

Implementasi YOLOv8 Pada Pengenalan Sistem Isyarat Bahasa Indonesia

Diterima: 10 Juni 2024
Revisi: 10 Juli 2024
Terbit: 1 Agustus 2024

^{1*}Ridho Kuncoro Adji Wibowo, ²Ardi Sanjaya, ³Umi Mahdiyah
¹⁻³Universitas Nusantara PGRI Kediri
¹kuncoroadjie17@gmail.com, ²ardisanjaya@unpkediri.ac.id,
³umimahdiyah@unpkediri.ac.id

Abstrak—Penelitian ini mengembangkan sistem pengenalan abjad Sistem Isyarat Bahasa Indonesia (SIBI) menggunakan YOLOv8 untuk meningkatkan komunikasi antara masyarakat umum dan penyandang tunarungu ataupun tunawicara. Dihasilkan 4 model dengan beberapa variasi parameter meliputi *Learning Rate*, *Batch Size*, dan *Epoch* untuk menentukan kombinasi optimal. Hasil menunjukkan model keempat dengan *Batch Size* 16 dan *Learning Rate* 0,001 dengan 20 *Epoch* memberikan hasil terbaik dengan mAP 0.853. Temuan ini penting untuk meningkatkan akurasi dan efisiensi deteksi isyarat abjad SIBI, serta membantu memperbaiki interaksi sosial penyandang tunarungu ataupun tunawicara

Kata Kunci—bahasa isyarat; pengenalan abjad; tunarungu; tunawicara; YOLOv8

Abstract— *This research develops an alphabet recognition system for the Indonesian Language Sign System (SIBI) using the YOLOv8 algorithm to improve communication between the general public and deaf / hard of hearing people. Four models were generated with several parameter variations including Learning Rate, Batch Size, and number of Epochs to determine the optimal combination. The results showed that the fourth model with Batch Size 16 and Learning Rate 0.001 with 20 Epochs gave the best result with mAP 0.853. These findings are important to improve the accuracy and efficiency of SIBI alphabetic sign detection, helping to improve the social interaction of deaf/blind people.*

Keywords— *sign language; alphabet recognition; deaf, speech impaired; YOLOv8*

This is an open access article under the CC BY-SA License.



Penulis Korespondensi:

Ridho Kuncoro Adji Wibowo,
Teknik Informatika,
Universitas Nusantara PGRI Kediri,
Email: kuncoroadjie17@gmail.com
ID Orcid: [<https://orcid.org/register>]
Handphone: 085158525432

I. PENDAHULUAN

Manusia sebagai makhluk sosial memiliki kebutuhan untuk berkomunikasi dan berinteraksi dengan orang lain untuk menyampaikan pikiran ataupun perasaan. Bahasa sebagai alat komunikasi dapat dilakukan secara verbal (lisan dan tulisan) dan non verbal (isyarat, ekspresi wajah, gerakan tubuh)[1] Penyandang tunarungu dan tunawicara menggunakan bahasa isyarat atau bahasa non verbal sebagai sarana berkomunikasi [2].

Bahasa isyarat menggunakan gerakan tangan yang membentuk simbol atau isyarat dengan aturan tertentu untuk menyampaikan makna [3]. Di Indonesia, terdapat dua jenis bahasa isyarat yang umum digunakan, yaitu Sistem Isyarat Bahasa Indonesia (SIBI) dan Bahasa Isyarat Indonesia (BISINDO). SIBI telah distandarisasi oleh Pemerintah untuk digunakan pada Sekolah Luar Biasa (SLB) dan berbagai institusi di Indonesia. SIBI menggunakan satu tangan dalam pembentukan isyarat [4].

Komunikasi antara masyarakat umum dengan penyandang tunarungu atau tunawicara sering terkendala karena kurangnya pemahaman bahasa isyarat [5]. Mayoritas masyarakat umum tidak mempelajari bahasa isyarat karena tidak menggunakannya untuk berkomunikasi sehari-hari, sehingga menyebabkan kesalahpahaman dan menghambat interaksi sosial antar kedua belah pihak. Untuk itu, diperlukan pengembangan model deteksi untuk mengenali abjad SIBI.

