

KLASIFIKASI SINYAL WICARA UNTUK GERAKAN KURSI RODA MENGUNAKAN METODE JARINGAN saraf TIRUAN BACKPROPAGATION

Arief Wisaksono, Hindarto^{*)}, Ade Efiyanti dan Ahmad Ahfas

Fakultas Sains dan Teknologi, Universitas Muhammadiyah Sidoarjo
Jl. Raya Gelam 250 Candi, Sidoarjo61271, Indonesia

^{*)}E-mail: hindarto@umsida.ac.id

Abstrak

Kursi Roda yang dikendalikan oleh sinyal wicara adalah teknologi bantuan yang inovatif dirancang untuk meningkatkan mobilitas dan kemandirian penyandang disabilitas. Penelitian ini bertujuan untuk mengklasifikasi sinyal wicara yang digunakan untuk Gerakan kursi roda yang dapat dioperasikan menggunakan perintah suara. Ada lima jenis data perintah wicara yang harus dikenali yaitu maju, mundur, kiri, kanan, dan berhenti. Terdapat beberapa ekstraksi fitur yang digunakan, yaitu menggunakan metode FFT dengan mengambil nilai energi rata-rata, dan menggunakan metode DWT dengan mengambil nilai subband energi dan nilai zero crossing threshold. Proses klasifikasi yang digunakan menggunakan metode Jaringan saraf Tiruan Backpropagation. Penelitian yang diujikan menggunakan metode Jaringan saraf Tiruan menghasilkan tingkat akurasi yang lebih baik dengan menggunakan hidden layer 2 dan hidden layer 3. Hasil akurasi yang didapatkan sebesar 100% untuk proses pelatihan dan 100% untuk proses pengujian.

Kata kunci: Sinyal Wicara, Backpropagation, FFT, DWT

Abstract

Speech signal-controlled wheelchairs are an innovative assistive technology designed to increase the mobility and independence of people with disabilities. This research aims to classify the speech signals used for wheelchair movement which can be operated using voice commands. There are five types of speech command data that must be recognized, namely forward, backward, left, right and stop. There are several feature extractions used, namely using the FFT method by taking the average energy value, and using the DWT method by taking the subband energy value and the zero crossing threshold value. The classification process used uses the Backpropagation Neural Network method. The research tested using the Artificial Neural Network method produced a better level of accuracy using hidden layer 2 and hidden layer 3. The accuracy results obtained were 100% for the training process and 100% for the testing process.

Keywords: Speech Signal, Backpropagation, FFT, DWT

1. Pendahuluan

Sinyal wicara adalah istilah yang digunakan untuk merujuk pada sinyal elektronik yang mengandung informasi suara. Dalam konteks komunikasi dan teknologi informasi, sinyal wicara sering kali berkaitan dengan transmisi dan pemrosesan suara manusia. Sinyal ini dapat berasal dari berbagai sumber, seperti mikrofon yang merekam suara manusia, dan kemudian diubah menjadi sinyal elektronik. Sinyal ini kemudian dapat diproses, disimpan, ditransmisikan, atau diubah kembali menjadi suara melalui speaker atau perangkat lainnya. Dalam konteks pengolahan sinyal, sinyal wicara sering kali dianalisis untuk berbagai keperluan, seperti pengenalan suara, sintesis suara, peningkatan kualitas suara, dan lain sebagainya. Teknologi ini penting dalam berbagai aplikasi,

termasuk telekomunikasi, asisten suara digital, alat bantu dengar, dan sistem keamanan [1], [2].

Sinyal wicara adalah bentuk komunikasi yang kompleks dan terdiri dari dua komponen utama: sinyal akustik dan aspek linguistik atau artikulasi [3], [4]. (1) Sinyal Akustik yaitu aspek fisik dari sinyal wicara. Komponen akustik mencakup frekuensi, intensitas, dan timbre suara yang dihasilkan oleh pita suara. Frekuensi berkaitan dengan nada atau tinggi suara, sedangkan intensitas berkaitan dengan kerasnya suara. Timbre adalah kualitas unik suara yang membedakan suara satu orang dari yang lain, bahkan jika mereka berbicara dengan nada dan intensitas yang sama. (2) Komponen Artikulatori yaitu berkaitan dengan cara suara dihasilkan secara mekanis. Komponen artikulatori melibatkan gerakan organ-organ bicara seperti bibir, lidah, rahang, dan palatum (langit-langit mulut). Gerakan-gerakan ini menentukan bagaimana aliran udara

diubah saat melewati traktus vokal untuk membentuk fonem yang berbeda, yang merupakan unit dasar suara dalam bahasa. Kedua komponen ini bekerja bersama untuk menghasilkan sinyal wicara yang khas dan dapat dikenali. Teknologi pengenalan suara modern sering menggunakan kedua aspek ini untuk menginterpretasi dan mengubah suara menjadi teks atau perintah.

Pengenalan sinyal wicara, juga dikenal sebagai pengenalan ucapan atau pengenalan suara, merujuk pada proses mengubah sinyal wicara menjadi teks atau representasi simbolik lainnya. Ini melibatkan penggunaan algoritma dan teknik pemrosesan sinyal digital untuk menganalisis sinyal suara dan mengidentifikasi unit-unit bahasa yang terkandung didalamnya. Pengenalan sinyal wicara adalah teknologi yang memungkinkan komputer atau perangkat untuk mengidentifikasi dan memproses ucapan manusia. Ini digunakan dalam berbagai aplikasi, termasuk Asisten Virtual, Transkripsi Otomatis, Sistem Kontrol Suara untuk Perangkat, Aplikasi Aksesibilitas, Sistem IVR (Interactive Voice Response), Pendidikan dan Pelatihan Bahasa, Pengamanan dan Verifikasi Identitas, Analisis Sentimen [5]–[7]. Penggunaan teknologi ini terus berkembang, membuka banyak kemungkinan baru dalam interaksi manusia dan komputer.

Pengenalan wicara untuk gerakan kursi roda dapat mengacu pada penggunaan sinyal wicara manusia untuk mengontrol gerakan atau navigasi kursi roda secara verbal. konversi perintah suara yang diucapkan oleh pengguna menjadi aksi fisik yang dilakukan oleh kursi roda [8]. Pengenalan wicara untuk gerakan kursi roda melibatkan langkah-langkah, diantaranya Perekaman suara, Pra-pemrosesan, Pengenalan ucapan, Penghubung dengan kursi roda, Eksekusi perintah, Umpan balik. Pengenalan wicara untuk gerakan kursi roda memungkinkan pengguna yang memiliki keterbatasan fisik untuk mengendalikan kursi roda secara verbal, memberikan mereka lebih banyak kemandirian dan aksesibilitas dalam mobilitas mereka sehari-hari [9]. Ini dapat menjadi bagian dari sistem kontrol yang lebih kompleks yang juga memperhitungkan faktor-faktor seperti navigasi, keamanan, dan penghindaran hambatan.

Penelitian tentang klasifikasi sinyal wicara untuk gerakan kursi roda melibatkan pengembangan dan penerapan teknik pengenalan sinyal wicara untuk mengidentifikasi perintah yang diucapkan oleh pengguna kursi roda dan menerjemahkannya menjadi gerakan yang sesuai. Beberapa penelitian yang telah dilakukan [10]. Penelitian ini menggunakan pendekatan model Hidden Markov untuk memperkirakan pengguna ke arah gerakan kursi roda yang diinginkan dengan menafsirkan pandangannya. Pemodelan tatapan yang khas pola menggunakan HMM tatapan orang yang paling mungkin keadaan kursi roda dapat ditentukan dan sesuai Gerakan [11], pada penelitian ini kursi roda dengan tenaga biasa menggunakan sensor untuk melihat sekeliling kursi roda, antarmuka ucapan untuk

menginterpretasikan perintah, perangkat nirkabel untuk penentuan lokasi tingkat ruangan, dan perangkat lunak kontrol motor untuk memengaruhi gerakan kursi roda. Pada penelitian ini terdapat Lima perintah suara (ya, tidak, kiri, kanan, dan berhenti) memandu dan mengendalikan kursi roda melalui Raspberry Pi dan penggerak motor DC [12]. Penelitian dengan mengembangkan kursi roda yang dikendalikan suara dengan menggunakan teknologi murah dan handal [13]. Teknologi ini menggunakan model pengenalan suara bawaan yang digabungkan dengan model CNN dan LSTM untuk melatih dan mengklasifikasikan lima perintah yang diucapkan. Penelitian-penelitian tersebut bertujuan untuk mengembangkan sistem yang lebih canggih dan akurat dalam mengklasifikasikan sinyal wicara pengguna dan mengontrol gerakan kursi roda berdasarkan perintah suara tersebut. Dalam banyak kasus, pendekatan yang melibatkan Deep Learning seperti Deep Neural Networks (DNN), Convolutional Neural Networks (CNN), dan Recurrent Neural Networks (RNN) telah memberikan hasil yang mengesankan dalam klasifikasi sinyal wicara untuk pengendalian kursi roda.

Begitu juga Terdapat banyak penelitian yang telah dilakukan tentang ekstraksi fitur sinyal wicara untuk menggerakkan kursi roda. Pada penelitian ini terdapat empat perintah untuk menggerakkan kursi roda yaitu empat jenis perintah suara yang harus dikenali yaitu go, left, right dan stop [14]. Penelitian menggunakan model simulator dan prototipe kontroler berbasis ANFIS bersama dengan sinyal sensor online yang dapat memaksimalkan kinerja kursi roda dan meningkatkan kualitas hidup penyandang disabilitas [15]. Penelitian bertujuan untuk mengekstraksi kumpulan fitur yang lebih berguna dari sinyal ucapan menggunakan teknik MFCC, dan pengoptimalan fitur menggunakan algoritma pengoptimalan ABC dan untuk pelatihan klasifikasi data, digunakan FFBN [16]. Penelitian dengan menerapkan model Markov tersembunyi (Hidden Markov Models) untuk mengenali emosi dalam sinyal wicara. Fitur-fitur seperti MFCC, energi, dan tingkat kecepatan bicara diekstraksi dan digunakan dalam pengenalan emosi [17]. Penelitian-penelitian di atas merupakan pendekatan dan teknik ekstraksi fitur yang digunakan dalam pengenalan sinyal wicara untuk berbagai tujuan, termasuk pengenalan ucapan otomatis dan pengenalan emosi berdasarkan suara. Fitur-fitur yang digunakan bervariasi tergantung pada tujuan penelitian dan konteks aplikasinya.

Pada penelitian tersebut, terdapat penelitian untuk menggerakkan kursi roda dengan empat perintah [14]. Sehingga untuk penelitian yang dilakukan yaitu menggunakan lima perintah (sinyal wicara yang digunakan ada lima perintah, yaitu: maju, mundur, kiri, kanan dan berhenti). Ekstraksi fitur yang digunakan menggunakan ekstraksi fitur yang berbeda dari peneliti sebelumnya, yaitu menggunakan penggabungan dua metode yaitu metode wavelet dikombinasikan dengan metode FFT.

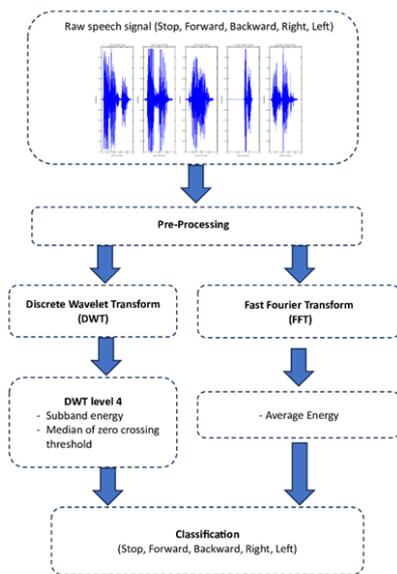
Pada penelitian yang dilakukan terdiri dari empat bagian yaitu bagian pertama berisi pendahuluan, bagian dua berisi tentang metode yang digunakan, bagian ketiga membahas tentang hasil, dan bagian empat kesimpulan.

2. Metode

Data yang digunakan dalam penelitian ini sebanyak 250 data sinyal wicara yang terdiri dari lima perintah maju, mundur, berhenti, kiri dan kanan. Data sinyal wicara diambil dari lima subyek, masing – masing subyek mengucapkan perintah majumundur, berhenti, kiri dan kanan sebanyak sepuluh kali. Subyek berumur antara 20 tahun sampai 40 tahun dengan jenis kelamin laki-laki. Durasi setiap rekaman kurang lebih 0.5 sampai 1 detik, dengan frekwensi sampling adalah 440000 Hz, direkam menggunakan aplikasi audacity.

2.1. Blok Diagram Sistem

Pada penelitian ini, untuk menggerakkan kursi roda dengan lima perintah (berhenti, maju, mundur, kiri, dan kanan) digunakan ekstraksi fitur yang berbeda dari peneliti sebelumnya [14], yaitu menggunakan metode FFT dikombinasikan dengan metode Wavelet. Terdapat satu nilai ekstraksi ciri yang diambil dari FFT yaitu nilai Rata-rata Energi serta diambil dari dua nilai ekstrasi ciri metode wavelet yaitu Nilai subband energi dan Nilai Median of zero crossing threshold.



Gambar 1. Desain sistem ekstraksi fitur sinyal wicara

2.2. Pre Processing

Proses preprocessing pada sinyal wicara melibatkan serangkaian langkah untuk mempersiapkan sinyal sebelum melakukan analisis atau pengolahan lebih lanjut. Berikut

adalah langkah-langkah dalam penelitian yang dilakukan dalam proses preprocessing pada sinyal wicara yaitu normalisasi, langkah ini untuk menormalkan amplitudo sinyal wicara agar memiliki level yang seragam. Langkah berikutnya yaitu pemotongan Awal dan Akhir, beberapa rekaman suara mungkin memiliki bagian awal atau akhir yang tidak diinginkan, seperti kebisingan sebelum atau setelah ucapan. Pemotongan awal dan akhir dilakukan untuk menghilangkan bagian-bagian tersebut agar sinyal hanya berisi ucapan yang relevan. Proses selanjutnya yaitu pemotongan rentang waktu, dalam kasus di mana sinyal wicara panjang, mungkin diperlukan pemotongan menjadi segmen yang lebih kecil. Ini dapat membantu dalam analisis atau pengolahan berdasarkan segmen-segmen tersebut.

2.3. Ekstraksi Fitur

Pada penelitian ini terdapat dua metode ekstraksi fitur yang dilakukan, yaitu penggabungan metode Fast Fourier Transform (FFT) dan metode transformasi Wavelet. Ekstrasi ciri yang didapatkan dari FFT dan Wavelet digabung menjadi satu.

Ekstraksi ciri menggunakan metode FFT pada sinyal wicara adalah proses untuk mengonversi sinyal dari domain waktu menjadi domain frekuensi. FFT digunakan untuk menganalisis spektrum frekuensi dari sinyal wicara dan mengekstraksi ciri-ciri yang berguna dalam pengenalan suara. Pada penelitian yang dilakukan ekstrasi ciri yang digunakan pada proses FFT yaitu dengan mengambil nilai spektrum frekwensi [18]. Metode FFT adalah algoritma yang digunakan untuk mengubah sinyal waktu-domain menjadi domain frekuensi. Dalam konteks sinyal, FFT dapat digunakan untuk menganalisis spektrum frekuensi dari sinyal wicara. Berikut adalah formula dasar untuk menghitung FFT dari sinyal:

Persiapan Sinyal Sebelum menghitung FFT, sinyal waktu-domain harus dipersiapkan. Misalkan sinyal waktu-domain memiliki panjang N dan ditandai dengan $x(n)$, di mana n adalah indeks dari elemen dalam sinyal (mulai dari 0 hingga N-1). Untuk menghitung FFT dari sinyal, digunakan formula berikut:

$$X(k) = \sum [x(n) * e^{(-j*2\pi*k*\frac{n}{N})}] \quad (1)$$

untuk $k = 0, 1, 2, \dots, N-1$

Di mana:
 $X(k)$ adalah komponen frekuensi di indeks k dalam domain frekuensi.
 $x(n)$ adalah elemen sinyal di indeks n.
 \exp adalah fungsi eksponensial kompleks ($e^{(j\theta)}$).
 j adalah unit imajiner ($\sqrt{-1}$).
 N adalah panjang sinyal.

Proses ini melibatkan perkalian setiap elemen sinyal dengan faktor eksponensial kompleks dan

menjumlahkannya untuk setiap frekuensi. Hasilnya adalah spektrum frekuensi dari sinyal, dengan nilai $X(k)$ yang merepresentasikan kontribusi frekuensi pada indeks k . Setelah FFT dihitung, spektrum frekuensi dapat dilihat dengan menggambarkan magnitudo (amplitudo) dari setiap komponen frekuensi $X(k)$.

Pada penelitian ini, ekstraksi ciri dari metode transformasi wavelet dengan mengambil nilai dari Energi Koefisien yaitu Menghitung energi koefisien pada setiap tingkat resolusi atau skala wavelet [19].

Metode Transformasi Wavelet adalah teknik yang digunakan untuk menganalisis sinyal dalam domain waktu-frekuensi. Transformasi Wavelet memungkinkan pemisahan informasi frekuensi tinggi dan rendah dalam sinyal dengan resolusi waktu yang baik. Berikut adalah formula dasar untuk metode Transformasi Wavelet diskret pada sinyal:

Dekomposisi (Analisis):

Dalam dekomposisi, sinyal asli difilter dengan filter low-pass (pendekatan) dan high-pass (detail) yang berbeda. Setelah pemfilteran, sinyal di-downsample (dengan mengambil setiap sampel kedua) untuk mengurangi ukurannya menjadi setengah dari ukuran asli.

Koefisien Pendekatan (Low-pass):

$$cA[k] = \sum_n x[n] \cdot h[2k - n] \quad (2)$$

Koefisien Detail (High-pass):

$$cD[k] = \sum_n x[n] \cdot g[2k - n] \quad (3)$$

Dimana :

$x[n]$ adalah sinyal input.

h adalah koefisien filter low-pass (pendekatan).

g adalah koefisien filter high-pass (detail).

Rekonstruksi (Sintesis):

Setelah mendapatkan koefisien pendekatan dan detail, kita bisa merekonstruksi sinyal asli dengan up-sampling koefisien dan kemudian melewatkannya melalui filter invers dari filter analisis.

$$cD[k] = \sum_n x[n] \cdot g[2k - n] \quad (4)$$

Dimana:

\tilde{h} adalah koefisien filter invers dari filter pendekatan.

\tilde{g} adalah koefisien filter invers dari filter detail.

Dalam Transformasi Wavelet diskret (DWT), sinyal waktu-domain dibagi menjadi skala dan posisi yang berbeda, dan setiap koefisien $W(a, b)$ dihitung dengan mengalikan elemen sinyal dengan fungsi wavelet yang telah digeser dan diskalakan. Proses ini berulang untuk setiap skala dan posisi yang sesuai. Transformasi Wavelet memungkinkan analisis sinyal dengan resolusi waktu-

frekuensi yang baik. Setiap koefisien $W(a, b)$ dalam Transformasi Wavelet mencerminkan kontribusi frekuensi dan waktu pada skala dan posisi yang sesuai.

DWT adalah teknik pengolahan sinyal yang digunakan untuk menganalisis dan merepresentasikan sinyal dalam domain waktu-frekuensi. DWT menggunakan fungsi-fungsi wavelet diskrit sebagai basis untuk merepresentasikan sinyal. