



Deteksi Objek Video Bahasa Isyarat Untuk Anak Tuna Rungu Dan Tuna Wicara Menggunakan YOLOv8

Muhammad Maheza Fresmanda¹, Istiadi², Syahroni Wahyu Iriananda³

^{1,2,3} Universitas Widyagama

Abstrak: Penelitian ini bertujuan untuk mengimplementasikan keberhasilan model algoritma YOLOv8 dalam mendeteksi abjad bahasa isyarat Indonesia (BISINDO) berbasis video. BISINDO merupakan adaptasi American Sign Language (ASL) yang disesuaikan dengan budaya Indonesia agar lebih mudah digunakan. Itu bergantung pada gerakan dua tangan untuk komunikasi. Model YOLOv8 mampu mendeteksi dan mengklasifikasikan bahasa isyarat Indonesia dengan tingkat akurasi yang cukup baik, mencapai precision 93,35%, recall 70,25%, mAP50 64,05% dan mAP50-95 45,88%. Temuan ini diharapkan dapat memberikan kontribusi positif dalam pendidikan dan komunikasi untuk anak tunarungu dan tunawicara.

Kata kunci: YOLOv8, Gerakan tangan, BISINDO

DOI: <https://doi.org/10.53697/jkomitek.v4i2.1895>

*Correspondence: Muhammad Maheza Fresmanda
Email: m.mahezaf3@gmail.com

Received: 18-10-2024

Accepted: 17-11-2024

Published: 01-12-2024



Copyright: © 2024 by the authors. Submitted for open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

Abstract : This study aims to implement the success of the YOLOv8 algorithm model in detecting Indonesian sign language alphabet (BISINDO) based on video. BISINDO is an adaptation of American Sign Language (ASL) that is adapted to Indonesian culture to make it easier to use. It relies on two-handed gestures for communication. The YOLOv8 model is able to detect and classify Indonesian sign language with a fairly good level of accuracy, achieving a precision of 93.35%, recall of 70.25%, mAP50 of 64.05% and mAP50-95 of 45.88%. These findings are expected to provide a positive contribution to education and communication for deaf and mute children.

Keywords: YOLOv8, Sign Language, BISINDO

Pendahuluan

Komunikasi melibatkan penyampaian ide atau pesan antar pihak, biasanya melalui cara verbal. Namun, bagi individu dengan gangguan pendengaran, dimana pendengaran tidak berfungsi secara normal, proses ini menjadi tantangan (Elyondri & Azizah, 2023; Rafikayati & Jauhari, 2020; Syah, 2022). Kondisi ini mempengaruhi interaksi sehingga pertukaran informasi menjadi sulit (Astriani, 2023; Izzalhaqqi, 2023; Limantoro, 2023; Yap, 2023). Penelitian menunjukkan bahwa lebih dari 360 juta orang di seluruh dunia, termasuk 328 juta orang dewasa dan 32 juta anak-anak, menderita tuli atau bisu, yang mencakup lebih dari 5% populasi global (Rasuandi, 2023)

Bahasa isyarat menggunakan isyarat manual, bahasa tubuh, dan gerakan bibir untuk berkomunikasi. Ini terutama digunakan oleh komunitas Tunarungu untuk berinteraksi,

tetapi juga berfungsi sebagai sarana bagi pengguna untuk mengekspresikan identitas mereka dan mengakses informasi. Di Indonesia, dikenal dua jenis bahasa isyarat: Sistem Bahasa Isyarat Indonesia (SIBI), yang terstandarisasi dan berasal dari Bahasa Isyarat Amerika (ASL) dengan tambahan awalan dan akhiran, dan Bahasa Isyarat Indonesia (BISINDO), yang merupakan bahasa asli. atau bahasa pertama yang biasa digunakan oleh komunitas Tunarungu dalam kehidupan sehari-hari (Susanty et al., 2021)

BISINDO merupakan adaptasi *American Sign Language* (ASL) yang disesuaikan dengan budaya Indonesia agar lebih mudah digunakan. Itu bergantung pada gerakan dua tangan untuk komunikasi (Abdullah, 2019; Zikky et al., 2019). SIBI merupakan bahasa isyarat turunan dari ASL dan resmi digunakan di seluruh sekolah luar biasa (SLB) di bawah Kementerian Pendidikan dan Kebudayaan. Penggunaannya diatur dengan Peraturan Menteri Pendidikan Nomor 0161/U/1994 (Rachardi, 2020)

Pada beberapa penelitian sebelumnya metode YOLOv8 (You Only Look Once Version 8) merupakan salah satu metode computer vision yang telah digunakan untuk mendeteksi isyarat dalam berbagai bidang, termasuk dalam pengenalan huruf Hijaiyah (Mulyana et al., 2022). Namun, penerapan metode YOLOv8 untuk mendeteksi isyarat tuna rungu dan tuna wicara masih terbatas dan belum diperkenalkan secara luas (Candra, 2024; Caraka, 2024; Saputra, 2024).

Tujuan dari konteks penelitian ini adalah untuk menerapkan metode YOLOv8 untuk mengidentifikasi spesies rungu dan wicara selain BISINDO. Model ini akan membantu meringankan kesulitan dalam berkomunikasi antara orang berkebutuhan khusus dan masyarakat umum serta mendorong pemahaman dan pengertian yang lebih positif antara masyarakat umum dan kelompok tuna rungu wicara (Ahmad, 2023; Andreas, 2023; Maheswara, 2023; Sutjiadi, 2023). Selain itu, diharapkan hasil penelitian ini akan meningkatkan kemampuan siswa untuk mempelajari dan memahami bahasa tuna rungu dan wicara, serta memberikan wawasan tentang potensi teknologi dalam komunikasi dan pendidikan (Enri, 2023; Joan, 2023; Sujatmiko, 2023).

Metode

Metode penelitian yang digunakan untuk deteksi objek dalam video Bahasa Isyarat Indonesia (BISINDO) menggunakan YOLOv8 adalah sebagai berikut:

Studi Literatur

Pada tahapan awal ini bertujuan untuk mempelajari literatur yang relevan dengan penelitian yang akan dilakukan. Studi pustaka dapat dilakukan dengan membaca literatur yang relevan, seperti buku, jurnal ilmiah, atau artikel online. Literatur yang dipelajari meliputi literatur tentang bahasa isyarat Indonesia, algoritma YOLOv8, dan metode deteksi objek video bahasa isyarat.

Pengumpulan Data

Langkah ini melibatkan pengumpulan sejumlah besar data yang akan digunakan untuk mendeteksi bahasa tulis Indonesia (BISINDO). Dua jenis data yang menyusun dataset tersebut adalah Data Primer dan Data Sekunder. Data Primer terdiri dari delapan video berdurasi satu hingga satu setengah menit yang direkam di UPT Rehabilitasi Bina Sosial Rungu Wicara Pasuruan, sedangkan Data Sekunder terdiri dari dua video berdurasi

satu setengah menit yang bersumber dari kagle.(Marília Prata, 2020). Berikut Gambaran dataset yang dimiliki

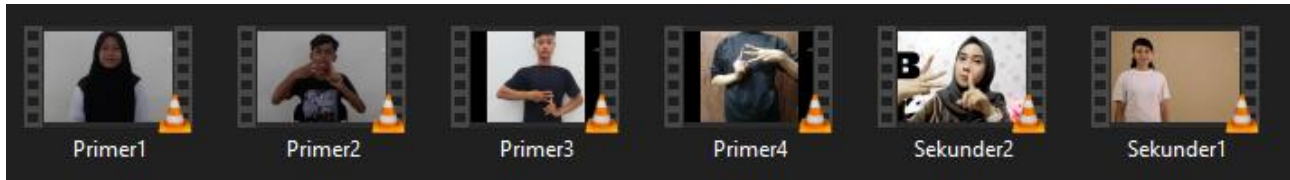


Figure 1. Dataset

Preprocessing Data

Langkah ini melibatkan beberapa proses, seperti pengetsaan frame untuk dijadikan frame tersendiri. Anotasi data dilakukan dengan membuat plot cross-hatch di mana setiap frame (atau gambar) yang dihasilkan dari video diberi label sesuai dengan gerakan tangan tertentu dan disesuaikan agar sesuai dengan format yang dibutuhkan oleh YOLOv8. Langkah selanjutnya melibatkan perubahan ukuran gambar agar sesuai dengan ukuran layar, penyempurnaan fitur seperti membalik secara horizontal dan vertikal, memutar antara -15 dan +15 derajat, menyesuaikan kecerahan antara -15 dan +15 derajat, mengaburkan hingga 2,5 persen, dan mengurangi noise hingga 100% pada layar. Pada fase ini, data yang terkumpul juga akan dibagi menjadi tiga kategori: data latih, data validasi, dan data penelitian. Tugas ini dilakukan dengan menggunakan Roboflow. Berikut pembagian datanya

Table 1: Split Ratio

Split Ratio		
Training	Validation	Testing
10%	10%	80%
20%	10%	70%
30%	10%	60%
40%	10%	50%
50%	10%	40%
60%	10%	30%
70%	10%	20%
80%	10%	10%

Pelatihan

Pada tahap ini mulai masuk ke proses pelatihan model. Pelatihan dilakukan di Google Colab menggunakan 100 epoch, ukuran gambar 640 piksel, ukuran batch 4, dan learning rate 1e-4. Waktu yang dihabiskan untuk setiap tahap per gambar adalah: 0.4 milidetik untuk pra-pemrosesan, 1.5 milidetik untuk inferensi, 0.0 milidetik untuk menghitung loss, dan 0.8 milidetik untuk pasca-pemrosesan yang meliputi penghapusan kotak pembatas duplikat dan perubahan format output.

Pengujian

Setelah proses pelatihan selesai, kemudian menerapkan model yang telah dilatih untuk mendeteksi objek dalam video yang diuji Pada tahap ini, model deteksi objek digunakan

untuk otomatis mengidentifikasi dan melabeli objek yang ada dalam setiap frame video. Proses ini sangat penting karena melibatkan penerapan model yang telah dilatih dalam situasi nyata untuk melihat seberapa baik model tersebut dapat mendeteksi objek dengan akurasi tinggi.

Evaluasi

Bagian ini bertujuan untuk mengevaluasi hasil kinerja model YOLOv8 dalam menganalisis terjemahan Bahasa Indonesia (BISINDO) yang telah diterjemahkan dan dievaluasi untuk menentukan hasil terbaik dan menghasilkan hasil yang paling banyak. Confusion Matrix digunakan dalam evaluasi kinerja model untuk membuat prediksi model. Berikut Gambaran rumus dari *Confusion Matrix*:

		Nilai Aktual	
		Positive	Negative
Nilai Prediksi	Positive	TP	FP
	Negative	FN	TN

Figure 2. Confusion Matrix

True Positive (TP) mengacu pada jumlah sampel yang jelas positif dan berhasil diklasifikasikan sebagai positif oleh model. True Negative (TN) mengukur jumlah sampel yang jelas negatif dan berhasil diklasifikasikan sebagai negatif oleh model. False Positive (FP) mengacu pada kuantitas sampel yang jelas negatif tetapi tetap diklasifikasikan sebagai positif. False Negative (FN) mengacu pada jumlah sampel yang tampak positif tetapi diklasifikasikan sebagai negatif oleh model.

Sedangkan untuk mengukur model menggunakan *Precision*, *Recall*, dan *Mean Average Precision* (mAP). Berikut rumus-rumus persamaan berikut:

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

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

$$mAp = \frac{1}{N} \int_{i=1}^N AP_i$$

Keterangan: N = Jumlah data AP | AP = Average Precision

Hasil dan Pembahasan

Dalam penelitian ini, hasil pelatihan model YOLOv8 akan dibahas terkait pengembangan huruf BISINDO. Percobaan dilakukan dengan membangun beberapa model yang menggunakan parameter Batch Size dan jumlah epoch untuk menilai

dampaknya terhadap kinerja model. Berikut ini adalah hasil dan analisis percobaan tersebut. Berikut adalah rincian hasil dan analisis dari percobaan tersebut.