Pada beberapa penelitian sebelumnya pernah menerapkan metode YOLOv5 untuk pengenalan bahasa isyarat SIBI [6], CNN untuk pengembangan aplikasi deteksi abjad isyarat SIBI [7]. YOLOv3 untuk deteksi gestur tangan berbasis citra [8], YOLO dan CNN untuk deteksi tangan otomatis pada video percakapan bahasa isyarat [9], CNN untuk pengenalan alfabet SIBI sebagai media pembelajaran bagi masyarakat umum [10].

Penelitian ini mengusulkan pengenalan bahasa isyarat SIBI menggunakan YOLOv8 dengan penggunaan parameter yang berbeda seperti *Learning Rate*, *Batch Size* dan *Epoch* untuk mendapatkan kombinasi parameter terbaik sehingga dapat memaksimalkan deteksi YOLOv8. Hasil analisa dan perbandingan ini dapat menjadi referensi bagi penelitian selanjutnya dalam menentukan kombinasi parameter YOLOv8 pada deteksi bahasa isyarat SIBI.

II. METODE

Adapun tahapan penelitian yang dilakukan untuk pengenalan abjad Sistem Isyarat Bahasa Indonesia (SIBI) menggunakan metode *You Look Only Once* versi 8 adalah sebagai berikut :

2.1. Studi Literatur

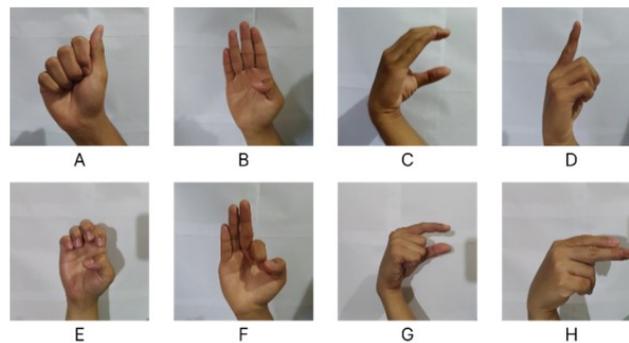
Penelitian ini diawali dengan eksplorasi literatur yang berkaitan mengenai pengolahan citra menggunakan YOLO. Informasi yang diperoleh tersebut digunakan untuk membangun dasar teoritis dan konteks penelitian.

2.2. Persiapan data

Dataset yang digunakan berasal dari Kaggle [11]. Dataset tersebut terdiri dari gambar isyarat abjad SIBI A-Z terdiri dari 26 kelas dan berjumlah 520 gambar. Spesifikasi dataset dapat dilihat pada Tabel 1 dibawah ini

Tabel 1. Spesifikasi Dataset

| Spesifikasi | Nilai |
|--------------------------------|-------------|
| Resolusi | 2000 x 2000 |
| Ekstensi | .jpg |
| Jumlah gambar | 520 |
| Jumlah <i>class</i> | 26 |
| Jumlah gambar per <i>class</i> | 20 |
| Channel | 3 (RGB) |



Gambar 1. Sampel Dataset

2.3 Preprocessing data

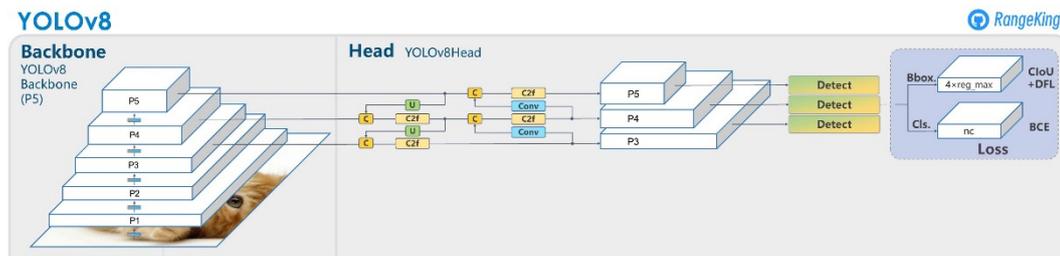
Tahap ini melibatkan beberapa proses yakni, *Resizing* gambar untuk mengubah ukuran gambar menjadi dimensi yang sama atau konsisten, augmentasi seperti *Flip* horizontal dan rotasi antara 0 – 20 derajat untuk membuat dataset agar lebih beragam, anotasi gambar dengan cara membuat kotak pembatas dan memberi kelas yang sesuai pada objek agar sesuai dengan format yang dibutuhkan oleh YOLO. Setiap citra yang telah dianotasi akan menghasilkan *file.txt* yang berisi data anotasi [kelas, x, y, lebar, tinggi]. Anotasi tersebut dilakukan menggunakan Roboflow.

