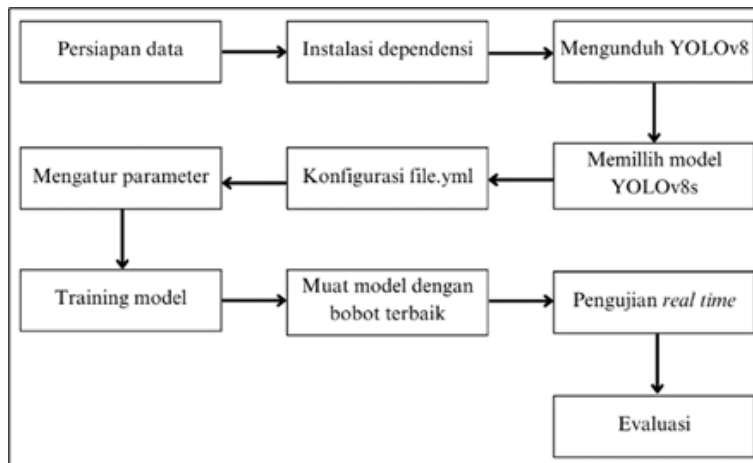
Gambar 3. Sampel Data Mentah

model, validasi, pengujian, dan evaluasi. YOLOv8s, yang merupakan versi terbaru dari pendekatan deteksi objek real-time YOLO, bekerja dengan membagi gambar menjadi grid dan memprediksi kotak pembatas serta kelas objek dalam setiap grid. Melalui proses pelatihan, model disesuaikan bobotnya untuk meningkatkan performa deteksi objek. Alurnya dapat dilihat pada Gambar 5.

1. Menyiapkan data, termasuk gambar dan label yang sudah diolah.
2. Menginstal dependensi dan library yang diperlukan.
3. Mengunduh YOLOv8.
4. Pilih model YOLOv8s .
5. Mengonfigurasi pada file yaml yang digunakan.
6. Mengatur parameter untuk proses pelatihan.
7. Melatih model dengan parameter yang telah ditetapkan. Berdasarkan arsitektur YOLOv8s yang dijelaskan dalam



Gambar 4. Pelabelan Gambar



Gambar 5. Metode yang Diusulkan
 Tabel 1. Parameter Training

Parameter	Jenis & Nilai
Image Size	640
Learning Rate	0,01
Batch Size	16
Epochs	100
Optimizer	AdamW

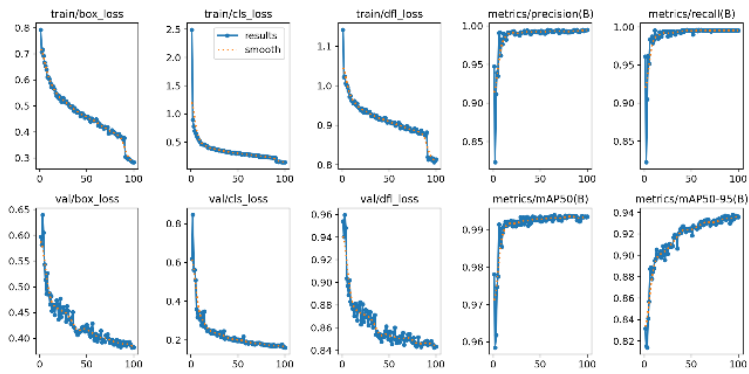
teori, proses pelatihan menggunakan tiga komponen utama: backbone, neck, dan head.

