

Sistem Pengenalan Objek Untuk Penyandang Tunanetra Menggunakan YOLOv8

^{1*} Abrham Galatiano Ohny ²Ardi Sanjaya, ³Risky Aswi Ramadhani

¹ Teknik Informatika, Universitas Nusantara PGRI Kediri

E-mail: ¹abrham.ohny@gmail.com, ²dersky@gmail.com

³riskyaswiramadhani@gmail.com

Penulis Korespondens : Abrham Galatiano Ohny

Abstrak— Tunanetra adalah istilah umum yang digunakan untuk menyebut seseorang yang mengalami gangguan atau hambatan pada indra penglihatan. Oleh karena itu, untuk mengenali keberadaan suatu objek, penyandang tunanetra akan lebih bergantung pada indra lain seperti pendengaran, penciuman, dan peraba. Tujuan penelitian ini yaitu, membuat sebuah sistem pengenalan objek menggunakan YOLOv8 yang digunakan untuk penyandang tunanetra mengenali objek di sekitarnya. *Input* dari sistem berupa *frame* yang ditangkap melalui kamera lalu diproses sehingga dapat mengenali objek. Untuk *output* dari sistem berupa suara dengan kata-kata objek apa yang dikenal. Sistem mampu mendeteksi secara *real-time* dengan resolusi 640x640 pada nilai rata-rata *fps* 30. Dari hasil uji coba yang dilakukan peneliti menggunakan 35 gambar dan 29 objek sebagai validasi, model deteksi objek memiliki nilai rata-rata akurasi diatas 85%.

Kata Kunci—tunanetra,yolov8,opencv

Abstract— Tunanetra is a common term used to describe a condition in which an individual experiences impairment or obstacles in their visual sense. People with this condition rely more on senses other than vision, such as hearing, smell, and touch, to perceive their surroundings. The goal of this research is to create an object recognition system using YOLOv8, specifically designed for visually impaired individuals to recognize objects around them. The system takes input from camera-captured frames, processes them, and identifies objects. The output of the system is provided through sound, verbally describing the recognized objects. The system achieves real-time detection with a resolution of 640x640 at an average frame rate of 30 fps. In validation tests using 35 images and 29 objects, the object detection model achieved an average accuracy above 85%.

Keywords—tunanetra;yolov8;opencv

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



I. PENDAHULUAN

Tunanetra adalah istilah umum yang digunakan untuk kondisi seseorang yang mengalami gangguan atau hambatan dalam indra penglihatannya [1], maka orang yang memiliki kondisi tersebut untuk mengetahui keberadaan objek, akan lebih bergantung kepada indra selain penglihatan yaitu pendengaran, pembau, peraba, dan lain-lain. Namun dengan menggunakan indra selain penglihatan, sulit untuk mengenali objek secara.

Dari berbagai perkembangan teknologi saat ini menunjang sebuah mesin untuk dapat belajar seperti manusia dalam mengenali sebuah objek. Mesin sendiri membutuhkan kecerdasan buatan untuk dapat mengenali objek citra digital [2]. Deteksi objek merupakan kecerdasan buatan yang dibuat layaknya ketika manusia mengenali sebuah benda [3]. Dalam beberapa tahun terakhir pendeteksian objek berdampak besar pada bidang visi komputer [4].

Algoritma yang sering digunakan dalam bidang deteksi objek adalah *YOLO (You Only Look Once)* [2]. *YOLO* pertama diperkenalkan pada tahun 2015 di sebuah paper yang dibuat oleh *Joseph Redmon*, hingga tahun 2023 telah dirilisnya *YOLOv8* [5]. Kelebihan utama *YOLOv8* adalah peningkatan performa dan akurasi dalam deteksi objek [6].

OpenCV merupakan *library computer vision* bersifat open-source yang populer, fungsi utamanya adalah untuk memanipulasi, memproses dan menganalisis gambar [7]. Terdapat lebih 2,500 algoritma yang bisa digunakan, selain itu OpenCV mendukung berbagai bahasa pemrograman [8].