2.4 Pembagian Data

Dataset kemudian dibagi menjadi data latih, dan data uji dengan rasio 8:2. Data latih digunakan untuk melatih model agar dapat mengenali pola dan fitur bahasa isyarat. Data validasi digunakan untuk melihat kemampuan model pada saat model dilatih apakah dapat mengenal pola secara umum atau belum dan memastikan model memberikan hasil yang konsisten dan akurat sebelum siap digunakan. Data uji digunakan untuk menguji peforma model dengan data nyata yang tidak digunakan untuk pelatihan [12]

2.4 YOLOv8

YOLO atau *You Look Only Once* versi 8 merupakan sebuah model *Deep Learning* yang termasuk dalam *detector* satu tahap karena langsung mendeteksi *Bounding Boxes* dan kelas objek dalam satu langkah dari *Input* gambar. Pendekatan tersebut membuat YOLOv8 sangat cepat dan efisien sehingga cocok untuk deteksi objek secara *Real-Time* [13] Arsitektur YOLOv8 terdiri dari tiga komponen utama yaitu *Backbone*, *Neck* dan *Head*.



Gambar 2. Arsitektur YOLOv8

Backbone berfungsi untuk ekstraksi fitur dari gambar Input menggunakan beberapa lapisan konvolusi dan modul C2f, menghasilkan fitur map pada berbagai resolusi (320x320, 160x160, 80x80, 40x40, dan 20x20). *Neck* mengumpulkan dan menggabungkan fitur dari berbagai tingkat *backbone* untuk memperkaya representasi fitur. Lapisan dalam Neck termasuk C2f, *Upsample*, dan *Concat*, yang membantu menangkap informasi multiskala dari gambar. Ukuran Output pada Neck adalah 80x80 dengan 768 *Filter*, 40x40 dengan 512 *Filter*, dan 20x20 dengan 256 *Filter*. *Head* bertanggung jawab untuk menghasilkan prediksi akhir yang mencakup *Bounding Box*, kelas objek, dan *Confidence Score* [14] Lapisan deteksi pada *Head* dirancang untuk mendeteksi objek pada berbagai skala, dengan ukuran *Output* 80x80, 40x40, dan 20x20.

Alur proses deteksi YOLOv8 membagi gambar menjadi *grid* dengan ukuran sel $S \times S$. Apabila titik pusat objek terdeteksi didalam suatu sel, maka sel tersebut bertanggung jawab untuk memprediksi *Bounding Box* dan kelas dari objek tersebut. Untuk menangani banyak prediksi *Bounding Box* yang tumpang tindih pada objek yang sama menggunakan *Non-Maximum Supression* (NMS). *Bounding Box* dengan *Confidence Score* tertinggi akan dibandingkan dengan *Bounding Box* lain dengan perhitungan *Intersection over Union* (IoU). IoU mengukur seberapa tumpang tindih tiap *Bounding Box* satu dengan yang lain. Nilai IoU memiliki rentang 0 sampai 1. Jika nilai IoU melebihi nilai batas yang telah ditentukan maka *Bounding Box* tersebut dianggap tumpang tindih kemudian dihapus, sehingga menyisakan prediksi *Bounding Box* yang relevan serta memiliki *Confidence* tertinggi untuk mewakili objek yang terdeteksi [15]

2.5 Evaluasi

Bagian ini bertujuan untuk mengevaluasi performa model YOLOv8 dalam mengenali abjad Sistem Isyarat Bahasa Indonesia yang telah dilatih menggunakan berbagai kombinasi parameter untuk menentukan kombinasi terbaik sehingga menghasilkan performa model yang maksimal. Parameter yang divariasikan meliputi *Batch Size*, *Learning Rate*, dan jumlah *Epoch*. Untuk mengukur performa model, digunakan *Confusion Matrix* yang gambaran menyeluruh tentang hasil prediksi model. Berikut adalah rumus untuk setiap komponen dalam *Confusion Matrix*:

| | | | |
|----------------|----------|----------------|----------------|
| | | Nilai Aktual | |
| | | Positive | Negative |
| Nilai Prediksi | Positive | True Positive | False Positive |
| | Negative | False Negative | True Negative |

Gambar 3. *Confusion Matrix*

True Positive (TP) adalah jumlah sampel yang sebenarnya positif dan berhasil diklasifikasikan dengan benar sebagai positif oleh model. *True Negative* (TN) menunjukkan jumlah sampel yang sebenarnya negatif dan berhasil diklasifikasikan negatif oleh model. *False Positive* (FP) adalah jumlah sampel yang sebenarnya negatif namun diklasifikasikan secara salah sebagai positif. *False Negative* (FN) adalah jumlah sampel yang sebenarnya positif namun diklasifikasikan sebagai negatif oleh model.

Untuk mengukur model menggunakan *Mean Average Precision* (mAP) yang merupakan rata-rata dari nilai *Average Precision* (AP) untuk semua kelas. AP mengukur akurasi model dalam deteksi dan klasifikasi dengan menghitung luas di bawah kurva *Precision-Recall* (AUC-PR) untuk setiap kelas [16]. AP biasanya dihitung sebagai:

$$AP = \int_0^1 p(r)dr \quad (1)$$

Dimana $p(r)$ adalah precision sebagai fungsi dari recall r .

1. *Mean Average Precision* (mAP). Rata-rata dari AP untuk semua kelas:

di mana

$$mAP = \frac{1}{N} \sum_{i=1}^N AP_i \quad (2)$$

N adalah jumlah kelas dan AP

AP_i adalah *Average Precision* untuk kelas ke- i .

III. HASIL DAN PEMBAHASAN

Pada penelitian ini akan dipaparkan hasil dari pelatihan model YOLOv8 untuk pengenalan abjad SIBI. Percobaan dilakukan dengan membuat beberapa model dengan variasi parameter *Batch Size*, *Learning Rate*, dan jumlah *Epoch* untuk melihat pengaruhnya terhadap performa model. Berikut rincian dari hasil dan analisis percobaan tersebut.

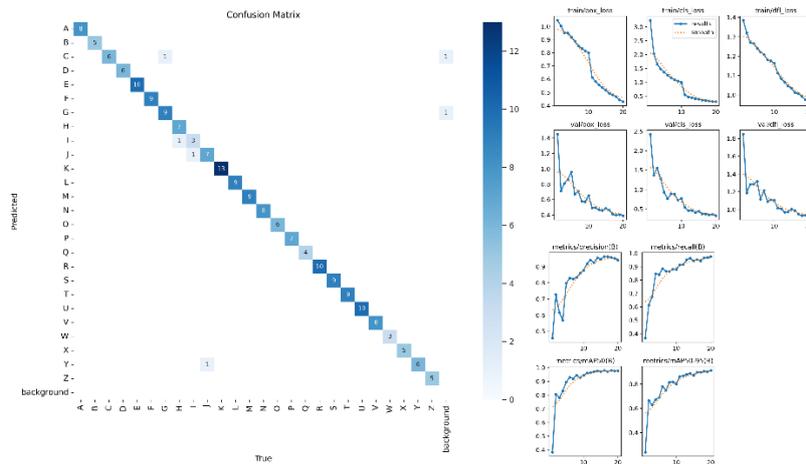
Tabel 2. Hasil Evaluasi Model

| Model | Learning Rate | Batch Size | Epoch | mAP |
|-------|---------------|------------|-------|-------|
| 1 | 0.001 | 16 | 20 | 0.899 |
| 2 | 0.0001 | 16 | 20 | 0.907 |
| 3 | 0.001 | 32 | 20 | 0.913 |
| 4 | 0.0001 | 32 | 20 | 0.899 |

Hasil evaluasi model menunjukkan bahwa model 3 dengan *Learning Rate* sebesar 0.001 dan *Batch Size* 32 memiliki nilai mAP tertinggi yaitu 0.913. Model 2 dengan *Learning Rate* sebesar 0.0001 dan *Batch Size* sebesar 16 juga menunjukkan performa yang cukup baik dengan mAP 0.907. Model 1 dan model 4 masing-masing memiliki mAP yang sama yaitu 0.899, meskipun memiliki kombinasi *Learning Rate* dan *Batch Size* yang berbeda.