Setelah proses training dan validation berhasil dilakukan, model berhasil dibangun dengan hasil yang baik. Nilai cls_loss sangat kecil yaitu 0,14. Sedangkan nilai F1 score mendapatkan angka 0,868. Pada kurva presisi, model ini mendapatkan nilai 0,967. Model memperoleh nilai 0,995 untuk mAP@0,5. Sedangkan untuk mAP@0,50-95 mendapatkan nilai 0,993. Hasil dari pelatihan dapat dilihat pada Gambar 6. Proses validasi menghasilkan confusion matrix yang dapat dilihat pada Gambar 7. Nilai akurasi pada proses validasi menggunakan model tersebut adalah sebagai berikut:

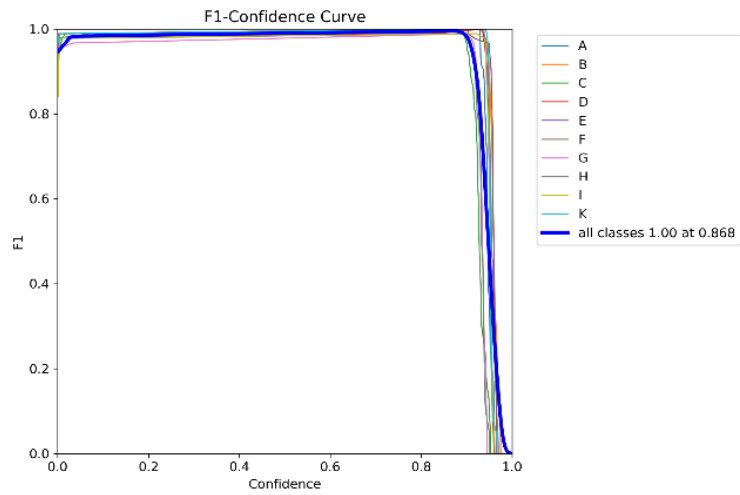
$$akurasi = \frac{TP+TN}{TP+TN+FP+FN} \times 100\% \quad (1)$$

$$akurasi = \frac{\begin{pmatrix} 1 + 1 + 1 + 1 + 1 + \\ 1 + 0,98 + 1 + 0,98 + 1 \\ + (0) \end{pmatrix}}{\begin{pmatrix} 1 + 1 + 1 + 1 + 1 + \\ 1 + 0,98 + 1 + 0,98 + 1 \\ + (0) + (0) + (0,02 + 0,02) \end{pmatrix}} \times 100\%$$

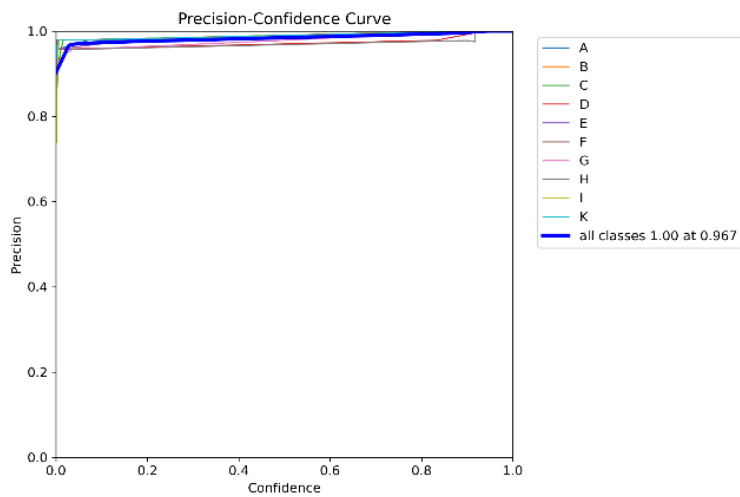
$$akurasi = \frac{9,96}{10} \times 100\% = 99,6\%$$



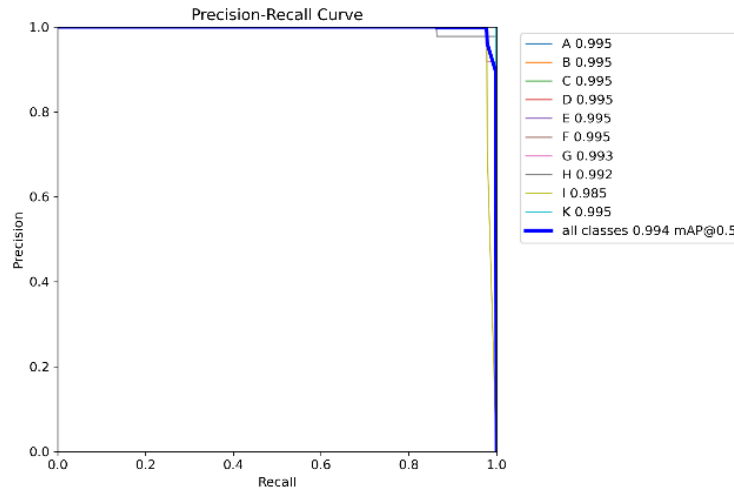
(a)



(b)



(c)



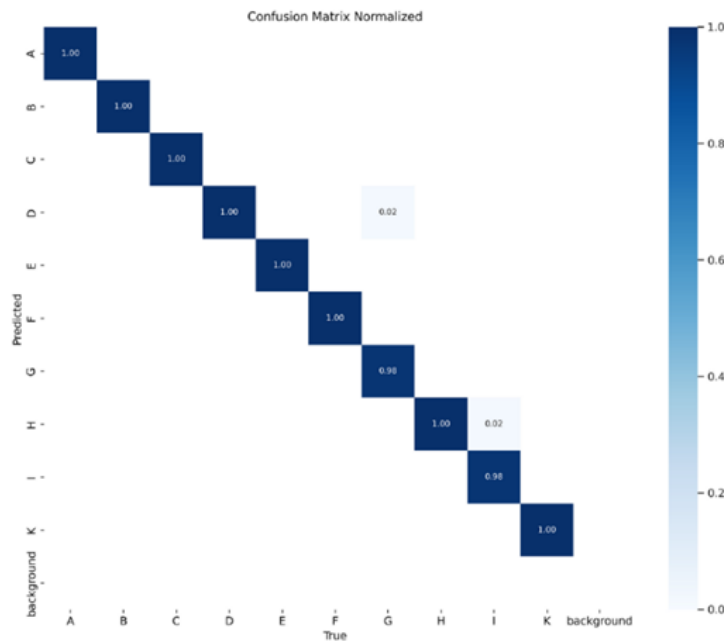
(d)

Gambar 6. Hasil Training, (a) training result, (b) nilai F1 score, (c) nilai presisi, (d) nilai mAP@50