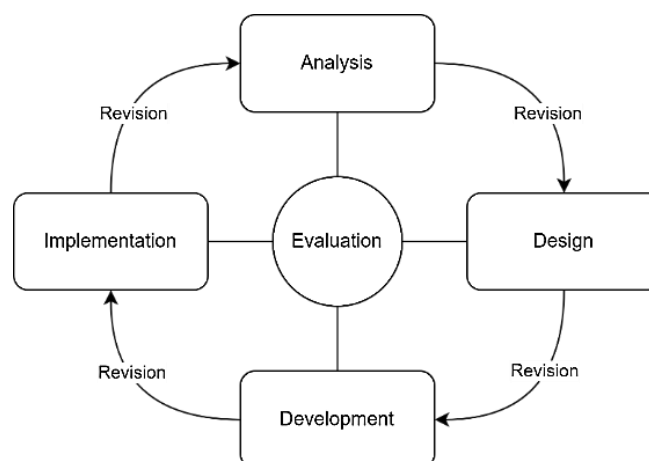
Text-to-Speech, merupakan kemampuan sistem untuk menghasilkan suara yang menyamai suara manusia dari teks tertulis [9]. *GTTS (Google Text-to-Speech)* adalah *library* pada bahasa pemrograman *python* yang berfungsi sama seperti *TTS*. Untuk proses pembuatannya membutuhkan koneksi internet agar terhubung ke *Google Translates text-to-speech API* [10].

Pada penelitian terdahulu yang telah menghasilkan aplikasi pendeteksi objek berbasis *Android* yang dirancang untuk membantu penyandang tunanetra [11]. Dengan memanfaatkan *library Text to speech*, *OpenCV* dan *Tensorflow Lite*, sebagai pendukung aplikasi untuk mendeteksi objek dan memberitahu pengguna. Disebutkan bahwa aplikasi telah berfungsi sesuai rancangan serta menjalani pengujian menggunakan 5 objek yang digunakan sehari-hari.

Dari uraian diatas ditawarkan sebuah solusi, yaitu membuat sebuah sistem pengenalan objek menggunakan *YOLOv8*. *Input* dari sistem berupa *frame* yang ditangkap melalui kamera lalu diproses sehingga dapat mengenali objek. Untuk *output* dari sistem berupa suara yang dengan kata-kata objek apa yang dikenal dan berapa banyak objek tersebut.

II. METODE

Metode yang digunakan dalam penelitian ini adalah metode *Research and Development (R&D) ADDIE*. Metode penelitian *Research and Development (R&D)* adalah metode penelitian yang bersifat analisis kebutuhan, bertujuan untuk menghasilkan dan menguji keefektifan produk tertentu supaya dapat berfungsi di masyarakat [12]. Terdapat banyak model dalam *R&D* salah satunya adalah *ADDIE*, (*Analyze, Design, Development, Implementation, dan Evaluation*) [13]. Berikut siklus penelitian dengan menggunakan model *ADDIE* [14]:



Gambar 1. Siklus model *ADDIE*

2.1 Analisa (*Analysis*)

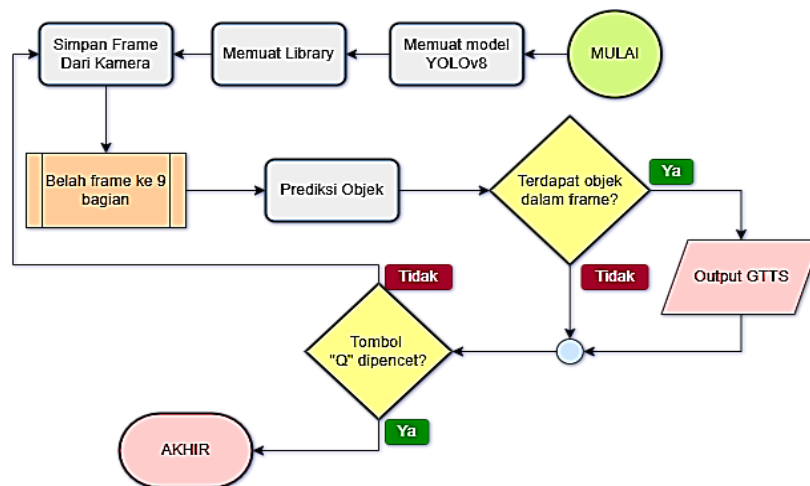
Dari hasil analisa peneliti pada pengembangan sistem deteksi objek terdahulu, peneliti menyimpulkan sebagai berikut:

- 1 Objek yang dapat dikenal sedikit, karena *dataset* yang digunakan bukan *dataset Microsoft COCO (Common Objects in Context)*. Dimana terdapat lebih dari 200 ribu gambar yang terlabel dan bersifat *open-source* [15].
- 2 Pelaporan objek belum menggunakan letak objek terdeteksi, sistem terdahulu masih terfokus pada pendeteksian objek
- 3 Pelaporan objek terdeteksi masih berupa teks atau getaran, sama seperti poin sebelumnya bahwa sistem terdahulu masih terfokus pada deteksi objek. Pada penelitian lain sistem yang dibuat tidak menggunakan modul *speaker* namun menggunakan getaran.

