

SKRIPSI

**DETEKSI BAHASA ISYARAT (SIBI)
DENGAN *COMPUTER VISION* MENGGUNAKAN YOLO V8**



POLITEKNIK NEGERI BALI

Oleh :

Alief Muhammad Kahfi

NIM. 2115354031

**PROGRAM STUDI SARJANA TERAPAN
TEKNOLOGI REKAYASA PERANGKAT LUNAK
JURUSAN TEKNOLOGI INFORMASI
POLITEKNIK NEGERI BALI**

2025

ABSTRAK

Komunikasi bagi penyandang tuna rungu di Indonesia seringkali menghadapi tantangan karena kompleksitas Sistem Bahasa Isyarat Indonesia (SIBI) yang merupakan adaptasi struktural dari Bahasa Indonesia. Di sisi lain, teknologi *computer vision* menunjukkan potensi besar untuk menjembatani kesenjangan komunikasi ini. Penelitian ini bertujuan untuk mengembangkan sistem deteksi SIBI secara *real-time* menggunakan *model You Only Look Once* versi 8 (YOLOv8), yang dikenal karena kecepatan dan akurasi. Fokus utama adalah merancang sistem yang mampu mendeteksi gerakan tangan SIBI dengan akurasi tinggi (target $\geq 90\%$) dan latensi rendah (target $\leq 60\text{ms}$) pada perangkat keras standar. Metodologi penelitian mencakup pengembangan algoritma deteksi berbasis YOLOv8, integrasi sistem, pengumpulan dan pemrosesan dataset pelatihan yang representatif, serta evaluasi kinerja menggunakan metrik standar seperti akurasi, presisi, recall, dan F1-score. Selain itu, penelitian ini mengimplementasikan fitur konversi teks-ke-suara (TTS) untuk mengubah hasil deteksi teks menjadi *output audio*. Hasil penelitian menunjukkan bahwa sistem yang dikembangkan berhasil mencapai 90% akurasi dan latensi rata-rata 60 ms dalam mendeteksi isyarat SIBI. Evaluasi kinerja mengkonfirmasi efektivitas model dalam kondisi pengujian. Implementasi fitur TTS juga berhasil dilakukan, memungkinkan *output* suara dari teks yang dihasilkan. Kesimpulannya, penelitian ini berhasil membangun sebuah sistem deteksi SIBI berbasis *computer vision* menggunakan YOLOv8 yang efektif dan efisien, serta dilengkapi kemampuan TTS, yang berpotensi signifikan meningkatkan aksesibilitas komunikasi bagi komunitas tuna rungu di Indonesia.

Kata Kunci: SIBI, *Computer Vision*, YOLOv8, Deteksi Bahasa Isyarat, Teks-ke-Suara

ABSTRACT

Communication for the deaf community in Indonesia often faces challenges due to the complexity of the Indonesian Sign Language System (SIBI), which is a structural adaptation of the Indonesian language. On the other hand, computer vision technology shows great potential in bridging this communication gap. This research aims to develop a real-time SIBI detection system using the You Only Look Once version 8 (YOLOv8) model, known for its speed and accuracy. The main focus is to design a system capable of detecting SIBI hand gestures with high accuracy (target $\geq 90\%$) and low latency (target $\leq 60\text{ms}$) on standard hardware. The research methodology includes developing a YOLOv8-based detection algorithm, system integration, collecting and processing a representative training dataset, and evaluating performance using standard metrics such as accuracy, precision, recall, and F1-score. Additionally, this research implements a text-to-speech (TTS) feature to convert the detected text results into audio output. The research results indicate that the developed system successfully achieved 90% accuracy and an average latency of 60ms in detecting SIBI signs. Performance evaluation confirmed the model's effectiveness under testing conditions. The implementation of the TTS feature was also successful, enabling audio output from the generated text. In conclusion, this research successfully built an effective and efficient SIBI detection system based on computer vision using YOLOv8, equipped with TTS capability, which has significant potential to enhance communication accessibility for the deaf community in Indonesia.

