

ABSTRAK

Pengenalan Bahasa Isyarat Otomatis Menggunakan Algoritma CNN, LSTM, dan N-Gram

Nama : Fardiansyah Futra

NPM : 2155201057

Pembimbing : Muntahanah, S.Kom., M.Kom

Penyandang tunarungu dan tunawicara di Indonesia masih menghadapi hambatan besar dalam berkomunikasi dengan masyarakat umum yang tidak menguasai bahasa isyarat. Penelitian ini bertujuan untuk merancang dan membangun sistem pengenalan bahasa isyarat otomatis berbasis web yang memanfaatkan algoritma Convolutional Neural Network (CNN) melalui model HandPose TensorFlow.js untuk mendeteksi dan mengklasifikasikan gesture tangan secara real-time, serta menghasilkan output suara berbahasa Indonesia sebagai respons komunikasinya.

Sistem dikembangkan sebagai aplikasi web satu halaman menggunakan HTML, CSS, dan JavaScript dengan mengintegrasikan library TensorFlow.js dan model HandPose berbasis CNN untuk mendeteksi 21 titik landmark tangan melalui kamera browser. Klasifikasi gesture dilakukan menggunakan pendekatan rule-based classification berdasarkan pola biner status jari yang dipetakan ke 8 gesture dasar. Output suara dihasilkan menggunakan Web Speech API dengan bahasa Indonesia (id-ID). Metode pengembangan yang digunakan adalah Waterfall.

Pengujian sistem dilakukan menggunakan metode Black Box Testing terhadap 15 skenario uji fungsional, dan seluruh fitur dinyatakan berjalan sesuai yang diharapkan. Sistem berhasil mengenali 8 gesture (Halo, Tidak, Aku, Oke, Dua, Tiga, Empat, Telepon) dengan latensi deteksi rata-rata kurang dari 100ms per frame dan dapat diakses dari browser modern tanpa instalasi tambahan. **Kata Kunci:** Pengukuran luas, lubang jalan, pengolahan citra digital, OpenCV, Python, confusion matrix.

Kata Kunci: CNN, Bahasa Isyarat, TensorFlow.js, HandPose, Web Speech API, Real-time, Aksesibilitas

ABSTRACT

THE AUTOMATIC SIGN LANGUAGE RECOGNITION USING CNN ALGORITHM

NAME : Fardiansyah Futra

STUDENT ID : 2155201057

SUPERVISOR : Muntahanah, S.Kom., M.Kom

People who are deaf or hard of hearing in Indonesia still face significant barriers in communicating with the public who do not know sign language. This study aims to design and develop a web-based automatic sign language recognition system that utilizes a Convolutional Neural Network (CNN) algorithm via the HandPose TensorFlow.js model to detect and classify hand gestures in real time, as well as generate Indonesian voice output as a communicative response. The system was developed as a single-page web application using HTML, CSS, and JavaScript, integrating the TensorFlow.js library and a CNN-based HandPose model to detect 21 hand landmark points via the browser's camera. Gesture classification is performed using a rule-based approach based on binary finger status patterns mapped to 8 basic gestures. Voice output is generated using the Web Speech API in Indonesian (id-ID). The Waterfall methodology was used for development. System testing was conducted using the Black Box Testing method across 15 functional test scenarios, and all features were found to function as expected. The system successfully recognized 8 gestures ("Hello," "No," "Me," "Okay," "Two," "Three," "Four," and "Phone") with an average detection latency of less than 100 ms per frame and is accessible from modern browsers without any additional installation.

Keywords: CNN, Sign language, TensorFlow.js, HandPose, Web Speech API, Real-time, Accessibility

BAB I

PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi kecerdasan buatan (Artificial Intelligence/AI), khususnya algoritma Convolutional Neural Network (CNN), telah membawa revolusi besar dalam bidang computer vision dan pengenalan pola. CNN terbukti sangat efektif dalam mengekstraksi fitur dari data visual secara hierarkis, mulai dari tepi dan tekstur sederhana hingga pola kompleks seperti bentuk tangan dan konfigurasi jari. Kemampuan ini menjadikan CNN sebagai algoritma terdepan dalam sistem pengenalan bahasa isyarat otomatis (Oyedotun & Khashman, 2017).

Selain CNN, perkembangan model berbasis sequence learning seperti Long Short-Term Memory (LSTM) juga berperan penting dalam meningkatkan akurasi sistem pengenalan, khususnya dalam memahami urutan gerakan atau konteks temporal dari gesture tangan. LSTM mampu mengingat informasi jangka panjang sehingga sangat cocok digunakan untuk memproses rangkaian gerakan bahasa isyarat yang tidak hanya bergantung pada satu frame, tetapi juga pada kontinuitas gerakan dari waktu ke waktu. Di sisi lain, pendekatan berbasis Natural Language Processing seperti N-gram dapat digunakan untuk meningkatkan kualitas hasil interpretasi dengan memprediksi kata atau frasa yang paling mungkin muncul berdasarkan urutan sebelumnya. Kombinasi CNN, LSTM, dan N-gram memberikan pendekatan yang lebih komprehensif dalam membangun sistem pengenalan bahasa isyarat yang tidak hanya akurat secara visual, tetapi juga kontekstual dalam penyusunan bahasa.

Di Indonesia, penyandang disabilitas tunarungu dan tunawicara masih menghadapi hambatan besar dalam berkomunikasi dengan masyarakat luas. Bahasa isyarat yang menjadi media utama komunikasi mereka belum dikuasai oleh sebagian besar masyarakat umum, sehingga interaksi sehari-hari di ruang publik, layanan kesehatan, pendidikan, maupun pekerjaan sering terhambat. Menurut data Badan Pusat Statistik (BPS) tahun 2023, jumlah penyandang disabilitas di

Indonesia mencapai sekitar 22,97 juta jiwa atau 8,5% dari total penduduk, dengan proporsi signifikan berasal dari kelompok tunarungu dan tunawicara.

Kemajuan teknologi web modern, khususnya hadirnya TensorFlow.js sebagai framework machine learning berbasis JavaScript, kini memungkinkan model CNN berjalan langsung di browser pengguna dengan memanfaatkan akselerasi GPU melalui WebGL. Model HandPose dari TensorFlow.js, yang berbasis arsitektur CNN, mampu mendeteksi 21 titik landmark pada tangan secara real-time melalui kamera browser. Hal ini membuka peluang besar untuk menghadirkan sistem pengenalan bahasa isyarat yang tidak memerlukan instalasi perangkat lunak tambahan dan dapat diakses dari berbagai perangkat (Zhang et al., 2020).

Berdasarkan permasalahan tersebut, penelitian ini bertujuan merancang dan membangun sistem pengenalan bahasa isyarat otomatis berbasis web yang memanfaatkan algoritma CNN melalui model HandPose TensorFlow.js untuk mendeteksi gesture tangan, mengintegrasikan LSTM untuk memahami urutan gerakan secara temporal, serta menggunakan pendekatan N-gram untuk menyusun hasil interpretasi menjadi kalimat yang lebih natural. Sistem ini juga dilengkapi dengan output suara berbahasa Indonesia menggunakan Web Speech API. Metode pengembangan yang digunakan adalah Waterfall.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, rumusan masalah dalam penelitian ini adalah:

1. Bagaimana merancang dan membangun sistem pengenalan bahasa isyarat otomatis berbasis web menggunakan kombinasi algoritma CNN, LSTM, dan N-Gram yang dapat berjalan secara *real-time* di *browser*?
2. Bagaimana mengimplementasikan model HandPose berbasis CNN untuk ekstraksi *landmark* tangan dan algoritma LSTM untuk mengklasifikasikan sekuens/urutan gerakan *gesture* tangan?

3. Bagaimana menerapkan pemodelan bahasa N-Gram untuk menyusun urutan kata hasil deteksi *gesture* menjadi susunan kalimat yang lebih terstruktur dan natural?
4. Bagaimana mengintegrasikan Web Speech API untuk menghasilkan *output* suara berbahasa Indonesia secara otomatis berdasarkan kalimat yang telah diproses?

1.3 Batasan Masalah

Batasan masalah pada penelitian ini:

1. Sistem dikembangkan sebagai aplikasi web menggunakan HTML, CSS, JavaScript dengan integrasi pustaka TensorFlow.js memanfaatkan arsitektur CNN (HandPose) untuk deteksi *landmark*, LSTM untuk klasifikasi data sekuensial gerakan, serta model N-Gram untuk pemrosesan teks.
2. *Gesture* yang dikenali dibatasi pada 8 *gesture* dasar: Halo, Tidak, Aku, Oke, Dua, Tiga, Empat, dan Telepon.
3. Model LSTM dibatasi pada pengenalan sekuens temporal *landmark* tangan dengan *window frame* tertentu, dan model N-Gram menggunakan pendekatan *Bi-gram/Tri-gram* untuk prediksi serta penyusunan kalimat sederhana.
4. *Output* suara dihasilkan menggunakan Web Speech API dengan bahasa Indonesia (id-ID).
5. Sistem hanya mengenali satu tangan dalam satu waktu (*single hand*) dan memerlukan kondisi pencahayaan yang memadai.
6. Penelitian ini tidak mencakup pengenalan seluruh alfabet bahasa isyarat (A–Z) maupun kosa kata bahasa isyarat lengkap (*BISINDO/SIBI* secara penuh).

1.4 Manfaat Penelitian

1. Memberikan solusi komunikasi berbasis CNN yang mudah diakses oleh penyandang tunarungu/tunawicara melalui browser tanpa instalasi.

2. Mengurangi hambatan komunikasi antara penyandang disabilitas dengan masyarakat melalui konversi isyarat tangan ke suara otomatis.
3. Menjadi referensi implementasi CNN berbasis TensorFlow.js untuk pengembangan aplikasi aksesibilitas web.
4. Mendukung inklusivitas digital di Indonesia dengan solusi yang tidak memerlukan biaya instalasi.

1.5 Tujuan Penelitian

1. Merancang dan membangun sistem pengenalan bahasa isyarat berbasis CNN (HandPose TF.js) yang berjalan real-time di browser.
2. Mengimplementasikan klasifikasi 8 gesture tangan berdasarkan 21 landmark CNN HandPose.
3. Mengintegrasikan Web Speech API untuk output suara berbahasa Indonesia secara otomatis.
4. Menyediakan antarmuka web yang ramah pengguna dengan fitur riwayat, highlight gesture, dan pengaturan volume.

1.6 Framework Penelitian

Tabel 1.1 Framework Penelitian

FASE	Uraian Kegiatan	Input	Output
1. Requirements Analysis (Analisis Kebutuhan)	Melakukan studi literatur dan observasi kebutuhan pengguna (tunarungu/tunawicara). Mengidentifikasi 8 gesture tangan yang akan dikenali. Menentukan kebutuhan teknis: TensorFlow.js, model	1. Kebutuhan komunikasi tunarungu/tunawicara 2. Referensi library TensorFlow.js & HandPose 3. Studi literatur CNN gesture recognition	Dokumen spesifikasi kebutuhan sistem & daftar 8 gesture beserta label suara

	HandPose/CNN, Web Speech API, WebRTC.		
2. System Design (Perancangan Sistem)	Merancang arsitektur sistem client-side penuh. Membuat DFD Level 0 & 1, flowchart alur deteksi, tabel pemetaan gesture ke pola biner jari, dan desain antarmuka menggunakan Figma.	1. Dokumen kebutuhan sistem 2. Spesifikasi library yang digunakan 3. Alur kerja sistem CNN	1. Rancangan arsitektur sistem 2. DFD Level 0 & Level 1 3. Flowchart deteksi gesture 4. Desain antarmuka (UI) 5. Tabel pemetaan gesture
3. Development / Implementation (Pengembangan)	Mengimplementasikan sistem dalam index.html menggunakan HTML, CSS, JavaScript. Mengintegrasikan model HandPose (CNN-based), logika klasifikasi gesture berbasis landmark jari, dan Web Speech API untuk output suara berbahasa Indonesia.	1. Rancangan UI 2. Model HandPose TF.js 3. Logika klasifikasi CNN 4. Web Speech API (id-ID)	Aplikasi web pengenalan isyarat tangan dengan fitur: kamera, deteksi skeleton 21 landmark, klasifikasi 8 gesture, output suara, highlight gesture,

			riwayat deteksi
4. Testing (Pengujian)	Melakukan pengujian Black Box Testing terhadap 15 skenario uji fungsional mencakup akurasi gesture, output suara, cooldown, highlight, riwayat, dan pengaturan volume. Pengujian dengan variasi kondisi pencahayaan.	1. Aplikasi versi awal 2. Skenario uji 15 test case 3. Kondisi pencahayaan bervariasi	Laporan hasil pengujian Black Box Testing dengan seluruh 15 fitur dinyatakan berhasil
5. Deployment (Penerapan)	Mengunggah aplikasi ke platform hosting HTTPS (GitHub Pages/Netlify) agar dapat diakses secara online. HTTPS diperlukan untuk akses kamera getUserMedia() di browser modern.	1. Aplikasi hasil pengujian 2. Platform hosting + HTTPS 3. Domain publik	Aplikasi web yang dapat diakses online melalui URL publik tanpa instalasi
6. Maintenance (Pemeliharaan)	Melakukan pemeliharaan berkala: memperbaiki bug, memperbarui versi library TensorFlow.js/HandPose, menambah gesture baru berdasarkan masukan pengguna, mengoptimalkan performa CNN.	1. Feedback pengguna 2. Pembaruan library TF.js 3. Data evaluasi penggunaan	Versi aplikasi yang diperbarui, lebih stabil, akurat, dan sesuai kebutuhan terbaru