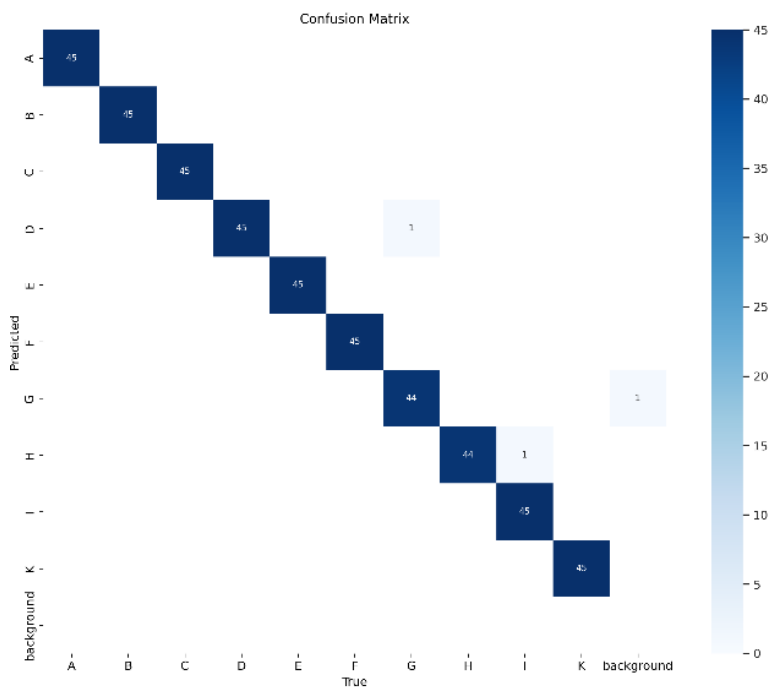
3.1. Testing

Model diuji menggunakan dataset pengujian. Hasil dari pengujian divisualisasikan menggunakan confusion matrix. Berdasarkan dari confusion matrix pada Gambar 8, maka nilai akurasi yang didapatkan pada proses testing adalah sebagai berikut:

$$akurasi = \frac{Jumlah\ data\ dikenali}{Total\ Data\ uji} \times 100\% \quad (2)$$



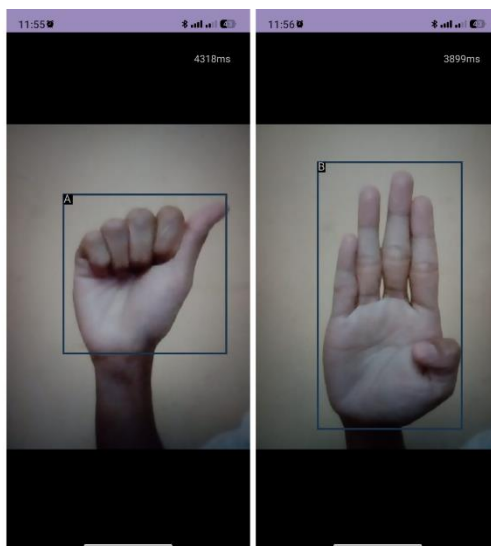
Gambar 7. Confussion Matrix Validation



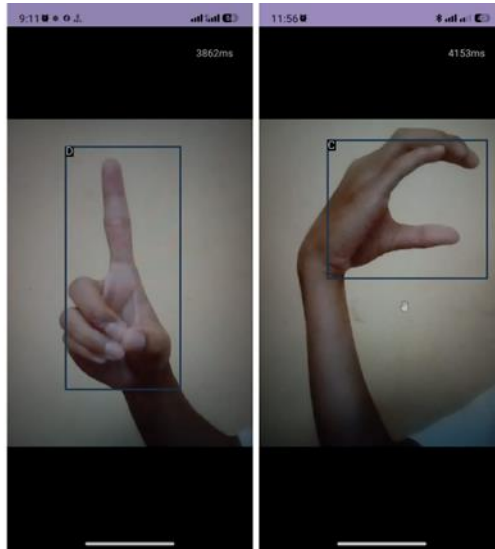
Gambar 8. Confussion Matrix Testing
 Maka nilai akurasi yang didapat adalah sebesar 99,5%.

3.2. Pengujian Berbasis Android

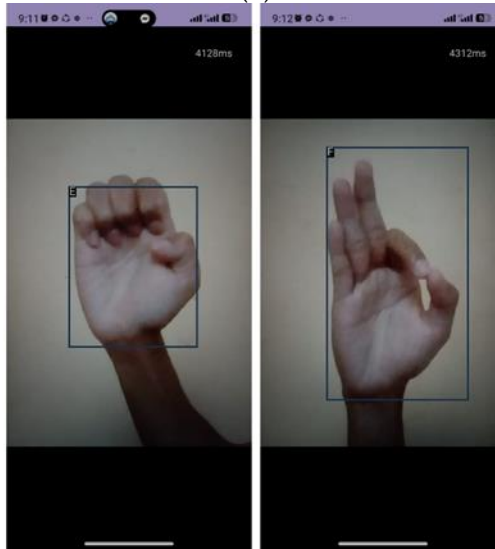
Model yang sudah jadi selanjutnya diterapkan dan akan dilakukan pengujian dengan aplikasi berbasis Android. Pengujian ini bertujuan untuk mengetahui kecepatan model dalam mendeteksi gestur SIBI. Berdasarkan hasil deteksi yang dilakukan berbasis android, model memiliki kecepatan yang bervariasi dalam mendeteksi objek gestur tangan. Kecepatan berkisar antara 3817ms sampai dengan 4318ms. Maka didapatkan rata-rata kecepatan model berbasis android dalam mendeteksi gestur SIBI adalah 4046ms. Hasil dari pengujian berbasis android ini dapat dilihat pada Gambar 9 dan Tabel II.



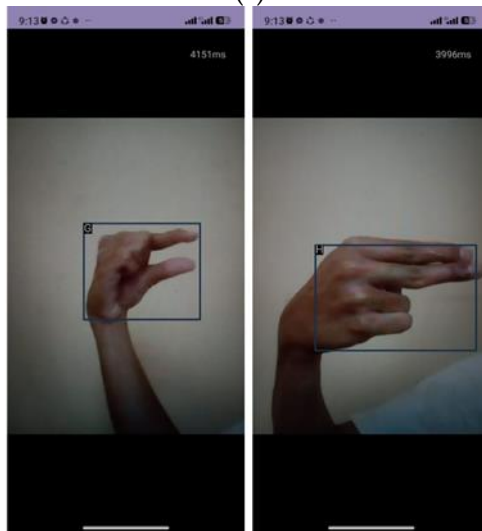
(a)



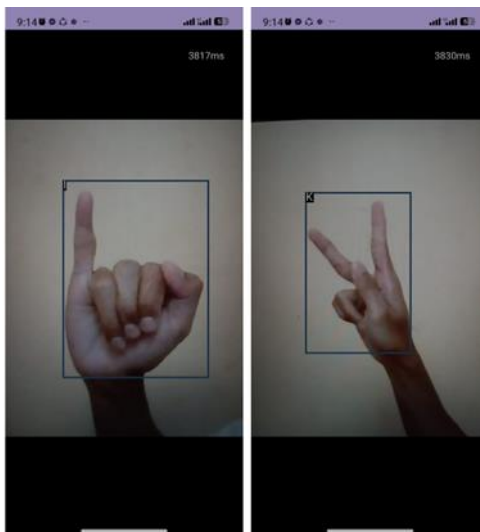
(b)



(c)



(d)



(e)

Gambar 9. (a) pengujian abjad A dan B, (b) pengujian abjad C dan D, (c) pengujian abjad E dan F, (d) pengujian abjad G dan H, (e) pengujian abjad I dan K

Tabel 2. Pengujian Berbasis Android

Pengujian	
Huruf	Kecepatan deteksi
A	4318 ms
B	3899 ms
C	4153 ms
D	3862 ms
E	4128 ms
F	4312 ms
G	4151 ms
H	3996 ms
I	3817 ms
K	3830 ms

4. KESIMPULAN

Penerapan metode YOLOv8s dalam penelitian ini menunjukkan kinerja yang sangat baik secara keseluruhan. Beberapa kegagalan yang terjadi selama tahap pengujian disebabkan oleh variasi latar belakang dalam proses deteksi objek, yang sering kali melibatkan banyak objek kompleks. Meskipun demikian, hasil penelitian menunjukkan potensi yang sangat menjanjikan. Dataset yang digunakan terdiri dari 3000 gambar yang mencakup 10 kelas gestur tangan, dengan masing-masing kelas berisi 300 gambar. Selama pengujian dengan dataset uji yang berjumlah 450 data, model mencapai akurasi yang tinggi sebesar 99,5%. Selain itu, pengujian real-time berbasis Android dilakukan dengan menguji setiap kelas. Kecepatan deteksi gestur SIBI menggunakan Android bervariasi, dengan rata-rata kecepatan sebesar 4046 ms.

UCAPAN TERIMA KASIH

Terimakasih Kepada LPPM Universitas Peradaban yang berkenan mendanai penelitian ini dalam program Hibah Internal Penelitian di Universitas Peradaban.

DAFTAR PUSTAKA

- [1] S. Shrenika and M. Madhu Bala, "Sign Language Recognition Using Template Matching Technique," 2020 Int. Conf. Comput. Sci. Eng. Appl. ICCSEA 2020, pp. 5-9, 2020, doi: 10.1109/ICCSEA49143.2020.9132899.
- [2] S. M. Kamal, Y. Chen, S. Li, X. Shi, and J. Zheng, "Technical Approaches to Chinese Sign Language Processing: A Review," IEEE Access, vol. 7, pp. 96926-96935, 2019, doi: 10.1109/ACCESS.2019.2929174.
- [3] B. R. Fajri and G. Kusumastuti, "Perceptions of 'Hearing' People on Sign Language Learning," Int. Conf. Educ. Technol., vol. 382, no. Icet, pp. 364-367, 2019, doi: 10.2991/icet-19.2019.91.
- [4] S. Apendi et al., "Sign Language Detection Indonesian Sign Language System Using the Single Shot Multibox Detector Method," vol. 10, no. 1, pp. 249-255, 2023.
- [5] J. Wang, T. Zhang, Y. Cheng, and N. Al-nabhan, "Deep Learning for Object Detection : A Survey," 2021, doi: 10.32604/csse.2021.017016.
- [6] J. Terven and D. Cordova-Esparza, "A Comprehensive Review of YOLO: From YOLOv1 to YOLOv8 and Beyond," pp. 1-27, 2023, [Online]. Available: <http://arxiv.org/abs/2304.00501>
- [7] D. Reis, J. Kupec, J. Hong, and A. Daoudi, "Real-Time Flying Object Detection with YOLOv8," 2023, [Online]. Available: <http://arxiv.org/abs/2305.09972>
- [8] M. Sholawati, K. Auliasari, and F. Ariwibisono, "Pengembangan Aplikasi Pengenalan Bahasa Isyarat Abjad Sibi Menggunakan Metode Convolutional Neural Network (Cnn)," JATI (Jurnal Mhs. Tek. Inform., vol. 6, no. 1, pp. 134-144, 2022, doi: 10.36040/jati.v6i1.4507.
- [9] N. Ilmi and H. Suryoprayogo, "Pengenalan Angka Bahasa Isyarat dengan Menggunakan Local Directional Pattern dan Klasifikasi K-Nearest Neighbour," J. Informatics Commun. Technol., vol. 4, no. 1, pp. 48-55, 2022, doi: 10.52661/j_ict.v4i1.103.
- [10] M. Dewi, T. Wahyuningrum, and N. A. Prasetyo, "Pengenalan Kata Bahasa Isyarat Indonesia (BISINDO) Menggunakan Augmented Reality (AR)," INISTA J. Informatics, Inf. Syst. Softw. Eng. Appl., vol. 3, no. 2, pp. 53-60, 2021, [Online]. Available: <https://journal.itelkom-pwt.ac.id/index.php/inista/article/view/256>
- [11] R. Y. Hakkun, M. Rizky, and B. Rafsanjani, "Game Edukasi Pengenalan Sistem Isyarat Bahasa Indonesia (SIBI) Menggunakan Myo Armband pada Arsitektur Client Server," J. Teknol. Inf. dan Ilmu Komput., vol. 9, no. 3, pp. 507-512, 2022, doi: 10.25126/jtiik.2021864149.
- [12] P. Kurnia Sari, G. Qorik Oktagalu Pratamasunu, and F. Nur Fajri, "Deteksi Tangan Otomatis Pada Video Percakapan Bahasa Isyarat Indonesia Menggunakan Metode Deep Gated Recurrent Unit (GRU)," J. Komput. Terap., vol. 8, no. 1, pp. 186-193, 2022, doi: 10.35143/jkt.v8i1.4901.