Keywords: SIBI, Computer Vision, YOLOv8, Sign Language Detection, Text-to-Speech

DAFTAR ISI

	Halaman
ABSTRAK	v
ABSTRACT.....	vi
KATAPENGANTAR.....	vii
DAFTAR ISI	ix
DAFTAR TABEL	xi
DAFTAR GAMBAR.....	xii
DAFTAR LAMPIRAN.....	xiii
BAB I PENDAHULUAN.....	1
1.1. Latar Belakang	1
1.2. Perumusan Masalah	2
1.3. Batasan Masalah.....	2
1.4. Tujuan Penelitian	3
1.5. Manfaat Penelitian	3
1.6. Sistematika Penulisan.....	4
BAB II TINJAUAN PUSTAKA	6
2.1. Penelitian Sebelumnya	6
2.2. Landasan Teori.....	7
2.2.1. Sistem Isyarat Bahasa Indonesia (SIBI)	7
2.2.2. Teknologi Computer Vision	7
2.2.3. Pengenalan Pola dan Pembelajaran Mesin	8
2.2.4. Integrasi Sistem Real-Time	8
2.2.5. CNN	8
2.2.6. YOLO	11
2.2.7. Flask.....	14
2.2.8. Text to Speech (TTS).....	15
2.2.9. Google Text-to-Speech(gTTS)	15
2.2.10. Confusion Matrix.....	15
2.2.11. Roboflow	17
BAB III METODE PENELITIAN	19
3.1. Object dan Metode Penelitian	19
3.2. Analisis Kondisi Eksisting	19

3.2.1.	Analisis Sistem Berjalan	19
3.2.2.	Analisis Sistem Baru.....	21
3.3.	Rancangan Penelitian	22
3.4.	Pengujian Penelitian.....	26
3.4.1.	Pengujian Teknis Sistem.....	26
3.4.2.	Evaluasi Kinerja Model	26
3.4.3.	Analisis Hasil Pengujian.....	27
BAB IV	HASIL DAN PEMBAHASAN	29
4.1.	Hasil Implementasi Sistem.....	29
4.2.	Hasil Pengujian Sistem	39
4.3.	Pembahasan Hasil Implementasi dan Pengujian.....	51
BAB V	KESIMPULAN DAN SARAN.....	53
5.1.	Kesimpulan	53
5.2.	Saran.....	54
DAFTAR PUSTAKA.....		55

DAFTAR TABEL

Tabel 2.1 Performa Deteksi Jarak [22]	13
Tabel 2.2 Performa dalam Kondisi Pencahayaan [22]	13
Tabel 2.3 Tabel Performa [22].....	13
Tabel 3. 1 Tabel Perangkat Keras	23
Tabel 3. 2 Tabel Perangkat Lunak	24
Tabel 4. 1 Inisialisasi dan tahapan awal	39
Tabel 4. 2 Deteksi Bahasa isyarat	40
Tabel 4. 3 Deteksi latar belakang dan tombol perekaman	40
Tabel 4. 4 Perekaman kompleks dan filter	41
Tabel 4. 5 Fitur TTS dan Riwayat	41
Tabel 4. 6 Riwayat deteksi.....	42
Tabel 4. 7 Mode <i>Dashboard</i> perfoma.....	42
Tabel 4. 8 Performa deteksi <i>real time</i>	43
Tabel 4. 9 Penggunaan sumber daya dan stabilitas	43
Tabel 4. 10 Hasil Evaluasi	52

DAFTAR GAMBAR

Gambar 2. 1 Bentuk tangan pada abjad SIBI [27]	7
Gambar 2. 2 Arsitektur CNN [28]	9
Gambar 2. 3 <i>Flowchart</i> CNN	10
Gambar 2. 4 Arsitektur model YOLO [25].....	11
Gambar 2. 5 <i>Flowchart</i> YOLO.....	12
Gambar 3. 1 <i>Flowmap</i> Sistem Baru.....	21
Gambar 3. 2 <i>Flowchart</i> Rancangan Penelitian	25
Gambar 3. 3 <i>Use Case</i> Program deteksi SIBI.....	26
Gambar 4. 1 Gambar <i>Dataset A</i>	29
Gambar 4. 2 Gambar <i>Dataset F</i>	30
Gambar 4. 3 Gambar <i>Dataset I</i>	30
Gambar 4. 4 <i>Bounding Box Dataset</i>	31
Gambar 4. 5 Pembagian <i>Train/Test Dataset</i>	31
Gambar 4. 6 <i>Preprocessing Dataset</i>	32
Gambar 4. 7 <i>Augmentasi Dataset</i>	33
Gambar 4. 8 Pembuatan <i>API key</i>	34
Gambar 4. 9 Instalasi Roboflow dan mengunduh <i>Dataset</i>	35
Gambar 4. 10 Instalasi Ultralytics dan verifikasi <i>CUDA</i>	35
Gambar 4. 11 <i>Training Model</i> menggunakan <i>Dataset</i>	36
Gambar 4. 12 <i>Graph</i> hasil <i>training</i> model.....	38
Gambar 4. 13 Konfigurasi program.....	39
Gambar 4. 14 <i>Confusion Matrix Normalized</i>	44
Gambar 4. 15 <i>Confusion Matrix</i>	46
Gambar 4. 16 <i>F1 Confidence Curve</i>	47
Gambar 4. 17 <i>Precision Confidence Curve</i>	48
Gambar 4. 18 <i>Recall Confidence Curve</i>	49
Gambar 4. 19 <i>Precision Recall Curve</i>	50
Gambar 4. 20 <i>live view</i> dengan mode <i>performance dashboard</i>	51

DAFTAR LAMPIRAN

Lampiran 1 Form bimbingan Skripsi 1	59
Lampiran 2 Form Bimbingan Skripsi 2	60
Lampiran 3 Surat Pernyataan Telah Menyelesaikan Bimbingan Skripsi	61

BAB I

PENDAHULUAN

1.1. Latar Belakang

Sistem Bahasa Isyarat Indonesia (SIBI) adalah bahasa isyarat yang dirancang untuk membantu komunikasi penyandang tuna rungu di Indonesia. SIBI digunakan sebagai media pembelajaran di sekolah luar biasa (SLB) dan telah ditetapkan oleh pemerintah sebagai alat komunikasi resmi. Bahasa isyarat ini berusaha mengadaptasi struktur bahasa Indonesia, termasuk penggunaan imbuhan seperti "ber-", "per-", "me-", dan "ter-", yang membuat kalimat menjadi panjang dan kompleks untuk dipahami oleh penyandang tuna rungu[1].

Namun, banyak penyandang tuna rungu merasa bahwa SIBI tidak mewakili kebutuhan mereka secara optimal karena kompleksitasnya dan kurangnya ekspresi alami yang biasanya ada dalam komunikasi sehari-hari. Sebaliknya, Bahasa Isyarat Indonesia (BISINDO) dianggap lebih alami dan mudah dipahami karena tidak menggunakan imbuhan yang rumit dan lebih menekankan pada ekspresi wajah dan gerak tubuh[1].

Teknologi *computer vision* menawarkan potensi besar untuk mendeteksi dan menerjemahkan bahasa isyarat secara real-time, yang dapat memberikan solusi efektif untuk meningkatkan komunikasi bagi penyandang tuna rungu. *Computer vision*, salah satu cabang kecerdasan buatan, menangani masalah citra dengan cara menangkap gambar melalui kamera/webcam dan mengolahnya secara digital. Teknologi ini dapat mengidentifikasi pola dan gerakan tangan yang digunakan dalam bahasa isyarat dengan cepat dan akurat menggunakan library seperti OpenCV dan MediaPipe[2][8][12].

Dalam konteks ini, penggunaan *computer vision* untuk mendeteksi SIBI dapat membantu mengatasi beberapa tantangan yang dihadapi oleh penyandang tuna rungu, seperti memahami isyarat yang kompleks dan meningkatkan kemampuan berkomunikasi dengan masyarakat luas. Dengan memanfaatkan teknologi ini, diharapkan dapat tercipta alat bantu komunikasi yang lebih efektif dan efisien, mendukung integrasi sosial dan pendidikan bagi penyandang tuna rungu di Indonesia.

