

Deteksi Bahasa Isyarat Indonesia (BISINDO) Pada Video dengan YOLOv7

¹Nurul Renaningtias, ²Ferzha Putra Utama, ³Azzahrah Nur Awaliah Sobri

^{1,2,3}Universitas Bengkulu, Indonesia

nurulrenaningtias@unib.ac.id; fputama@unib.ac.id; azzahrahawaliah@gmail.com;

Article Info

Article history:

Received, 2024-09-13

Revised, 2024-10-04

Accepted, 2024-11-20

Kata Kunci:

Bahasa isyarat
BISINDO
YOLOv7
deteksi

Keywords:

Sign Language
BISINDO
YOLOv7
Detection

ABSTRAK

Bahasa isyarat merupakan metode berkomunikasi yang tidak menggunakan suara, namun menggunakan gerakan fisik seperti tangan, tubuh, dan bibir. Salah satu bahasa isyarat yang banyak digunakan penyandang tuli untuk berkomunikasi adalah Bahasa Isyarat Indonesia (BISINDO). Pada penelitian ini bertujuan menerapkan model deteksi objek deteksi dan klasifikasi citra gestur abjad BISINDO dengan menggunakan algoritma *You Only Look Once* (YOLO) versi 7. Penelitian ini menggunakan data citra video yang terdiri dari 26 abjad alfabet. Pada penelitian ini dilakukan 3 percobaan dengan nilai parameter yang berbeda. Evaluasi dilakukan menggunakan matriks *mean Average Precision (mAP)*, *precision*, *recall*, dan *F1-Score*. Berdasarkan percobaan yang dilakukan nilai akurasi terbaik didapatkan pada percobaan 1 dengan nilai parameter *epoch* = 100, *batch size* = 64, *learning rate* = 0.001, *weight decay* = 0.0001, dan *momentum* = 0.9 menghasilkan nilai *mAP@IoU 0.5* sebesar 0.995, *recall* 1.00, *precision* 1.00, *F1-Score* 1.00. Meskipun demikian, ditemukan bahwa pada penerapan model kedalam model *real-time*, hasil deteksi tidak sebaik hasil yang didapatkan pada proses pelatihan.

ABSTRACT

Sign language is a method of communication that does not use sound but uses physical movements such as hands, body, and lips. One of the sign languages that many deaf people use to communicate is Indonesian Sign Language (BISINDO). This study aims to implement an object detection and image classification model for BISINDO alphabet gestures using the You Only Look Once (YOLO) version 7 algorithm. This research uses video image data consisting of 26 alphabet letters. In this study, three experiments were conducted with different parameter values. Evaluation was carried out using the metrics of mean Average Precision (mAP), precision, recall, and F1-Score. Based on the experiments conducted, the best accuracy was obtained in experiment 1 with parameter values of epoch = 100, batch size = 64, learning rate = 0.001, weight decay = 0.0001, and momentum = 0.9, resulting in mAP@IoU 0.5 value of 0.995, recall 1.00, precision 1.00, F1-Score 1.00. However, it was found that in the application of the model to real-time scenarios, the detection results were not as good as the results obtained during the training process.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-nc-nd/4.0/) license.



Penulis Korespondensi:

Azzahrah Nur Awaliah Sobri,
Program Studi Informatika,
Universitas Bengkulu,
Email: azzahrahawaliah@gmail.com

1. PENDAHULUAN

Indonesia menjadi salah satu negara dengan jumlah penduduk terbanyak di antara negara-negara Asia lainnya. Dalam kategori disabilitas khususnya Tuna Rungu, terdapat sekitar 16,8% atau setara dengan 35 juta jiwa penduduk yang menyandang Tuna Rungu. Hal ini menjadikan Indonesia berada di urutan ke empat setelah Srilanka, Myanmar, dan India. Diperkirakan setiap tahunnya terdapat lebih dari 5 ribu kelahiran bayi penyandang disabilitas dengan kondisi tuli [1].