Table 2: Hasil Evaluasi Model

Split Ratio	Precision	Recall	mAP50	mAP50-95
80%/10%/10%	0,956	0,96	0,96	0,729
70%/10%/20%	0,909	0,84	0,823	0,642
60%/10%/30%	1,00	0,7	0,699	0,438
50%/10%/40%	1,00	0,81	0,72	0,501
40%/10%/50%	0,987	0,8	0,675	0,424
30%/10%/60%	0,934	0,7	0,487	0,359
20%/10%/70%	0,725	0,38	0,375	0,293
10%/10%/80%	0,957	0,43	0,385	0,284
Rata-Rata	0,9335	0,7025	0,6405	0,45875

Skor analisis data sebesar 80% untuk pelatihan, 10% untuk validasi, dan 10% untuk pengujian menghasilkan hasil terbaik, dengan presisi 0,956 dan recall 0,96. Kinerja model pada mAP50 dan mAP50-95 masing-masing mencapai 0,96 dan 0,729, yang menunjukkan tingkat efektivitas deteksi objek yang tinggi. Sebaliknya, rasio 20%/10%/70% menghasilkan yield tertinggi, dengan Presisi 0,725 dan Recall hanya 0,38, yang menunjukkan bahwa model mengalami kesulitan dalam melakukan deteksi yang akurat. Perincian kinerja ini secara jelas menunjukkan betapa pentingnya memiliki data pelatihan yang cukup untuk mencapai hasil deteksi yang baik. Secara umum, semakin banyak data yang diberikan untuk pelatihan, semakin tinggi pula kinerja model, seperti yang terlihat pada metrik rata-rata yang diperoleh: Precision 0,9335, Recall 0,7025, mAP50 0,6405, dan mAP50-95 0,45875.

Analisis lebih lanjut menunjukkan bahwa model tersebut memiliki kinerja yang baik dalam pendeteksian objek, seperti yang ditunjukkan oleh matriks kebingungan metrik (Gambar 3) dan skala evaluasi metrik (Gambar 4). Visualisasi hasil pendeteksian pada Gambar 5 juga memperkuat gagasan ini, dengan kotak pembatas yang ketat dan tingkat kepercayaan yang tinggi.

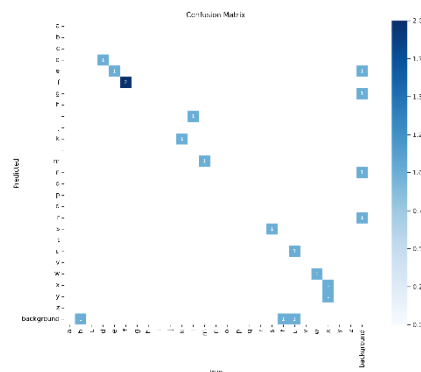


Figure 3. Confusion Matrix

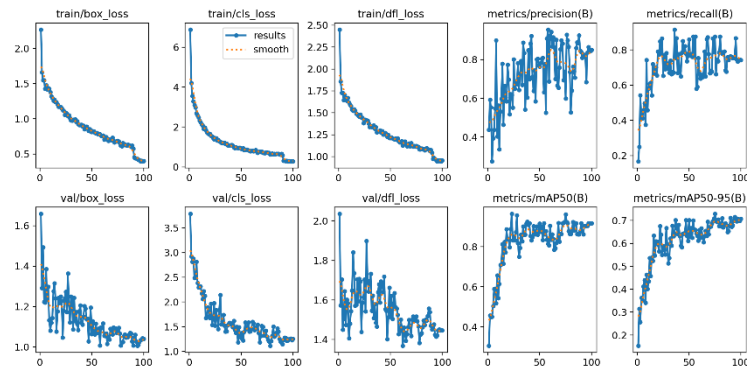


Figure 4. Grafik Evaluasi



Figure 5. Hasil Deteksi Bahasa Isyarat Indonesia

Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan bahwa model yang dikembangkan menunjukkan kinerja yang menjanjikan. Model YOLOv8 mampu mendeteksi dan mengklasifikasikan bahasa isyarat Indonesia dengan tingkat akurasi yang cukup baik, mencapai precision 93,35%, recall 70,25%, mAP50 64,05% dan mAP50-95 45,88%. Tersedia matriks untuk pengembangan dan penelitian lebih lanjut berdasarkan temuan studi dan data yang dikumpulkan. Fitur utamanya meliputi perluasan kumpulan data dengan lebih banyak variabel, pengoptimalan model hiperparameter, kelas ketidakseimbangan penanganan, pengembangan aplikasi waktu nyata, evaluasi khusus konteks dengan pengguna runtu dan wicara tuna, eksplorasi model arsitektur eksploratori, dan interpretasi model yang disempurnakan.

Referensi

- Marília Prata. (2020). *Bisindo Indonesian Sign Language MP4*. <https://www.kaggle.com/code/mpwolke/bisindo-indonesian-sign-language-mp4/notebook>.
- Mulyana, D. I., Lazuardi, M. F., & Yel, M. B. (2022). Deteksi Bahasa Isyarat Dalam Pengenalan Huruf Hijaiyah Dengan Metode YOLOV5. *Jurnal Teknik Elektro Dan Komputasi (ELKOM)*, 4(2), 145–151. <https://doi.org/10.32528/elkom.v4i2.8145>
- Rachardi, F. (2020). Deteksi Gambar Gestur Kosakata Bahasa Isyarat Indonesia dengan Convolutional Neural Network. *Institutional Repository UIN Syarif Hidayatullah*, 192. <https://repository.uinjkt.ac.id/dspace/handle/123456789/56075>
- Rasuandi, M. (2023). *PENGENALAN ALFABET A-Z BAHASA ISYARAT AMERICAN SIGN LANGUAGE MENGGUNAKAN HALAMAN PENGESAHAN SKRIPSI PENGENALAN ALFABET A-Z BAHASA ISYARAT AMERICAN SIGN LANGUAGE MENGGUNAKAN METODE SUPPORT VECTOR MACHINE*.
- Susanty, M., Fadillah, R. Z., & Irawan, A. (2021). Model Penerjemah Bahasa Isyarat Indonesia (BISINDO) Menggunakan Pendekatan Transfer Learning. *Petir*, 15(1), 1–9. <https://doi.org/10.33322/petir.v15i1.1289>
- Abdullah, A. (2019). Strategi Penanganan Guru Pendidikan Agama Islam Bagi Anak Berkebutuhan Khusus Di Sekolah Dasar Luar Biasa Negeri Keleyan Socah Bangkalan. *Al-Ibrah: Jurnal Pendidikan Dan Keilmuan Islam*. <https://www.ejournal.stital.ac.id/index.php/alibrah/article/view/81>
- Ahmad, N. (2023). Transforming Sign Language using CNN Approach based on BISINDO Dataset. *2023 International Conference on Informatics, Multimedia, Cyber and Information Systems, ICIMCIS 2023*, 543–548. <https://doi.org/10.1109/ICIMCIS60089.2023.10349011>
- Andreas, R. (2023). Mobile Application for Children to Learn BISINDO Sign Language. *6th International Conference on Inventive Computation Technologies, ICICT 2023 - Proceedings*, 774–780. <https://doi.org/10.1109/ICICT57646.2023.10134183>
- Astriani, M. S. (2023). TELEMEDICINE SIGN LANGUAGE CLASSIFICATION FOR COVID-19 PATIENTS WITH DISABILITY BASED ON LSTM MODEL. *Journal of Engineering Science and Technology*, 18, 140–152.
- Candra, A. (2024). Development of Machine Learning-based Sign Language Translator for Bahasa Isyarat Indonesia (BISINDO). *AIP Conference Proceedings*, 2987(1). <https://doi.org/10.1063/5.0199747>
- Caraka, R. E. (2024). Empowering deaf communication: a novel LSTM model for recognizing Indonesian sign language. *Universal Access in the Information Society*. <https://doi.org/10.1007/s10209-024-01095-1>
- Elyondri, N., & Azizah, N. (2023). Analisis Pengembangan Komunikasi, Persepsi, Bunyi, dan Irama (PKPBI) Anak Tunarungu dan Kebutuhan Media Pembelajarannya. *Jurnal Obsesi: Jurnal Pendidikan Anak Usia* <https://www.academia.edu/download/108761726/pdf.pdf>

- Enri, U. (2023). Sign Language Detection Using Mediapipe and Long-Short Term Memory Network. *2023 International Conference on Informatics, Multimedia, Cyber and Information Systems, ICIMCIS 2023*, 617–622. <https://doi.org/10.1109/ICIMCIS60089.2023.10349016>
- Izzalhaqqi, M. Y. D. (2023). Gesture Recognition in Indonesian Sign Language Using Hybrid Deep Learning Models. *Proceedings - IWIS 2023: 3rd International Workshop on Intelligent Systems*. <https://doi.org/10.1109/IWIS58789.2023.10284666>
- Joan, D. (2023). BISINDO Hand-Sign Detection Using Transfer Learning. *8th International Conference on Recent Advances and Innovations in Engineering: Empowering Computing, Analytics, and Engineering Through Digital Innovation, ICRAIE 2023*. <https://doi.org/10.1109/ICRAIE59459.2023.10468194>
- Limantoro, E. (2023). Indonesian Sign Language Recognition Using Kinect Sensor Based on Deep Neural Network. *AIP Conference Proceedings*, 2689(1). <https://doi.org/10.1063/5.0114983>
- Maheswara, Y. D. (2023). Real-Time BISINDO Sign Language Recognition: A Dynamic Approach with GRU and LSTM Models Leveraging MediaPipe. *6th International Seminar on Research of Information Technology and Intelligent Systems, ISRITI 2023 - Proceeding*, 226–232. <https://doi.org/10.1109/ISRITI60336.2023.10467586>
- Rafikayati, A., & Jauhari, M. N. (2020). Studi tentang Pelaksanaan Pelatihan Terapi Wicara Anak Tunarungu Usia 3-5 Tahun untuk Orangtua dalam Setting Blended Learning di SLB Karya Mulia Surabaya. *Jurnal Leverage, Engagement* <https://journal.uc.ac.id/index.php/LeECOM/article/view/1417>
- Saputra, M. A. (2024). Recognizing Indonesian sign language (Bisindo) gesture in complex backgrounds. *Indonesian Journal of Electrical Engineering and Computer Science*, 36(3), 1583–1593. <https://doi.org/10.11591/ijeecs.v36.i3.pp1583-1593>
- Sujatmiko, D. (2023). AlexNet Architecture Based Convolution Neural Network for Realtime Audio to Text Translator of Bisindo Hand Sign. *2023 International Seminar on Application for Technology of Information and Communication: Smart Technology Based on Industry 4.0: A New Way of Recovery from Global Pandemic and Global Economic Crisis, ISemantic 2023*, 429–434. <https://doi.org/10.1109/iSemantic59612.2023.10295322>
- Sutjiadi, R. (2023). Android-Based Application for Real-Time Indonesian Sign Language Recognition Using Convolutional Neural Network. *TEM Journal*, 12(3), 1541–1549. <https://doi.org/10.18421/TEM123-35>
- Syah, A. R. (2022). APLIKASI PENERJEMAH BAHASA ISYARAT MENGGUNAKAN METODE K-NN (K-NEAREST NEIGHBOUR). *Jurnal Teknologi Pintar*. <http://teknologipintar.org/index.php/teknologipintar/article/view/143>
- Yap, S. (2023). Enhancing BISINDO Recognition Accuracy Through Comparative Analysis of Three CNN Architecture Models. *Proceedings of 2023 International Conference on Information Management and Technology, ICIMTech 2023*, 732–737. <https://doi.org/10.1109/ICIMTech59029.2023.10277780>
- Zikky, M., Akbar, Z. F., & Utomo, S. (2019). Kamus sistem isyarat bahasa Indonesia (KASIBI)

dengan voice recognition sebagai pendukung belajar bahasa isyarat berbasis android.
JST (Jurnal Sains Terapan). <http://jurnal.poltekba.ac.id/index.php/jst/article/view/732>