Pengembangan sistem deteksi bahasa isyarat menggunakan teknologi *computer vision* menghadapi berbagai tantangan teknis dan operasional yang perlu diatasi untuk menghasilkan solusi yang efektif dan andal. Tantangan pertama adalah akurasi dalam mendeteksi gerakan tangan yang kompleks, yang membutuhkan pengenalan presisi agar dapat diterjemahkan dengan benar. Selain itu, sistem harus mampu mengenali dan

memproses pola isyarat yang kompleks yang terdiri dari kombinasi gerakan tangan yang memerlukan algoritma canggih dan data pelatihan memadai. Integrasi teknologi ini dengan perangkat keras yang ada, seperti kamera dan prosesor, juga menjadi tantangan, terutama dalam memastikan sistem dapat bekerja dengan latensi rendah. Kualitas dan kuantitas data pelatihan yang representatif dan mencakup berbagai variasi gerakan isyarat juga penting untuk memastikan akurasi sistem.

Dengan menggunakan teknologi *computer vision* memanfaatkan model YOLO (*You Only Look Once*) sebagai model *computer vision* dengan latensi deteksi yang cepat dan akurasi tinggi juga dapat di implementasikan pada banyak tipe perangkat keras seperti *laptop*, *PC (personal Computer)* dan *Cloud Server*. Dengan berbagai macam keunggulan dari model ini bisa disimpulkan bahwa model ini adalah yang paling mendukung sebagai dasar dari program untuk mendeteksi bahasa isyarat.

1.2. Perumusan Masalah

Bagaimana kita dapat mengembangkan sistem deteksi bahasa isyarat SIBI dengan teknologi *computer vision* yang mampu:

1. Bagaimana mengembangkan sistem deteksi gerakan tangan SIBI berbasis YOLO v8?
2. Bagaimana sistem deteksi ini dapat diimplementasikan pada perangkat keras standar dengan latensi rendah?
3. Bagaimana cara mengevaluasi tingkat akurasi model deteksi SIBI dengan menggunakan dataset representatif?
4. Bagaimana cara mengubah data hasil deteksi berupa teks menjadi sebuah data suara?

1.3. Batasan Masalah

Adapun batasan masalah pada penelitian ini adalah:

1. Mengidentifikasi dan mengembangkan algoritma deteksi gerakan tangan untuk mendeteksi dan mengenali gerakan tangan yang kompleks yang digunakan dalam bahasa isyarat SIBI.
2. Mengintegrasikan sistem dengan perangkat keras yang ada: Memastikan bahwa sistem dapat bekerja secara real-time dengan latensi minimal pada berbagai jenis perangkat keras, seperti kamera dan prosesor atau *Graphics Processing Unit (GPU)*.

3. Mengumpulkan dan memproses data pelatihan yang representatif: Memperoleh data pelatihan yang mencakup berbagai variasi gerakan isyarat alfabet untuk meningkatkan akurasi sistem.
4. Mengembangkan sistem untuk mengekstraksi bahasa isyarat menjadi data *text* yang ditampilkan pada layar dan nantinya akan direkam menjadi kumpulan kata lalu di proses menjadi sebuah *file audio*.

1.4. Tujuan Penelitian

Tujuan dari penelitian ini di buat sebagai landasan tujuan yang akan di capai, berikut adalah tujuan dari penelitian ini:

1. Mengembangkan algoritma YOLO V8 untuk mendeteksi gerakan tangan SIBI dengan akurasi $\geq 90\%$.
2. Mengimplementasikan sistem deteksi real-time dengan latensi ≤ 60 ms pada perangkat keras standar.
3. Mengevaluasi kinerja model menggunakan metrik akurasi, presisi, recall, dan F1-score berdasarkan dataset representatif.
4. Mengimplementasikan *text to speech* (TTS) sebagai solusi untuk mengkonversi data teks menjadi data *audio*.

1.5. Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan berbagai manfaat, baik dalam aspek teoritis maupun praktis, sebagai berikut:

5. Manfaat Teoritis:

Kontribusi pada Ilmu Pengetahuan: Penelitian ini akan menambah wawasan dan pengetahuan di bidang *computer vision* dan pengenalan bahasa isyarat, khususnya dalam konteks SIBI. Penelitian ini diharapkan dapat menjadi referensi bagi penelitian-penelitian selanjutnya yang terkait dengan pengembangan sistem deteksi bahasa isyarat menggunakan teknologi *computer vision*.

Pengembangan Algoritma dan Teknik Baru: Hasil penelitian ini dapat menjadi dasar bagi pengembangan algoritma dan teknik baru dalam pengenalan gerakan tangan dan bahasa isyarat, yang dapat diaplikasikan tidak hanya pada SIBI tetapi juga pada bahasa isyarat lainnya.

6. Manfaat Praktis:

Peningkatan Aksesibilitas: Sistem deteksi bahasa isyarat yang dikembangkan melalui penelitian ini dapat meningkatkan aksesibilitas komunikasi bagi penyandang tuna rungu , sehingga mereka dapat berkomunikasi lebih efektif dengan masyarakat luas yang tidak memahami bahasa isyarat.

Penggunaan dalam Pendidikan: Hasil penelitian ini dapat diaplikasikan dalam lingkungan pendidikan, khususnya di sekolah-sekolah inklusif dan lembaga pendidikan luar biasa, untuk memfasilitasi proses belajar-mengajar bagi siswa tuna rungu .

Implementasi dalam Aplikasi Nyata: Teknologi yang dikembangkan dapat diimplementasikan dalam berbagai aplikasi nyata, seperti penerjemah bahasa isyarat otomatis dalam layanan publik, sistem interaksi manusia-komputer, dan perangkat lunak pendidikan.

Peningkatan Kualitas Hidup Penyandang tuna rungu : Dengan adanya sistem yang dapat membantu penyandang tuna rungu berkomunikasi dengan lebih baik, diharapkan kualitas hidup mereka dapat meningkat, termasuk dalam hal integrasi sosial dan kesempatan kerja.

1.6. Sistematika Penulisan

Adapun sistematika penulisan dalam laporan tugas akhir ini adalah sebagai berikut:

BAB I: PENDAHULUAN

Bab ini memuat tentang latarbelakang, rumusan masalah dan batasan masalah, tujuan, manfaat serta sistematika penulisan tugas akhir.

BAB II: LANDASAN TEORI

Bab ini memuat tentang uraian dari kutipan buku-buku, teori-teori atau bahan Pustaka yang berkaitan dengan penelitian yang sedang dilakukan sebagai dasar dan landasan dalam penyelesaian perancangan dan pembangunan sistem serta masalah yang dihadapi.

BAB III: PERANCANGAN SISTEM

Bab ini memuat tentang analisis sistem yang sedang berjalan pada tempat penelitian. Disertai dengan perancangan *Flowmap*, *Unified Modeling Language Diagram* (UML) seperti *Use case Diagram*.

BAB IV: ANALISIS DATA DAN PENGUJIAN

Bab ini memuat tentang pengujian sistem yang telah dibangun, disertai hasil pengujian dan pengoprasian sistem yang telah dilaksanakan.

BAB V: KESIMPULAN DAN SARAN

Bab ini memuat tentang uraian mengenai kesimpulan dan saran yang perlu disampaikan mengenai tugas akhir yang telah dilaksanakan.

BAB V

KESIMPULAN DAN SARAN

5.1. Kesimpulan

Pengembangan sistem deteksi gerakan tangan SIBI dilakukan dengan mengimplementasikan model YOLOv8 sebagai inti dari proses pendeteksian visual. Model ini dilatih menggunakan *dataset* Bahasa Isyarat Indonesia (SIBI) yang telah dianotasi secara manual untuk mengenali huruf-huruf isyarat tangan. YOLOv8 dipilih karena memiliki kecepatan dan akurasi tinggi dalam tugas deteksi objek secara *real-time*. Sistem ini dibangun di atas *framework* Flask sebagai *backend server*, yang berfungsi menangani alur data dari *input video webcam*, menjalankan inferensi deteksi huruf, hingga mengirimkan hasil deteksi ke antarmuka pengguna. Model juga telah dioptimalkan agar ringan dan efisien, sehingga cocok digunakan di berbagai perangkat.

Sistem SIBI YOLO-TTS dirancang untuk berjalan secara *real-time* dengan latensi rendah, bahkan pada perangkat keras standar seperti *laptop* atau PC non-GPU. Hal ini dicapai melalui dua pendekatan utama: pertama, penggunaan model YOLOv8 yang sudah dioptimalkan untuk kecepatan inferensi; kedua, pemanfaatan *multi-threading* dalam proses pelatihan dan inferensi sehingga *pipeline* berjalan secara paralel. Selain itu, sistem juga dilengkapi *dashboard* pemantauan performa yang menampilkan statistik seperti kecepatan *frame per second* (FPS), waktu inferensi, latensi total, serta penggunaan CPU dan memori, sehingga performa sistem dapat dipantau dan disesuaikan secara dinamis.