Dari hasil tersebut, dapat disimpulkan bahwa *Batch Size* yang lebih besar yakni 32 cenderung menghasilkan mAP yang lebih tinggi dibandingkan dengan *Batch Size* yang lebih kecil yakni 16, terlepas dari nilai *Learning Rate* yang digunakan. *Batch Size* yang lebih besar memungkinkan model untuk mempelajari pola yang lebih beragam dalam setiap iterasi, sehingga dapat meningkatkan kemampuan generalisasi model. Selain itu, *Learning Rate* yang lebih besar tampaknya lebih efektif ketika digabungkan dengan *Batch Size* yang lebih besar, seperti yang terlihat pada model 3. *Learning Rate* yang lebih besar memungkinkan model untuk melakukan update parameter yang lebih signifikan, sehingga dapat mempercepat proses konvergensi ke solusi optimal.

Selain itu, dari hasil *Confusion Matrix* yang dapat dilihat pada gambar 4 menunjukkan performa model yang baik dalam deteksi objek secara benar dengan tren metrik seperti *loss* yang menurun serta *precision*, *recall*, *mAP* yang meningkat seiring dengan *Epoch* pada grafik evaluasi pada gambar 4. Selain itu, pada gambar 5 menunjukkan hasil deteksi isyarat abjad yang baik dengan *bounding box*, serta nilai *confidence* yang tinggi.



Gambar 4. *Confusion Matrix* dan Grafik Evaluasi



Gambar 5. Hasil Deteksi Isyarat Tangan

IV. KESIMPULAN

Penelitian ini bertujuan untuk mengimplementasikan YOLOv8 dalam pengenalan abjad Sistem Isyarat Bahasa Indonesia dan mengevaluasi performanya dengan berbagai kombinasi parameter pelatihan. Dari hasil penelitian, dapat disimpulkan bahwa model pertama dengan dengan parameter *Batch Size* sebanyak 16 dan *Learning Rate* sebesar 0,001 selama 20 *Epoch* menunjukkan performa terbaik, menghasilkan mAP sebesar 0.853. Kombinasi ini menghasilkan deteksi yang akurat dan stabil.

Sebagai saran, penelitian selanjutnya dapat fokus pada peningkatan dataset dengan variasi lebih banyak, serta pengujian model dalam kondisi pencahayaan dan latar belakang yang lebih bervariasi untuk meningkatkan kemampuan deteksi. Selain itu, eksplorasi parameter lain dapat dilakukan untuk mencapai performa yang lebih tinggi.

DAFTAR PUSTAKA

- [1] O. Mailani, I. Nuraeni, S. A. Syakila, and J. Lazuardi, "Bahasa Sebagai Alat Komunikasi Dalam Kehidupan Manusia," *Kampret Journal*, vol. 1, no. 1, pp. 1–10, Jan. 2022, doi: 10.35335/kampret.v1i1.8.
- [2] A. S. Nugraheni, A. P. Husain, and H. Unayah, "OPTIMALISASI PENGGUNAAN BAHASA ISYARAT DENGAN SIBI DAN BISINDO PADA MAHASISWA DIFABEL TUNARUNGU DI PRODI PGMI UIN SUNAN KALIJAGA," *Jurnal Holistika*, vol. 5, no. 1, p. 28, Feb. 2023, doi: 10.24853/holistika.5.1.28-33.
- [3] D. I. Mulyana, M. F. Lazuardi, and M. B. Yel, "Deteksi Bahasa Isyarat Dalam Pengenalan Huruf Hijaiyah Dengan Metode YOLOV5," *Jurnal Teknik Elektro dan Komputasi (ELKOM)*, vol. 4, no. 2, pp. 145–151, Aug. 2022, doi: 10.32528/elkom.v4i2.8145.
- [4] M. Zikky and Z. F. Akbar, "Kamus Sistem Isyarat Bahasa Indonesia (KASIBI) dengan Voice Recognition sebagai Pendukung Belajar Bahasa Isyarat Berbasis Android," *JST (Jurnal Sains Terapan)*, vol. 5, no. 2, Oct. 2019, doi: 10.32487/jst.v5i2.732.
- [5] S. A. Sanjaya and H. Faustine Ilone, "BISINDO Sign Language Recognition: A Systematic Literature Review of Deep Learning Techniques for Image Processing," *Indonesian Journal of Computer Science*, vol. 12, no. 6, Dec. 2023, doi: 10.33022/ijcs.v12i6.3539.
- [6] D. Permana and J. Sutopo, "APLIKASI PENGENALAN ABJAD SISTEM ISYARAT BAHASA INDONESIA (SIBI) DENGAN ALGORITMA YOLOv5," *Jurnal Simantec*, vol. 11, no. 2, pp. 231–240, Jul. 2023, doi: 10.21107/simantec.v11i2.19783.
- [7] M. Sholawati, K. Auliasari, and FX. Ariwibisono, "PENGEMBANGAN APLIKASI PENGENALAN BAHASA ISYARAT ABJAD SIBI MENGGUNAKAN METODE CONVOLUTIONAL NEURAL NETWORK (CNN)," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 6, no. 1, pp. 134–144, Mar. 2022, doi: 10.36040/jati.v6i1.4507.
- [8] A. Sani and S. Rahmadinni, "Deteksi Gestur Tangan Berbasis Pengolahan Citra," *Jurnal Rekayasa Elektrika*, vol. 18, no. 2, 2022.
- [9] I. I. Arifah, F. N. Fajri, and G. Q. O. Pratamasunu, "Deteksi Tangan Otomatis Pada Video Percakapan Bahasa Isyarat Indonesia Menggunakan Metode YOLO Dan CNN," *Journal of Applied Informatics and Computing*, vol. 6, no. 2, pp. 171–176, Nov. 2022, doi: 10.30871/jaic.v6i2.4694.
- [10] I. I. Arifah, F. N. Fajri, and G. Q. O. Pratamasunu, "Deteksi Tangan Otomatis Pada Video Percakapan Bahasa Isyarat Indonesia Menggunakan Metode YOLO Dan CNN," *Journal of Applied Informatics and Computing*, vol. 6, no. 2, pp. 171–176, Nov. 2022, doi: 10.30871/jaic.v6i2.4694.
- [11] M. L. Afkaar, "Datasets SIBI Sign Language Alphabets," Kaggle.
- [12] G. B. Prananta, H. A. Azzikri, and C. Rozikin, "REAL-TIME HAND GESTURE DETECTION AND RECOGNITION USING CONVOLUTIONAL ARTIFICIAL NEURAL NETWORKS," *METHODIKA: Jurnal Teknik Informatika dan Sistem Informasi*, vol. 9, no. 2, pp. 30–34, Sep. 2023, doi: 10.46880/mtk.v9i2.1911.
- [13] R. F. Putra and D. I. Mulyana, "Optimasi Deteksi Objek Dengan Segmentasi dan Data Augmentasi Pada Hewan Siput Beracun Menggunakan Algoritma You Only Look Once (YOLO)," *Jurnal JTIK (Jurnal Teknologi Informasi dan Komunikasi)*, vol. 8, no. 1, pp. 93–103, Jan. 2024, doi: 10.35870/jtik.v8i1.1391.
- [14] D. H. Kusuma and M. Mauizah, "Deteksi Lampu Lalu Lintas Menggunakan YOLO untuk Autonomous Car," *Journal of Computer Engineering, Network, and Intelligent Multimedia*, vol. 1, no. 1, pp. 21–27, Feb. 2023, doi: 10.59378/jcenim.v1i1.5.
- [15] T. Abuzairi, Nurdina Widanti, Arie Kusumaningrum, and Yeni Rustina, "Implementasi Convolutional Neural Network Untuk Deteksi Nyeri Bayi Melalui Citra Wajah Dengan YOLO," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 5, no. 4, pp. 624–630, Aug. 2021, doi: 10.29207/resti.v5i4.3184.
- [16] D. Arifadilah, Asriyanik, and A. Pambudi, "Sunda Script Detection Using You Only Look Once Algorithm," *Journal of Artificial Intelligence and Engineering Applications (JAIEA)*, vol. 3, no. 2, pp. 606–613, Feb. 2024, doi: 10.59934/jaiea.v3i2.443.