

PENDETEKSI KATA DALAM BAHASA ISYARAT MENGGUNAKAN ALGORITMA YOLO VERSI 8

Dennis Saputra Ariansyah^{1*}

¹Universitas Muhammadiyah Sukabumi; Jl. R. Syamsudin, S.H. No. 50, Cikole; Telpon/Fax (0266) 218345

Received: 12 Juli 2024

Accepted: 31 Juli 2024

Published: 7 Agustus 2024

Keywords:

YOLOv8; Deteksi; Bahasa isyarat; Real-time.

Correspondent Email:

dennisptr063@ummi.ac.id

Abstrak. Bahasa isyarat adalah sistem komunikasi yang digunakan oleh individu dengan gangguan pendengaran atau bicara, mengandalkan gerakan tangan, ekspresi wajah, dan gerakan tubuh. Di Indonesia, terdapat dua sistem bahasa isyarat: SIBI dan BISINDO, dengan BISINDO lebih umum digunakan oleh tunarungu dan tunawicara karena lebih mudah dipahami. Saat ini, Indonesia memiliki lebih dari 223.000 penyandang tuli dan 73.500 penyandang bisu dan tuli, yang sering mengalami kesulitan dalam berkomunikasi di tempat umum. Rata-rata masyarakat belum memahami bahasa isyarat, disebabkan kurangnya sumber informasi yang memadai. Masyarakat biasanya hanya dapat memahami melalui kamus, yang dimana cara tersebut tidak cukup efektif. Penelitian ini bertujuan mengembangkan sebuah model machine learning untuk deteksi bahasa isyarat secara real-time menggunakan algoritma YOLOv8. Dataset yang digunakan terdiri dari 25.000 gambar, di mana dilakukan perbandingan parameter dan variasi YOLOv8. Model terbaik diperoleh dari varian YOLOv8m dengan akurasi 93,8%, menggunakan optimizer Adam dengan learning rate 0.001 sebagai parameter optimisasi yang optimal.

Abstract. Sign language is a communication system used by individuals with hearing or speech impairments, relying on hand movements, facial expressions, and body movements. In Indonesia, there are two sign language systems: SIBI and BISINDO, with BISINDO being more commonly used by the deaf and mute community because it is easier to understand. Currently, Indonesia has over 223,000 deaf individuals and 73,500 individuals who are both deaf and mute, who often face difficulties communicating in public spaces. The average person lacks an understanding of sign language due to insufficient information sources. Generally, people can only understand through dictionaries, which is not an effective method. This research aims to develop a machine learning model for real-time sign language detection using the YOLOv8 algorithm. The dataset consists of 25,000 images, and comparisons were made regarding the parameters and variations of YOLOv8. The best model was obtained from the YOLOv8m variant with an accuracy of 93.8%, using the Adam optimizer with a learning rate of 0.001 as the optimal optimization parameter.

1. PENDAHULUAN

Manusia hidup dengan berdampingan dan akan butuh berintraksi serta berkomunikasi dengan sesama, umumnya mereka akan berintraksi menggunakan bahasa lisan [1].

Tidak semua manusia dapat berinteraksi menggunakan bahasa lisan, sebagian individu memiliki keterbatasan dalam berkomunikasi seperti tunarungu dan tunawicara. Mereka yang tidak menggunakan bahasa lisan biasanya

menggunakan bahasa isyarat sebagai metode utama untuk berinteraksi dengan sesama. intraksi yang dilakukan oleh disabilitas dalam berkomunikasi akan melalui gerakan tubuh seperti tangan, bibir, dan wajah [2].

Bahasa isyarat menjadi salah satu metode sebagai jembatan untuk intraksi individu yang memiliki gangguan komunikasi dengan masyarakat umum. Setiap negara memiliki variasi bahasa isyarat masing-masing [3]. Indonesia memiliki dua sistem bahasa isyarat yaitu Bahasa Isyarat Indonesia (SIBI) dan Bahasa isyarat Indonesia (BISINDO). Tercatat menurut data penyandang tunarunggu di indonesia memiliki lebih dari 223 ribu orang penyandang tuli dan 73,5 ribu penyandang tuli dan bisu. Penyandang bisu dan tuli di indonesia memiliki kendala dalam mendapatkan informasi di tempat umum [4], masyarakat umum di indonesia masih belum mengenal mengenai interaksi menggunakan bahasa isyarat kurangnya edukasi serta sumber pengetahuan menjadikan hal tersebut kendala bagi masyarakat penyandang bisu dan tuli.

Masalah tersebut dapat diselesaikan dengan bantuan kecerdasan buatan (AI). Dengan adanya kecerdasan buatan, pembuatan sistem yang dapat dilatih untuk mengenali bahasa isyarat menjadi lebih cepat dan efisien, sehingga masyarakat umum tidak perlu repot-repot menghafal bahasa isyarat. kecerdasan buata (AI) memiliki beberapa bidang ilmu, salah satunya adalah computer vision. Computer vision menjadi solusi bagi sistem untuk memprediksi sebuah input visual dan mengenali objek tersebut. Algoritma YOLO merupakan salah satu algoritma yang dapat mendeteksi objek secara real-time.

2. TINJAUAN PUSTAKA

2.1 Bahasa Isyarat

Bahasa Isyarat sebuah metode untuk berintraksi tanpa menggunakan suara, seperti ucapan manusia atau tukisan [5]. Metode ini biasa digunakan untuk berintraksi oleh tunarunggu dan tunawicara.

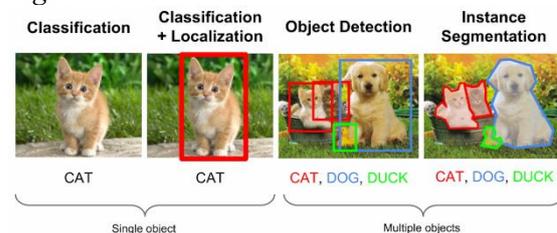
2.2 Computer Vision

Computer Vision adalah bagian dari machine learning yaitu berfungsi untuk membaca objek serta pengenalan pola hasilnya akan memberikan sebuah pemahaman terhadap citra [6]. Tujuan computer vision yaitu untuk

mengembangkan model serta mengestрак informasi lebih dalam terhadap gambar.

2.3 Object Detection

Object detection memiliki peran penting dalam computer vision yaitu bertujuan untuk mengidentifikasi serta mengklasifikasikan gambar dalam satu frame seperti video atau foto dengan memberikan penanda persegi terhadap objek yang dideteksi [7]. Object Detection memiliki dua cara yaitu single object melakukan klasifikasi terlebih dahulu dan memberikan lokasi penanda object, kedua multiple object akan melakukan object detection dilanjut dengan penandaan instance segmentation.



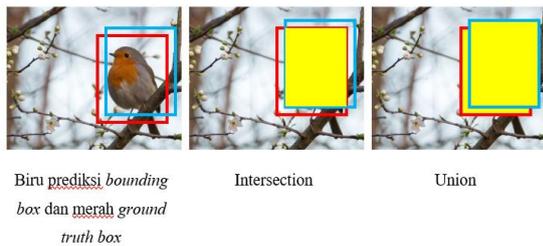
Gambar 1 Deteksi objek

2.4 You Only Look Once (YOLO)

YOLO merupakan salah satu metode object detection yang memiliki pendekatan ini merevolusi deteksi objek dengan memandangnya sebagai masalah regresi [8]. Hal ini memungkinkan YOLO untuk membagi bounding box secara spasial dan memperkirakan probabilitas kelas yang sesuai dengan bounding box tersebut. YOLO bekerja dengan membagi citra pada gambar menjadi beberapa sel grid berukuran s X s. Dari setiap sel grid pada gambar bertanggung jawab untuk memprediksi bounding box serta confidence scores pada setiap kotak pembatas perhitungan persamaan dapat dilihat menggunakan rumus berikut:

$$Confidence = p_r(object) \times IOU_{pred}^{truth} \quad (1)$$

Selain menghitung menggunakan persamaan nilai confidence yolo sendiri dapat dihitung menggunakan Intersection Over Union (IoU). Intersection Over Union (IoU) akan memprediksi dengan cara membandingkan nilai labelling pada gambar secara manual (Ground Truth Box) dengan luas area dari prediksi label prediksi (bounding box) dapat dilihat seperti ilustrasi berikut:



Gambar 2 contoh *intersection over union* Mengukur baik buruknya deteksi objek dapat dilihat untuk ilustrasi berikut



Gambar 3 Ilustrasi Skor IoU

2.5 Convolutional Neural Network(CNN)

YOLO memiliki jaringan arsitektur dari CNN, Convolutional Neural Network (CNN) merupakan jaringan neural yang paling populer dalam mengenal dan menganalisis gambar [9]. Pada konsep YOLO digunakan sebagai ekstraksi fitur penting dari gambar secara otomatis. Berikut penjelasan layer pada CNN:

A. Lapisan Konvolusi

Lapisan Konvolusi adalah komponen dalam jaringan syaraf CNN yang berfungsi melakukan operasi konvolusi pada *matriks* gambar objek dengan menggunakan *matriks* filter sebagai *input*.

B. Lapisan Pooling

Lapisan Pooling adalah elemen dalam jaringan syaraf konvolusional (CNN) yang berfungsi untuk mengurangi dimensi peta fitur (*feature map*).

C. Layer Sepenuhnya Terhubung

Lapisan Sepenuhnya Terhubung terdiri dari lapisan input, lapisan tersembunyi, dan lapisan output, yang semuanya dikenal sebagai lapisan Fully-Connected.

D. Lucky ReLU

Lucky ReLU ini berfungsi untuk memperbaiki fungsi layer dari aktivitas ReLU.

E. Fungsi Aktivitas Linear

Fungsi Aktivitas Linear ini berfungsi untuk mengenalkan unsur *non-linear* pada keluaran *layer neuron*

2.6 Mean Average Precision(mAP)

mAP merupakan nilai rata-rata yang digunakan dalam evaluasi yang biasa digunakan oleh Faster R-CNN, Mobile Net, SSD, dan YOLO. mAP itu sendiri adalah nilai

diambil dari *Average Precision* pada setiap kelas dalam deteksi [10].

2.7 Precision

Precision merupakan evaluasi dengan menghitung nilai rata-rata dari model. setiap model dilatihkan untuk mengukur presisi, yang kemudian akan menghitung nilai rata-rata dalam setiap kelas. Rumus persamaan sebagai berikut:

$$Precision = \frac{True\ Positive}{True\ Positive + false\ Positive} \quad (2)$$

2.8 Recall

Recall adalah evaluasi model yang digunakan untuk mengukur jumlah prediksi positif yang teridentifikasi dalam dataset. Rumus persamaannya adalah sebagai berikut:

$$Recall = \frac{True\ Positive}{True\ Positive + false\ Positive} \quad (3)$$

3. METODE PENELITIAN

CRIPS-DM (Cross Industry Standard Process for Data Mining) adalah sebuah kerangka atau struktur yang biasanya ada dalam proyek kecerdasan buatan. CRIPS-DM memiliki tujuan untuk memberikan blue print berupa tahapan dalam memproses sebuah data [11], adapun tahapan pada CRISP-DM sebagai berikut:

3.1 Business Understanding

Fase awal pada tahap business understanding berfokus pada pemahaman tujuan serta persyaratan proyek dari perspektif bisnis, lalu pengetahuan tersebut menjadi definisi masalah data dan rancangan rencana untuk mencapai tujuan

3.2 Data Understanding

fase kedua adalah pengumpulan data awal, pemahaman pada data, mengidentifikasi kualitas data, dan Menemukan wawasan awal dari data atau mendeteksi subset yang menarik untuk membentuk hipotesis mengenai informasi tersembunyi.

3.3 Data Pre-processing

Fase ketiga yaitu persiapan data mencakup kegiatan menyusun dataset dari data mentah menjadi data yang dapat digunakan pada tahap modelling. fase ini dilaklkan sebuah proses perubahan data seperti augmentasi pada dataset bertujuan agar data memiliki data yang bersih.

3.4 Modelling

Fase keempat adalah pemodelan, di mana percobaan pemodelan dilakukan untuk

Tabel 1 *Augmentasi*

NO	Augemntasi	Hasil
1.	Resize	
2.	Grayscale	
3.	Contrast	
4.	Rotate	

4. Modelling

Fase ini melibatkan pemodelan menggunakan YOLOv8 dengan membandingkan berbagai variasinya, yaitu YOLOv8n, YOLOv8s, YOLOv8m, dan YOLOv8l. Sebelum melakukan perbandingan variasi YOLOv8, pengaturan parameter akan dilakukan dengan membandingkan beberapa *optimizer* terlebih dahulu.

Optimizer yang dibandingkan adalah Adam, SGD, dan RMSProp. Ketiga optimizer tersebut memiliki tingkat akurasi yang tidak terlalu berbeda. Namun, optimizer terbaik

adalah Adam dengan akurasi 93,6%. Detail hasilnya adalah sebagai berikut:

Tabel 2 Perbandingan 3 *Optimizer*

Optimizer	Precision	Recal	mAP50	mAP50-95
Adam	0.993	1	0.995	0.936
SGD	0.992	1	0.995	0.931
RMSprop	0.986	1	0.995	0.919

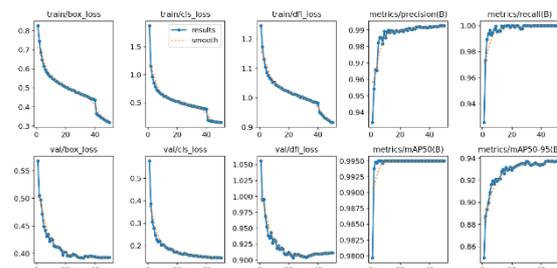
Setelah membandingkan tiga optimizer, Adam dipilih sebagai optimizer untuk proses pelatihan YOLOv8 berikutnya. Perbandingan variasi YOLOv8 menunjukkan akurasi yang tidak jauh berbeda, dengan hasil perbandingan dari keempatnya sebagai berikut:

Tabel 3 Perbandingan variasi YOLOv8

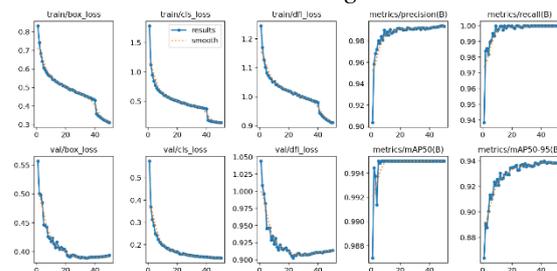
Varian YOLO v8	Precision	Recal	mAP50	mAP50-95
YOLOv8n	0.992	1	0.995	0.931
YOLOv8s	0.992	1	0.995	0.936
YOLOv8m	0.992	1	0.995	0.938
YOLOv8l	0.994	1	0.995	0.938

5. Evaluation

Fase ini merupakan tahap evaluasi model yang sudah dilakukan, evaluasi dilakukan dengan melihat beberapa *matric diagram* seperti Nilai *mean Average Precision*, *Precision*, *Recall*, serta *accuracy*. Hasil dari YOLOv8 dapat dilihat dari table 3, hasil tersebut menunjukan YOLOv8m dan YOLOv8l memiliki *mean Average Precision* (mAP) yang sama namun perbandingan tersebut dapat dilihat lebih lanjut dengan membandingkan *matric diagram* lainnya seperti *matirc diagram* berikut:



Gambar 6 Hasil *Matric diagram* YOLOv8m



Gambar 7 Hasil *Matric diagram* YOLOv8l

Hasil pada *matriks diagram* menunjukkan bahwa YOLOv8m memiliki performa yang lebih baik dibandingkan YOLOv8l, yang menunjukkan tanda-tanda *overfitting*. Oleh karena itu, YOLOv8m dengan parameter digunakan *imgsize* sebesar 600x600 piksel, optimizer Adam dengan *learning rate* 0.001, dan jumlah pelatihan sebanyak 50 *epoch*. dipilih sebagai model untuk permasalahan deteksi bahasa isyarat, dengan rincian akurasi untuk setiap kelas sebagai berikut:

Tabel 4 hasil model YOLOv8m

NO	Nama Kelas	mAP50-95
1.	Adik	0.995
2.	Apa	0.918
3.	Ayah	0.995
4.	Baik	0.973
5.	Bau	0.915
6.	Berapa	0.995
7.	Berhenti	0.934
8.	Buruk	0.981
9.	Cantik	0.924
10.	Cinta	0.97
11.	Datang	0.88
12.	Diam	0.982
13.	Dimana	0.893
14.	Ganteng	0.974
15.	Hai	0.986
16.	Ibu	0.907
17.	Janji	0.965
18.	Kacamata	0.941
19.	Kaget	0.903
20.	Kakak	0.962
21.	Kakek	0.961
22.	Kamu	0.871
23.	Kapan	0.873
24.	Kemana	0.897
25.	Lagi	0.995
26.	Lari	0.978
27.	Maaf	0.979
28.	Makan	0.934
29.	Marah	0.836
30.	Mendengar	0.938
31.	Minum	0.964
32.	Mobil	0.909
33.	Mohon	0.974
34.	Motor	0.941
35.	Ngantuk	0.974
36.	Pergi	0.995
37.	Pintar	0.98
38.	Rumah	0.946
39.	Sama-sama	0.889
40.	Saya	0.907
41.	Sedih	0.868
42.	Sedikit	0.915

NO	Nama Kelas	mAP50-95
43.	Selamat pagi	0.944
44.	Selamat tinggal	0.974
45.	Senang	0.937
46.	Takut	0.92
47.	Telepone	0.984
48.	Terimakasih	0.909
49.	Tidur	0.919
50.	Tolong	0.879

6. Deployment

Fase ini adalah proses pengujian model untuk mengetahui apakah model berfungsi sesuai dengan yang diharapkan.

5. KESIMPULAN

- Membandingkan ketiga *optimizer* dengan menggunakan parameter 50 *epoch*, *batch size* sebesar 32, dengan *leaning rate* 0,001. Menghasilkan nilai akurasi sebesar 93,6% untuk Adam, 93,1% untuk Stochastic Gradient Descent (SGD), dan 91,9% untuk RMSProp. Perbandingan dari ketiganya memiliki hasil yang tidak berbeda jauh, namun, *optimizer Adam* menunjukkan kinerja paling optimal dengan akurasi 93,6%.
- Membandingkan 4 variasi YOLOv8 dengan menggunakan *optimizer Adam* dan *learning rate* sebesar 0.001 menghasilkan akurasi sebesar 93,1% untuk YOLOv8n, 93,6% untuk YOLOv8s, serta 93,8% untuk YOLOv8m dan YOLOv8l. Meskipun YOLOv8m dan YOLOv8l memiliki nilai akurasi yang sama, YOLOv8m menunjukkan hasil yang lebih baik berdasarkan *matriks diagram*, menjadikannya pilihan yang paling optimal di antara keempat variasi tersebut.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Allah Subhanahu Wa Ta'ala atas karunia-Nya yang telah melimpahkan ilmu serta kelancaran dalam menyelesaikan penelitian ini. Terima kasih juga disampaikan kepada orang tua dan semua pihak yang telah memberikan dukungan dan kontribusi dalam pengerjaan penelitian ini.

DAFTAR PUSTAKA

- [1] M. L. Hakim, S. Devi, and A. Suprayit, "Pengaruh Model Pembelajaran Demonstrasi Terhadap Peningkatan Kemampuan Hasil Belajar Bahasa Indonesia Tentang Teks Deskripsi," *J. Ilm. Wahana Pendidik.*, vol. 9, no. 6, pp. 325–332, 2023, [Online]. Available: <https://doi.org/10.5281/zenodo.7783615>
- [2] J. S. Saputro, F. K. Anggarani, and A. Anggrellangi, "Desain Aplikasi 'SAPA' Berbasis Computer-Mediated Communication (CMC) untuk Efektivitas Komunikasi Mahasiswa Tunarungu," *J. Basicedu*, vol. 6, no. 1, pp. 386–393, 2021, doi: 10.31004/basicedu.v6i1.1931.
- [3] S. I. Tenrisara, "Meneliti Bahasa Isyarat Dalam Perspektif Variasi Bahasa," *Kongr. Bhs. Indones.*, vol. 1, no. 2, pp. 1–14, 2018, [Online]. Available: http://kbi.kemdikbud.go.id/kbi_back/file/dokumen_makalah/dokumen_makalah_1540468871.pdf
- [4] D. Arlinta, "Menjembatani Keterbatasan Komunikasi Lewat Aplikasi - Kompas.id," *Kompas.id*, 2020. <https://www.kompas.id/baca/ilmu-pengetahuan-teknologi/2020/09/28/menjembatani-keterbatasan-komunikasi-lewat-aplikasi/> (accessed Aug. 22, 2023).
- [5] Rima Dias Ramadhani, A. Nur Aziz Thohari, C. Kartiko, A. Junaidi, T. Ginanjar Laksana, and N. Alim Setya Nugraha, "Optimasi Akurasi Metode Convolutional Neural Network untuk Identifikasi Jenis Sampah," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 5, no. 2, pp. 312–318, 2021, doi: 10.29207/resti.v5i2.2754.
- [6] K. Telaumbanua, S. Sudarto, F. Butar-Butar, and P. S. Bilqis, "Identifikasi Sampah Berdasarkan Tekstur Dengan Metode GLCM dan GLRLM Menggunakan Improved KNN," *Explorer (Hayward)*, vol. 1, no. 2, pp. 45–52, 2021, doi: 10.47065/explorer.v1i2.94.
- [7] Z. Q. Zhao, P. Zheng, S. T. Xu, and X. Wu, "Object Detection with Deep Learning: A Review," *IEEE Trans. Neural Networks Learn. Syst.*, vol. 30, no. 11, pp. 3212–3232, 2019, doi: 10.1109/TNNLS.2018.2876865.
- [8] M. J. Shaifee, B. Chywl, F. Li, and A. Wong, "Fast YOLO: A Fast You Only Look Once System for Real-time Embedded Object Detection in Video," *J. Comput. Vis. Imaging Syst.*, vol. 3, no. 1, 2017, doi: 10.15353/vsnl.v3i1.171.
- [9] R. R. Karim, "Implementasi Klasifikasi Senjata Tradisional Jawa Barat Menggunakan Convolutional Neural Network (Cnn) Dengan Metode Transfer Learning," *J. Inform. dan Tek. Elektro Terap.*, vol. 12, no. 2, 2024, doi: 10.23960/jitet.v12i2.4166.
- [10] S. Yohanandan, "mAP (mean Average Precision) might confuse you! | by Shiviy Yohanandan | Towards Data Science," May 09, 2020. <https://towardsdatascience.com/map-mean-average-precision-might-confuse-you-5956f1bfa9e2> (accessed Feb. 21, 2024).
- [11] S. Navisa, H. Luqman, and N. Aulia, "Komparasi Algoritma Klasifikasi Genre Musik pada Spotify Menggunakan CRISP-DM," *J. Sist. Cerdas*, vol. 4, no. 2, pp. 114–125, 2021, doi: 10.37396/jsc.v4i2.162.
- [12] P. Jiang, D. Ergu, F. Liu, Y. Cai, and B. Ma, "A Review of Yolo Algorithm Developments," *Procedia Comput. Sci.*, vol. 199, pp. 1066–1073, 2021, doi: 10.1016/j.procs.2022.01.135.