Tunarungu-wicara merupakan kondisi yang diakibatkan oleh berkurangnya pendengaran dan atau kesulitan dalam berbicara yang dialami oleh seseorang [2]. Penyandang tunarungu cenderung mempunyai alat bantu pendengaran atau menggunakan gerakan tangan atau tubuh dalam melakukan komunikasi dengan sekitar. Gerakan tersebut dikenal sebagai bahasa isyarat. Bahasa isyarat merupakan metode komunikasi yang tidak menggunakan suara, tetapi menggunakan gerakan fisik seperti tangan, tubuh, dan bibir untuk menyampaikan sebuah informasi.

Indonesia sendiri memiliki 2 bahasa isyarat yang digunakan, yaitu Bahasa Isyarat Indonesia (BISINDO) dan Sistem Isyarat Bahasa Indonesia (SIBI). BISINDO merupakan bahasa isyarat yang banyak digunakan penyandang tuli untuk berkomunikasi sehari – hari dengan menggunakan gerakan kedua tangan dan juga ekspresi wajah. Gerkatina merupakan Gerakan untuk Kesejahteraan Tunarungu Indonesia (GERKATINA) yang merupakan organisasi yang melakukan gerakan untuk mengupayakan kesejahteraan masyarakat dengan kondisi tunarungu yang ada di Indonesia. Salah satu cara yang dapat dilakukan adalah dengan pemanfaatan *artificial intelligence* [3]

You Only Look Once (YOLO) merupakan salah satu metode deteksi objek yang paling populer digunakan. Kerangka kerja metode ini didasarkan pada *Convolutional Neural Network* (CNN) yang telah dimodifikasi untuk secara bersamaan dapat memprediksi kotak pembatas (*bounding box*) untuk deteksi objek dan probabilitas kelas untuk klasifikasi objek. Algoritma YOLO membanggakan deteksi kecepatan tinggi sambil mempertahankan presisi dan akurasi. Algoritma YOLO juga menunjukkan hasil yang luar biasa untuk aplikasi *real-time* karena latensinya yang rendah [4] dan mampu mendapatkan deteksi waktu video [5]. YOLOv7 mampu mengungguli pendeteksi objek dengan kecepatan dan akurasi kisaran dari 5 FPS hingga 160 FPS [6][7]

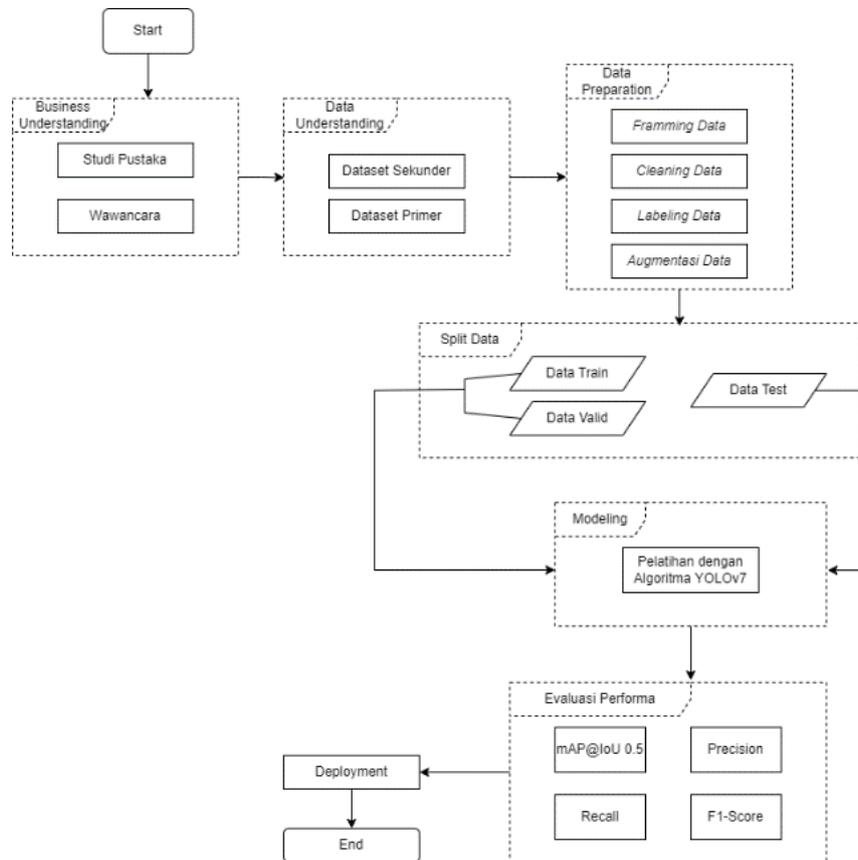
Menurut R Hartanto (2014) dalam Darmatasia [8], permasalahan yang sering terjadi ketika berkomunikasi menggunakan bahasa isyarat adalah ketika orang biasa melakukan komunikasi dengan penyandang tunarungu, orang biasa seringkali kesulitan dalam memahami bahasa isyarat yang digunakan oleh penyandang tunarungu. Beberapa penelitian mengenai pengenalan bahasa isyarat untuk penyandang tunarungu sudah banyak yang dilakukan [9][10][11]. Putri dkk [12] dengan penelitiannya yang berjudul “Pendeteksian Bahasa Isyarat Indonesia Secara *Real-Time* menggunakan *Long Short Term Memory* (LSTM)” pada video rekaman 30 abjad isyarat BISINDO memberikan hasil akurasi sebesar 92% dalam mendeteksi *real-time* menggunakan *bidirectional layer* LSTM, *epoch* 1000, *hidden layer* 64, *batch size* 32 untuk model 10 kelas abjad bahasa isyarat. Talukder dan Jahara [13] dengan penelitiannya yang berjudul “*Real-time Bangla Sign Language Detection Sentence and Speech Generation*” menggunakan model objek deteksi *You Only Look Once* versi 4 (YOLOv4) membangun sistem yang dapat mengenali bahasa isyarat Bangla (BdSL) dari urutan gambar atau *video streaming*. Hasil penelitian ini didapatkan akurasi deteksi objek keseluruhan sebesar 97,95%.

Penelitian yang dilakukan oleh Mulyana dkk [14] dengan judul “Deteksi Bahasa Isyarat Dalam Pengenalan Huruf Hijayah dengan metode YOLOv5” menggunakan metode *You Only Look Once* (YOLO) *Version 5* untuk deteksi bahasa isyarat dalam pengenalan huruf hijayah. Pada penelitian ini digunakan sebanyak 1014 gambar huruf hijayah dan hasil menunjukkan model terbukti secara konsisten dapat mengenali objek dengan nilai tingkat akurasi yang cukup tinggi sebesar 95%. Selain itu Iriani Sapitri dkk [15] dalam penelitiannya yang berjudul “*Deep Learning-Based Real-time Detection for Cardiac Object With Fetal Ultrasound Video*” mengembangkan model deteksi struktur jantung janin dengan memanfaatkan algoritma YOLOv7 dengan menggunakan 40 video ekokardiografi. Penelitian ini mendapatkan hasil mAP sebesar 92% untuk empat objek substruktur jantung dan 82% untuk sembilan objek substruktur jantung. Pan dkk [16] pada penelitiannya yang berjudul “*Research on License Plate Detection and Recognition System based on YOLOv7 and LPRNet*” mengembangkan sebuah sistem deteksi plat kendaraan menggunakan algoritma YOLOv7. Plat nomor yang telah dikenali akan di masukkan ke model STN dan LPRNet untuk mewujudkan pengenalan end-to-end. Penelitian ini mampu mencapai akurasi rata-rata pada lingkungan yang kompleks sebesar 96,1%, *recall* 98,33%, dan *presisi* 99,55%.

Berdasarkan beberapa penelitian terkait sebelumnya, pada penelitian ini dilakukan deteksi gestur abjad Bahasa Isyarat Indonesia (BISINDO) dengan menggunakan metode *You Only Look Once* (YOLO) versi 7. Penelitian ini menggunakan data awal dalam bentuk video dengan format MP4. Pemilihan data dalam bentuk video dikarenakan visualisasi pergerakan dari objek pada video lebih realistis dibandingkan dengan objek pada bentuk gambar. Penggunaan YOLOv7 dipilih karena mampu mendeteksi objek dengan akurasi yang lebih baik dibanding versi sebelumnya.

2. METODE PENELITIAN

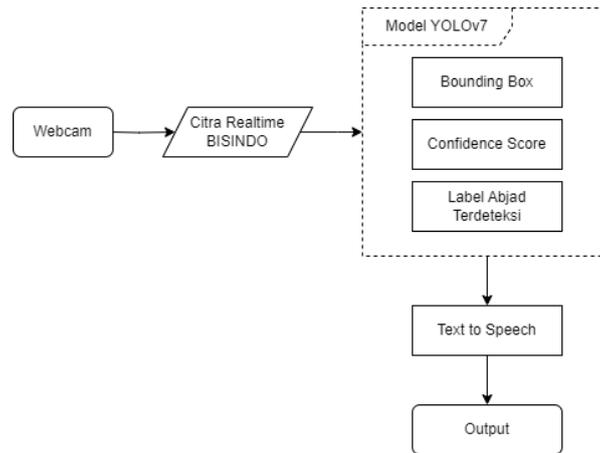
Penelitian sistem deteksi bahasa isyarat dengan menggunakan YOLOv7 dilakukan secara bertahap untuk meminimalisir kesalahan dalam membangun model deteksi bahasa isyarat, alur penelitian yang ditunjukkan pada Gambar 1 berikut ini.



Gambar 1. Diagram Alur Penelitian

Pada tahap *business understanding* dilakukan studi pustaka dan wawancara terkait penggunaan bahasa isyarat di masyarakat. Studi pustaka dilakukan dengan mencari referensi, seperti jurnal terkait bahasa isyarat dan deteksi objek. Sedangkan wawancara dilakukan untuk mengetahui penggunaan bahasa isyarat di masyarakat. Penelitian ini menggunakan dua sumber data, yaitu data primer dan data sekunder. Pada penelitian ini, data sekunder digunakan sebagai data awal. Data sekunder yang digunakan berasal dari dataset *opensource* yang diperoleh dari laman *Kaggle.com*, yaitu dataset *BISINDO Indonesia Sign Language MP4*. Sedangkan data primer yang digunakan pada penelitian ini diambil oleh peneliti dengan melakukan perekaman video gestur abjad BISINDO. Data-data tersebut disatukan menjadi dataset final yang terbagi menjadi data *train*, data validasi, dan data *test* dengan rasio 70:20:10. Setelah *dataset* siap selanjutnya dilakukan beberapa tahap *data preparation* seperti *framing data*, *cleaning data*, *labelling data*, *augmentasi data*, dan *splitting data*. Pada *data preparation* menggunakan bantuan *tools* dalam proses *framing data* dan *labelling data* yang bernama *Roboflow*.

Tahap selanjutnya adalah *modeling*. Pada tahap ini data dilatih menggunakan algoritma YOLOv7. Proses pelatihan dilakukan dengan 3 kali percobaan yang menggunakan parameter berbeda. Parameter yang digunakan dalam penelitian ini, yaitu *epoch*, *batch size*, *learning rate*, *weight decay*, *momentum*, dan *patience*. Selanjutnya dilakukan evaluasi model, tahap ini dilakukan untuk pengujian model yang telah dilatih dengan menggunakan data uji yang telah disiapkan. Pada tahap evaluasi ini menggunakan nilai *mean Average Precision (mAP)*, *Precision*, *Recall*, *f1-Score*. Pada penelitian ini dilakukan proses *deployment* dengan memanfaatkan *framework flask*. *Framework* ini menyediakan *libraries* dan kumpulan kode yang dapat digunakan untuk membangun halaman *website* [17]. Alur kerja sistem setelah dilakukan *deployment* ditunjukkan pada Gambar 2 berikut ini.



Gambar 2. Diagram Alur Sistem

Gambar 2 menunjukkan alur kerja sistem pendeteksian bahasa isyarat secara *realtime*. Pendeteksian dimulai memilih menu deteksi dengan menyalakan *webcam*, kemudian dilanjutkan dengan peragaan gestur bahasa isyarat secara *realtime*. Dengan memanfaatkan model yang telah dilatih, sistem akan menampilkan *bounding box* pada objek terdeteksi dengan menampilkan *confidence score* dan label abjad yang terdeteksi. Label abjad yang terdeteksi akan di *generate* menjadi teks abjad yang tampil pada laman. Selanjutnya abjad yang terdeteksi akan dikonversikan menjadi suara.

3. HASIL DAN ANALISIS

Terdapat 26 kelas yang digunakan dalam penelitian ini, yaitu kelas A-Z. Rincian data awal dapat dilihat pada Tabel 1 berikut.

Tabel 1. Jumlah Data Awal

No	Jenis Data	Jumlah
1.	Data Sekunder	78
2.	Data Primer	26
Total		104

Tahap pertama yang dilakukan untuk proses deteksi adalah dengan melakukan konversi citra dari bentuk video ke dalam bentuk citra. Proses konversi dilakukan dengan memecah video menjadi *frame-frame* citra gambar menggunakan *library OpenCV*. Proses pemecahan citra ini menyebabkan peningkatan jumlah dataset. Namun hasil data yang telah di *frame* belum bisa langsung digunakan untuk proses selanjutnya dikarenakan pada data tersebut masih terdapat citra yang tidak berisikan objek deteksinya. Oleh sebab itu, dilakukan proses pembersihan data terlebih dahulu. Dataset yang telah dibersihkan selanjutnya akan diberikan label sesuai dengan kelas objek yang terdeteksi.

Proses pelabelan data menggunakan bantuan *tools Roboflow* dengan format anotasi YOLOv7. Proses pelabelan difokuskan pada objek tangan yang menggambarkan gestur abjad BISINDO. Hasil keluaran dari proses pelabelan ini yaitu *bounding box* dan file dengan format *.txt* yang berisikan koordinat dari objek terdeteksi. Proses augmentasi data dilakukan dengan bantuan *tools Roboflow*. Teknik augmentasi data yang digunakan dapat dilihat pada Tabel 2 berikut.

Tabel 2. Augmentasi Data

Teknik	Nilai
<i>Flip</i>	<i>Horizontal</i>
<i>Rotasi</i>	-5° dan +5°
<i>Shear</i>	±5° <i>horizontal</i> dan ±5° <i>vertical</i>

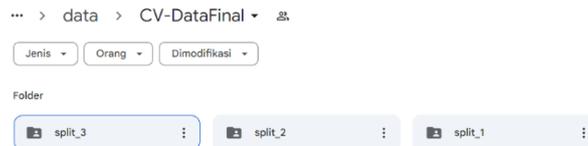
Langkah selanjutnya dari *preparation data* adalah melakukan *splitting* data atau pembagian data. Tabel 3 merupakan tabel total final jumlah data yang digunakan pada penelitian ini dengan memanfaatkan teknik *Cross Validation*.

Tabel 3. Hasil pembagian data menggunakan Cross Validation

Jenis Data	Jumlah
Train	12061

Validasi	6031
Test	2209
Total	20301

Penelitian ini menerapkan metode *Cross Validation*. *Cross Validation* merupakan teknik yang digunakan pada *machine learning* untuk menilai kinerja model dengan membagi data menjadi subset *train* dan validasi. Teknik ini membantu mencegah kebocoran data dan *overfitting* dengan menggunakan data terpisah. Gambar 3 merupakan hasil pembagian data menjadi 3 *fold* terpisah. Folder tersebut diberi nama *split_1*, *split_2*, dan *split_3*.



Gambar. 3. Folder hasil K-Fold

Pada penelitian ini proses *modelling* dilakukan dengan memanfaatkan algoritma YOLOv7 yang diujikan dengan beberapa pengaturan untuk mendapatkan hasil yang baik. Adapun beberapa pengaturan *hyperparameter* yang dilakukan dapat dilihat pada Tabel 4 berikut ini.

Tabel 4. *Hyperparameter*

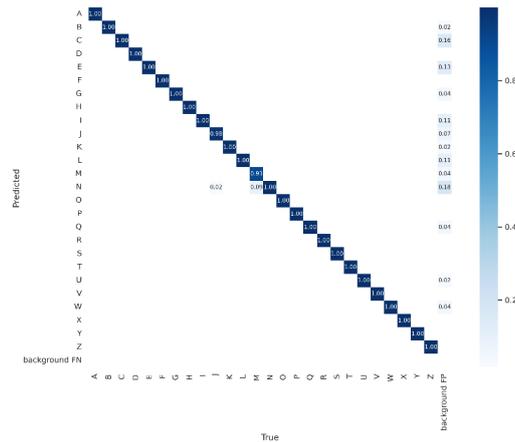
<i>Hyperparameter</i>							
	<i>Fold</i>	<i>Epoch</i>	<i>Batch</i>	<i>Patience</i>	<i>Learning Rate</i>	<i>Weight decay</i>	<i>momentum</i>
P1	3	100	64	-	0.001	0.0001	0.9
P2	3	100	64	50	0.01	0.005	0.92
P3	3	100	64	50	0.1	0.0005	0.8

Tahap selanjutnya dilakukan *evaluation* untuk pengujian performa model dari beberapa percobaan yang dilakukan. Tahap ini memberikan nilai untuk *mean Average Precision (mAP)*, *Precision*, *Recall*, *f1-Score*. Berikut merupakan hasil percobaan menggunakan dataset yang telah melalui tahap *Cross Validation (CV)*.

Tabel 5. Tabel evaluasi performa Yolov7 dengan CV

	Hasil			
	<i>mAP@IoU 0.5</i>	<i>Recall</i>	<i>Precision</i>	<i>F1-Score</i>
P1	0.995	1.00	1.00	1.00
P2	0.995	1.00	1.00	0.99
P3	0.994	1.00	1.00	0.99

Percobaan pertama (P1) menggunakan nilai *epoch = 100*, *batch = 64*, *learning rate = 0.001*, *weight decay = 0.0001*, dan *momentum = 0.9*. Hasil menunjukkan bahwa nilai benar pada setiap kelas telah menunjukkan nilai 1.00 yang dapat diartikan bahwa objek deteksi terdeteksi benar sesuai dengan kelasnya. Namun pada percobaan pertama masih terdapat kelas yang belum sepenuhnya terdeteksi benar. Seperti pada kelas J, terdapat 0.02 citra yang terdeteksi sebagai kelas N. Begitu pula pada kelas M, terdapat 0.09 citra yang terdeteksi sebagai kelas N. Hal ini dapat disebabkan oleh kemiripan gestur seperti pada huruf M dan huruf N sehingga terdapat beberapa citra huruf M yang terdeteksi sebagai huruf N. Hasil *confusion matrix* P1 dengan *cross validation* ditunjukkan pada Gambar 4 berikut ini.



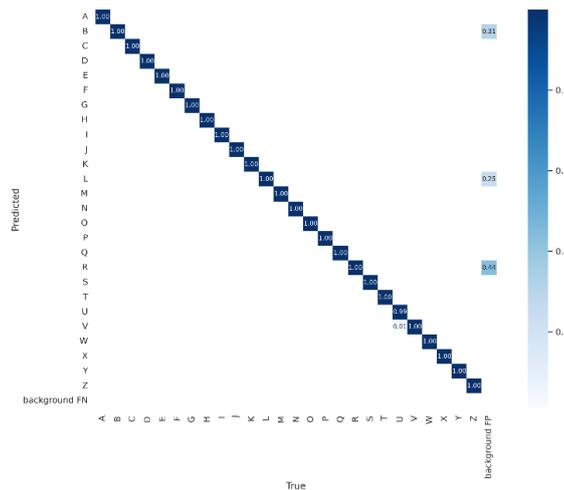
Gambar 4. Confusion Matrix P1 dengan CV

Selanjutnya dilakukan percobaan lain dengan tanpa melalui tahap *cross valudation*. Berikut merupakan hasil percobaan menggunakan dataset tanpa melalui tahap *Cross Validation*.

Tabel 6. Tabel evaluasi performa Yolov7 tanpa CV

	Hasil			
	mAP@IoU 0.5	Recall	Precision	F1-Score
P1	0.995	1.00	1.00	1.00
P2	0.992	1.00	1.00	0.99
P3	0.986	1.00	1.00	0.97

Percobaan pertama (P1) menggunakan nilai *epoch = 100*, *batch = 64*, *learning rate = 0.001*, *weight decay = 0.0001*, dan *momentum = 0.9*. Hasil menunjukkan bahwa nilai benar pada setiap kelas telah menunjukkan nilai 1.00 yang dapat diartikan bahwa objek deteksi terdeteksi benar sesuai dengan kelasnya. Namun pada percobaan pertama masih terdapat kelas yang belum sepenuhnya terdeteksi benar. Seperti pada kelas U, terdapat 0.01 citra yang terdeteksi sebagai kelas V. Begitu pula pada kelas B, L, R, terdeteksi di *background FP* atau dapat dikatakan sebagai model yang keliru mendeteksi adanya bahasa isyarat. Hasil *confusion matrix* P1 dengan tanpa *cross validation* ditunjukkan pada Gambar 5 berikut ini.



Gambar 5. Confusion Matrix P1 tanpa CV

Setelah didapatkan model dengan performa terbaik, selanjutnya memasuki tahap *deployment*. Tahap *deployment* merupakan proses pengimplementasian model yang telah dibangun. Proses deteksi yang dilakukan secara *realtime*. Gambar 6 merupakan tampilan hasil dari deteksi yang dilakukan. Pada gambar terlihat bahwa Bisilator mampu memprediksi huruf dengan ketepatan 0.978 yang menandakan bahwa model berhasil mengenali huruf dengan tepat dan akurat. Pada kolom deteksi ditampilkan hasil deteksi yang merupakan hasil konversi dari *output* yang terdeteksi dan waktu yang dibutuhkan model untuk memprediksi huruf tersebut. Selain itu *output* juga dikonversikan ke dalam bentuk suara.

Gambar 6. Tampilan Deteksi *Realtime*

4. KESIMPULAN

Berdasarkan hasil dan pembahasan yang telah dilakukan, penelitian ini menunjukkan bahwa algoritma *YOLOv7* mampu mendeteksi objek tangan dengan gestur abjad BISINDO dengan performa yang tinggi. Penelitian ini dilakukan dengan 3 kali percobaan dengan nilai parameter yang berbeda dengan membandingkan nilai parameter *learning rate*, *weight decay*, dan *momentum*. Hasil akurasi terbaik yang diperoleh dalam penelitian ini didapatkan pada percobaan 1 dengan $mAP@IoU 0.5$ sebesar 0.995, presisi sebesar 1.00, *recall* sebesar 1.00, dan *f1-score* sebesar 1.00.

REFERENSI

- [1] M. B. S. Bakti and Y. M. Pranoto, "Pengenalan Angka Sistem Isyarat Bahasa Indonesia Dengan Menggunakan Metode Convolutional Neural Network," *Seminar Nasional Inovasi Teknologi*, pp. 11–16, 2019.
- [2] G. Q. O. Pratamasunu, Nur Fajri, F., & Kurnia Sari, P., "Deteksi Tangan Otomatis Pada Video Percakapan Bahasa Isyarat Indonesia Menggunakan Metode Deep Gated Recurrent Unit (GRU)," *Jurnal Komputer Terapan*, vol. 8, no. 1, 2022, pp. 186–193. DOI: <https://doi.org/10.35143/jkt.v8i1.4901>
- [3] Fadillah, R. Z., Irawan, A., & Susanty, M., "Model Penerjemah Bahasa Isyarat Indonesia (BISINDO) Menggunakan Pendekatan Transfer Learning," *PETIR: Jurnal Pengkajian dan Penerapan Teknik Informatika*, vol. 15, no.1, pp. 1–9, 2022.
- [4] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2016-Decem, pp. 779–788, 2016, DOI: 10.1109/CVPR.2016.91.
- [5] Jiang, P., Ergu, D., Liu, F., Cai, Y., & Ma, B., "A Review of Yolo Algorithm Developments". *Procedia Computer Science*, 199, pp. 1066–1073, 2021 DOI: <https://doi.org/10.1016/j.procs.2022.01.135>
- [6] Wang, C.-Y., Bochkovskiy, A., & Liao, H.-Y. M., "*YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors*", pp. 1–15, 2022, DOI: <http://arxiv.org/abs/2207.02696>
- [7] Olorunshola, O. E., Irhebhude, M. E., & Eseoghene, A., "A Comparative Study of *YOLOv5* and *YOLOv7* Object Detection Algorithms", vol. 2, no. 1, 2023.
- [8] D. Darmatasia, "Pengenalan Sistem Isyarat Bahasa Indonesia (IBI) Menggunakan Gradient-Convolutional Neural Network," *Jurnal INSTEK (Informatika Sains dan Teknologi)*, vol. 6, no. 1, p. 56, 2021, DOI: 10.24252/instekv6i1.18637.
- [9] Aziz, A. N. "Image Recognition Alfabet Bahasa Isyarat Indonesia (BISINDO) Menggunakan Metode Convolutional Neural Network", *Skripsi*, Yogyakarta, 2021.
- [10] Abdurrahman, M. F., "Pengenalan Bahasa Isyarat Indonesia Dengan Algoritma Convolutional Neural Network (CNN) Menggunakan Kinect 2.0", *Skripsi*, Yogyakarta, 2019.
- [11] Bastomi, A., & Fadlilah, U., "Alat Penerjemah Bahasa Isyarat Penyandang Tunarungu berbasis Raspberry Pi Kamera". *Simposium Nasional RAPI XIX Tahun 2020 FT UMS, December 2*(ISSN 1412-9612), pp. 15–22, 2020.
- [12] H. Moetia Putri, W. Fuadi, and Fadlisyah, "Pendeteksian Bahasa Isyarat Indonesia Secara Real-Time Menggunakan Long Short-Term Memory (LSTM)," *Tts 4.0*, vol. 3, no. 1, pp. 14–25, 2022.

- [13] D. Talukder and F. Jahara, "Real-Time Bangla Sign Language Detection with Sentence and Speech Generation," *ICCIT 2020 - 23rd International Conference on Computer and Information Technology, Proceedings*, pp. 19–21, 2020, doi: 10.1109/ICCIT51783.2020.9392693.
- [14] D. I. Mulyana, M. F. Lazuardi, and M. B. Yel, "Deteksi Bahasa Isyarat Dalam Pengenalan Huruf Hijaiyah Dengan Metode YOLOV5," vol. 4, pp. 145–151, 2022.
- [15] A. Iriani Sapitri *et al.*, "Deep learning-based real time detection for cardiac objects with fetal ultrasound video," *Inform Med Unlocked*, vol. 36, no. December 2022, p. 101150, 2023, DOI: 10.1016/j.imu.2022.101150.
- [16] S. Pan, J. Liu, and D. Chen, "Research on License Plate Detection and Recognition System based on YOLOv7 and LPRNet", vol. 4, no. 2, pp. 62–68, 2022.
- [17] Mufid, M. R., Basofi Arif, Al Rasyid, M. U. H., Rochimansyah, I. F., & Rokhim, A., *Design an MVC Model using Python for Flask Framework Development*, 2019, DOI: <https://doi.org/10.1109/ELECSYM.2019.8901656>.