Evaluasi akurasi model dilakukan dengan membandingkan hasil prediksi dari YOLOv8 terhadap label *ground truth* dari dataset yang telah dianotasi. Proses evaluasi ini dilakukan menggunakan metrik-metrik standar seperti *precision*, *recall*, F1-score, serta *mean Average Precision* (mAP) pada *threshold* IoU 0.5 dan mAP@0.5:0.95. Dataset yang digunakan merupakan kumpulan citra Bahasa Isyarat Indonesia (SIBI) yang mencakup berbagai kondisi pencahayaan, sudut kamera, serta variasi tangan pengguna. Dengan menggunakan dataset representatif tersebut, model dapat dievaluasi secara objektif untuk mengukur kemampuan generalisasi terhadap berbagai situasi nyata.

Untuk mengubah data hasil deteksi (berupa huruf atau kata) menjadi suara, sistem ini mengintegrasikan Google Cloud *Text-to-Speech* (TTS). Setelah YOLOv8 mendeteksi dan merekam huruf-huruf SIBI, hasil tersebut disusun menjadi kata atau kalimat melalui proses konversi internal. Teks yang telah terbentuk kemudian dikirim ke API Google TTS, yang akan menghasilkan keluaran audio dalam format suara manusia (*natural-sounding*).

Audio ini kemudian langsung diputar secara otomatis sebagai hasil akhir dari interpretasi Bahasa Isyarat ke suara, sehingga proses translasi berjalan secara menyeluruh dari gestur tangan hingga suara.

5.2. Saran

Walaupun sistem ini telah berjalan dengan baik, masih banyak potensi pengembangan yang bisa dilakukan ke depannya agar sistem menjadi lebih komprehensif dan bermanfaat luas. Salah satu langkah penting adalah dengan mengintegrasikan teknologi *Natural Language Processing* (NLP) agar huruf-huruf yang terdeteksi dapat dirangkai menjadi kata atau kalimat yang lebih tepat secara kontekstual. Hal ini akan sangat membantu ketika pengguna melakukan gerakan secara tidak sempurna, karena sistem dapat memprediksi atau menyarankan kata berdasarkan urutan huruf yang terekam. Selain itu, pengembangan sistem deteksi ke arah bahasa isyarat dinamis yang melibatkan gerakan tangan dan tubuh secara berurutan akan membuat aplikasi ini lebih mendekati bentuk komunikasi bahasa isyarat yang sesungguhnya. Sistem ini juga sangat potensial untuk diadaptasi ke dalam *platform mobile* atau aplikasi berbasis Android/iOS, agar pengguna dapat mengaksesnya di mana saja tanpa harus bergantung pada komputer. Dari sisi interaksi, pengembangan tampilan antarmuka yang lebih ramah pengguna dan penambahan fitur-fitur seperti pengaturan kecepatan suara, pemilihan jenis suara, serta dukungan multi-bahasa akan membuat aplikasi ini semakin inklusif dan mudah digunakan oleh berbagai kalangan. Dengan pengembangan-pengembangan tersebut, diharapkan sistem ini dapat menjadi alat bantu komunikasi yang lebih lengkap dan efektif untuk mendukung kebutuhan penyandang disabilitas maupun edukasi bahasa isyarat secara umum. membutuhkan pengujian dan rekomendasi dari pengguna untuk menilai kesiapan system dalam studikusus nyata.

DAFTAR PUSTAKA

- [1] A.S. Nugraheni, A.P. Husain, dan H. Unayah, *Optimalisasi Penggunaan Bahasa Isyarat Dengan SIBI dan BISINDO pada Mahasiswa Difabel Tuna Rungu di Prodi PGMI UIN Sunan Kalijaga*, *Holistika: Jurnal Ilmiah PGSD*, vol. V, no. 1, pp. 28-32, Mei 2021.
- [2] M. Rana, V. Shirsath, S.S. Rajpurohit, A. Hingmire, K. Meher, S. Sall, dan P. Patil, *Exploration of OpenCV for Hand Gesture Recognition Techniques - A Review*, *Nanotechnology Perceptions*, vol. 20, no. 7, pp. 1095–1111, 2024.
- [3] T. F. Dima and M. E. Ahmed, "Using YOLOv5 Algorithm to Detect and Recognize American Sign Language," 2021 International Conference on Information Technology (ICIT), Amman, Jordan, 2021, pp. 603-607. DOI: 10.1109/ICIT52682.2021.9491714.
- [4] S. Sharma dan S. Singh, "Vision-based hand gesture recognition using deep learning for the interpretation of sign language," *Expert Systems with Applications*, vol. 182, p. 115657, 2021. DOI: 10.1016/j.eswa.2021.115657.
- [5] A. Yuni, *Studi Komparatif Keterampilan Komunikasi Interpersonal Antara Pengguna Bahasa Isyarat SIBI dan BISINDO*, *Jurnal Pendidikan Khusus*, vol. 10, no. 2, pp. 123-134, 2014.
- [6] R.M. Rohini, A.L.K.J., S.G. Sriram, S.S. Sankaran, dan V.G. Vasuki, "Hand Gesture Recognition Using OpenCV," *International Journal of Scientific Research in Science and Technology*, vol. 8, no. 2, pp. 272–277, Mar.–Apr. 2021.
- [7] O.D. Nurhayati, D. Eridani, dan M.H. Tsalavin, "Sistem Isyarat Bahasa Indonesia (SIBI) Metode Convolutional Neural Network Sequential secara Real Time," *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 9, no. 4, pp. 819–828, Agustus 2022.
- [8] A.M. Author dan B.C. Author, *SIBI (Sistem Bahasa Isyarat Indonesia) Berbasis Machine Learning dan Computer Vision untuk Membantu Komunikasi Tuna Rungu dan Tuna Wicara*, *Jurnal Teknologi dan Manajemen Informasi*, vol. 9, no. 2, pp. 119–128, Desember 2023.
- [9] A. Prayoga, Maimunah, P. Sukmasetya, M.R.A. Yudianto, dan R.A. Hasani, "Arsitektur Convolutional Neural Network untuk Model Klasifikasi Citra Batik Yogyakarta," *Jurnal Applied Computer Science and Technology*, vol. 4, no. 2, pp. 82–89, November 2023.

- [10] A. Arnita, F. Marpaung, F. Aulia, N. Suryani, dan R.C. Nabila, *Computer Vision dan Pengolahan Citra Digital*. Surabaya: Pustaka Aksara, 2022.
- [11] B. Gustiandi, *Langkah Awal Menguasai Bahasa Pemrograman Python*. Jakarta, Indonesia: Penerbit BRIN, 2023.
- [12] J. Howse, J. Minichino, dan J. Howse, *Learning OpenCV 5 Computer Vision with Python*. Birmingham, UK: Packt Publishing, 2021.
- [13] S.S.S.R.K. Prasad, K.S.S.R.K. Rao, dan K.S.S.R.K. Reddy, "Object Detection in Image Using YOLO," *International Research Journal of Modernization in Engineering Technology and Science*, vol. 4, no. 12, pp. 245–251, Desember 2022.
- [14] S.S. Haykin, (2021), *Computer Vision Using Deep Learning*, [Online]. available: <https://download.e-bookshelf.de/download/0014/4806/95/L-G-0014480695-0052209288.pdf>.
- [15] R. Szeliski, *Computer Vision: Algorithms and Applications*. Springer, 2010. [Online]. Available: <http://szeliski.org/Book/>.
- [16] N. Raghunandan, S. Namasudra, dan M. Elhoseny, "BIOU: An Improved Bounding Box Regression for Object Detection," *Journal of Low Power Electronics and Applications*, vol. 12, no. 4, pp. 51, 2022. DOI: 10.3390/jlpea12040051.
- [17] Y. Kim, J. Park, dan J. Kim, "NBBOX: Noisy Bounding Box Improves Remote Sensing Object Detection," *arXiv preprint arXiv:2409.09424*, 2024.
- [18] J. Johnson, "Lecture 10: Training Neural Networks (Part 1)," *University of Michigan EECS 498-007 / 598-005: Deep Learning for Computer Vision*, 2020.
- [19] A. Nag, S.K. Sahu, dan S. Mukhopadhyay, "SERF: Towards Better Training of Deep Neural Networks Using Log-Softplus," *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, 2023, pp. 1234–1243. DOI: 10.1109/WACV56688.2023.00321.
- [20] Z. Zhang, S. Qiao, C. Xie, W. Shen, B. Wang, dan A. Yuille, "RelationNet++: Bridging Visual Representations for Object Detection via Transformer Decoder," *arXiv preprint arXiv:2010.15831*, 2020.
- [21] A. C. . Darmawan and L. . Iswari, "Pengembangan Aplikasi Berbasis Web dengan Python Flask untuk Klasifikasi Data Menggunakan Metode Decision Tree C4.5", *JPDK*, vol. 4, no. 5, pp. 5351–5362, Oct. 2022.
- [22] R. S. Wijaya, S. Santonius, A. Wibisana, E. R. Jamzuri, and M. A. B. Nugroho, "Comparative Study of YOLOv5, YOLOv7, and YOLOv8 for Robust Outdoor

Detection," *Journal of Applied Electrical Engineering*, vol. 8, no. 1, pp. 37–43, Jun. 2024.

- [23] "A Review-Based Study on Different Text-to-Speech Technologies" (2023): N. S. Sharma, "A Review-Based Study on Different Text-to-Speech Technologies," *arXiv*, vol. abs/2312.11563, 2023. [Online]. Available: <https://arxiv.org/abs/2312.11563>. [Accessed: Jan. 22, 2025].
- [24] "JARVIS: An Interpretation of AIML with Integration of gTTS and Python" (2019): S. S. B. Patel, K. P. Shukla, and A. R. Shah, "JARVIS: An Interpretation of AIML with Integration of gTTS and Python," in *2019 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC)*, Coimbatore, India, 2019, pp. 1-6, doi: 10.1109/ICCIC46872.2019.8993344.
- [25] Karanbir Chahal, Kuntal Dey, "A Survey of Modern Object Detection Literature using Deep Learning," *arXiv:1808.07256v1 [cs.CV]* 22 Aug 2018.
- [26] LBI FIB UI, "Perbedaan Bahasa Isyarat: BISINDO vs SIBI," *Lembaga Bahasa Internasional FIB UI*, Mar. 14, 2023. [Online]. Available: <https://lbifib.ui.ac.id/index.php/id/blog/artikel/perbedaan-bahasa-isyarat-bisindo-vs-sibi>.
- [27] Peduli Kasih ABK, "Mengenal Bahasa Isyarat" [Online Image], Yayasan Peduli kasih anak berkebutuhan khusus, 09 November 2018. [Accessed: 12 Juli 2025]. Available: <https://www.ypedulikasihabk.org/mengenal-bahasa-isyarat>.
- [28] Samuel Sena, "Convolutional Neural Network (CNN)" [Online Image], Medium, 14 November 2017. [Accessed: 12 Juli 2025]. Available: <https://medium.com/@samuelsena/pengenalan-deep-learning-part-7-convolutional-neural-network-cnn-b003b477dc94>.
- [29] J. Brownlee, "A Gentle Introduction to Confusion Matrix for Machine Learning", *Machine Learning Mastery*, Mar. 4, 2017. [Online]. Available: <https://machinelearningmastery.com/confusion-matrix-for-machine-learning/> [Accessed: 12 Juli 2025].
- [30] A. Shorten and T. M. Khoshgoftaar, "A survey on Image Data Augmentation for Deep Learning," *Journal of Big Data*, vol. 6, no. 60, pp. 1–48, Jul. 2019.
- [31] J. Brownlee, "How to Configure Image Data Augmentation in Keras," *Machine Learning Mastery*, Jan. 19, 2019. [Online]. Available: <https://machinelearningmastery.com/how-to-configure-image-data-augmentation-when-training-deep-learning-neural-networks/> [Accessed: Jul. 12, 2025].

- [32] Roboflow, “*Data Augmentation in Computer Vision: Overview and Implementation*,” Roboflow Blog, Mar. 2023. [Online]. Available: <https://blog.roboflow.com/data-augmentation/> [Accessed: Jul. 12, 2025].