Dengan menggunakan *YOLO* versi 8, *Text-to-Speech*, dan fungsi navigasi yang telah dirancang oleh peneliti, berkemungkinan mengurangi masalah-masalah tersebut.

2.2 Desain (*Design*)

Berikut alur sistem yang digunakan sistem dipresentasikan menggunakan *flowchart*, sebagai penjelasan bagaimana sistem berjalan:



Gambar 2 Desain *Flowchart* Alur Sistem

Dari desain *flowchart* diatas, terdapat poin penting dalam sistem. Penyimpanan *frame* untuk *input model YOLOv8*, setelah itu dibelah ke 9 bagian (kiri atas, depan, kanan bawah, dan seterusnya), dan yang terakhir menggunakan *GTTS* sebagai *output* atau pelaporan. Sistem akan tetap berjalan hingga tombol (q) dipencet

2.3 Pengembangan (*Development*)

Seperti yang tertulis pada analisa dan desain, peneliti mengembangkan sistem agar dapat mengenali objek menggunakan model *YOLOv8*. Dengan menambahkan *GTTS* pada sistem, pelaporan deteksi objek yang dihasilkan berupa suara layaknya manusia. Selain itu peneliti mengembangkan fitur navigasi, dengan memanfaatkan *image manipulation* menggunakan *library OpenCV*. Untuk mengukur performa model, peneliti menggunakan *Confusion Matrix*. Dimana teknik ini mempresentasikan hasil dari identifikasi antara banyaknya data yang terprediksi benar dan salah [16]. Berikut bagaimana hasil dari *confusion matrix* dipresentasikan:

Tabel 1. Struktur tabel *Confusion Matrix*

Aktual	Terprediksi		Jumlah
	Negative	Positive	
	Negative	Positive	Jumlah
	TN	FP	N
	FN	TP	P
	N'	P'	P + N

Dimana:

TN (*True Negative*) : Memprediksi label “*negative*” dengan tepat

FN (*False Negative*) : Memprediksi “*negative*” namun label aktual “*positive*”

FP (*False Positive*) : Memprediksi “*positive*” namun label aktual “*negative*”

TP (*True Positive*) : Memprediksi label “*positive*” dengan tepat

Dengan hasil *Confusion Matrix* diatas dapat menghasilkan tingkat akurasi sistem dalam memprediksi label dengan penjumlahan sebagai berikut:

$$Akurasi = \frac{TP + TN}{P + N} \quad (1)$$

Implementasi sistem pada penelitian ini dilakukan di ruangan kamar, guna mereplika skenario penggunaan dalam sehari-hari. Pengguna diharuskan mendapatkan objek yang disebutkan oleh pihak lain, dengan kondisi pengguna hanya dapat memanfaatkan fungsi dari sistem, peletakan objek secara acak oleh pihak selain pengguna. Keakuratan pengguna direkam sebagai penilaian uji coba. Karena sistem belum dapat mendeteksi halangan seperti tembok. Pihak lain diharuskan memberitahu pengguna jika terdapat halangan di dekat arah jalan pengguna.

III. HASIL DAN PEMBAHASAN

3.1 Hasil Pengembangan

Pada tahap ini, peneliti telah mengimplementasikan sistem seperti skenario yang dirancang pada tahap sebelumnya.



Gambar 3. Pelaksanaan Uji Coba Implementasi Sistem

3.2 Evaluasi

Uji coba performa dilakukan menggunakan 35 *file* gambar yang terdapat beragam objek. *File* gambar ini dikumpulkan dari *dataset microsoft coco* [15] sebagai validasi, dan juga label aktual tertulis sebagai perbandingan. Berikut disajikan evaluasi performa menggunakan confusion matrix:

Tabel 2. Hasil Uji Coba Menggunakan *Confusion Matrix*

Objek	Confusion Matrix								Akurasi
	TN	FN	TP	FP	Total				
					P	P'	N	N'	
Bis	27	0	6	2	6	8	29	27	0.9429
Botol	25	3	7	0	10	7	25	28	0.9143
Buku	30	3	2	0	5	2	30	33	0.9143
Cangkir	26	3	5	1	8	6	27	29	0.8857
Dasi	34	0	1	0	1	1	34	34	1.0000
Gajah	34	0	1	0	1	1	34	34	1.0000
Gelas	34	0	1	0	1	1	34	34	1.0000
Kursi	29	2	3	1	5	4	30	31	0.9143
Laptop	31	1	2	1	3	3	32	32	0.9429
Lemari Es	29	1	5	0	6	5	29	30	0.9714
Mangkuk	31	0	4	0	4	4	31	31	1.0000
Meja Makan	34	0	1	0	1	1	34	34	1.0000
Microwave	32	0	3	0	3	3	32	32	1.0000
Mobil	30	0	5	0	5	5	30	30	1.0000
Mouse	30	2	3	0	5	3	30	32	0.9429
Orang	25	0	9	1	9	10	26	25	0.9714
Oven	33	1	0	1	1	1	34	34	0.9429
Papan Ketik	32	0	2	1	2	3	33	32	0.9714
Pemanggang Roti	34	1	0	0	1	0	34	35	0.9714
Ransel	34	0	1	0	1	1	34	34	1.0000
Remote	33	0	1	1	1	2	34	33	0.9714
Sofa	32	0	3	0	3	3	32	32	1.0000
Tanaman Pot	34	0	1	0	1	1	34	34	1.0000
Telepon Selular	34	0	1	0	1	1	34	34	1.0000
Televisi	30	0	4	1	4	5	31	30	0.9714
Tempat Tidur	30	0	4	1	4	5	31	30	0.9714
Toilet	33	0	2	0	2	2	33	33	1.0000
Truk	32	0	3	0	3	3	32	32	1.0000
Wastafel	31	1	3	0	4	3	31	32	0.9714

IV. KESIMPULAN

Berdasarkan hasil uji coba sistem, dapat disimpulkan bahwa sistem mampu mendeteksi objek dengan nilai keakuratan maksimal sebesar 1.00, minimal 0.88, dan rata-rata 0,97 serta berjalan secara *real-time*. Melalui pengujian fungsi navigasi dan pelaporan menggunakan fitur *Text-to-Speech*, sistem menunjukkan kinerja yang lancar selama implementasi pada 20 kali putaran uji coba. Namun, dari hasil pengujian tersebut, masih ditemukan beberapa kendala, seperti objek yang diambil oleh pengguna tidak sesuai, tidak dikenali oleh sistem, atau sistem memberikan lebih dari satu hasil dalam satu putaran. Dengan demikian, dapat disimpulkan bahwa sistem belum sepenuhnya akurat dan masih memerlukan penyempurnaan lebih lanjut.

Adapun kekurangan pada penelitian ini, yang dipaparkan untuk penelitian lanjutan ataupun pengembangan baru sebagai saran. Karena keterbatasan alat pada penelitian ini, sistem yang dikembangkan hanya dapat bekerja di perangkat *desktop* dan memerlukan koneksi *internet*. Sistem juga hanya mendukung resolusi maksimal 640x640, sehingga jika digunakan sebagai navigasi tidaklah optimal. Terdapat beberapa objek yang memiliki nilai akurasi yang tinggi yaitu 1.00 atau 100% hal ini disebabkan oleh *file* gambar untuk uji coba yang sedikit. Implementasi pada penelitian ini berfokus pada ruangan tertutup atau *indoor*.

DAFTAR PUSTAKA

- [1] M. Siahaan, C. Harsana Jasa, K. Anderson, M. V. Rosiana, S. Lim, dan W. Yudianto, "Penerapan Artificial Intelligence (AI) Terhadap Seorang Penyandang Disabilitas Tunanetra," 2020. doi: <https://doi.org/10.37253/joint.v1i2.4322>.
- [2] A. Setiyadi, E. Utami, dan D. Ariatmanto, "Analisa Kemampuan Algoritma YOLOv8 Dalam Deteksi Objek Manusia Dengan Metode Modifikasi Arsitektur," *Jurnal Sains Komputer & Informatika (J-SAKTI)*, vol. 7, no. 2, hlm. 891–901, 2023, doi: <http://dx.doi.org/10.30645/j-sakti.v7i2.694>.
- [3] P. Ratna Aningtiyas, A. Sumin, dan S. Wirawan, "Pembuatan Aplikasi Deteksi Objek Menggunakan TensorFlow Object Detection API dengan Memanfaatkan SSD MobileNet V2 Sebagai Model Pra - Terlatih," *Jurnal Ilmiah Komputasi*, vol. 19, no. 3, Mar 2020, doi: 10.32409/jikstik.19.3.68.
- [4] Z. Zou, K. Chen, Z. Shi, Y. Guo, dan J. Ye, "Object Detection in 20 Years: A Survey," *Proceedings of the IEEE*, vol. 111, no. 3, hlm. 257–276, 2023, doi: 10.1109/JPROC.2023.3238524.
- [5] M. Hussain, "YOLO-v1 to YOLO-v8, the Rise of YOLO and Its Complementary Nature toward Digital Manufacturing and Industrial Defect Detection," *Machines*, vol. 11, no. 7. Multidisciplinary Digital Publishing Institute (MDPI), 1 Juli 2023. doi: 10.3390/machines11070677.
- [6] Yanto, F. Aziz, dan Irmawati, "YOLO-V8 PENINGKATAN ALGORITMA UNTUK DETEKSI PEMAKAIAN MASKER WAJAH," *Jurnal Mahasiswa Teknik Informatika*, vol. 7, no. 3, 2023, doi: <https://doi.org/10.36040/jati.v7i3.7047>.
- [7] S. Gollapudi, *Learn Computer Vision Using OpenCV*. Apress, 2019. doi: 10.1007/978-1-4842-4261-2.
- [8] B. Santoso dan R. P. Kristianto, "IMPLEMENTASI PENGGUNAAN OPENCV PADA FACE RECOGNITION UNTUK SISTEM PRESENSI PERKULIAHAN

- MAHASISWA,” *SISTEMASI*, vol. 9, no. 2, hlm. 352, Mei 2020, doi: 10.32520/stmsi.v9i2.822.
- [9] N. Kaur dan P. Singh, “Conventional and contemporary approaches used in text to speech synthesis: a review,” *Artif Intell Rev*, vol. 56, no. 7, hlm. 5837–5880, 2023, doi: 10.1007/s10462-022-10315-0.
- [10] S. Kurlekar, O. A. Deshpande, A. V. Kamble, A. A. Omanna, dan D. B. Patil, “Reading Device for Blind People using Python, OCR and GTTS,” *International journal of Science and Engineering Applications*, vol. 9, no. 4, hlm. 049–052, Apr 2020, doi: 10.7753/IJSEA0904.1003.
- [11] T. Pradana dan E. H. Sula, “PENGEMBANGAN APLIKASI PENDETEKSI OBJEK UNTUK TUNANETRA MENGGUNAKAN OPERATOR TEPI CANNY DENGAN LIBRARY OPENCV BERBASIS ANDROID,” *SPIRIT*, vol. 15, no. 1, hlm. 23–27, Mei 2023, doi: 10.53567/spirit.v15i1.285.
- [12] Okpatrioka, “Research And Development (R&D) Penelitian Yang Inovatif Dalam Pendidikan,” *DHARMA ACARIYA NUSANTARA: Jurnal Pendidikan, Bahasa dan Budaya*, vol. 1, no. 1, 2023, doi: <https://doi.org/10.47861/jdan.v1i1.154>.
- [13] E. Widyastuti dan Susiana, “Using the ADDIE model to develop learning material for actuarial mathematics,” *J Phys Conf Ser*, vol. 1188, hlm. 012052, Mar 2019, doi: 10.1088/1742-6596/1188/1/012052.
- [14] B. A. Mukmin dan N. Primasatya, “Pengembangan Multimedia Interaktif Macromedia Flash Berbasis K-13 Sebagai Inovasi Pembelajaran Tematik Untuk Siswa Sekolah Dasar,” *JURNAL PENDIDIKAN DASAR NUSANTARA*, vol. 5, no. 2, hlm. 211–226, Feb 2020, doi: 10.29407/jpdn.v5i2.13854.
- [15] T.-Y. Lin *dkk.*, “Microsoft COCO: Common Objects in Context,” Mei 2014, doi: 10.48550/arXiv.1405.0312.
- [16] L. Syafa’ah, Z. Zulfatman, I. Pakaya, dan M. Lestandy, “Comparison of Machine Learning Classification Methods in Hepatitis C Virus,” *Jurnal Online Informatika*, vol. 6, no. 1, hlm. 73, Jun 2021, doi: 10.15575/join.v6i1.719.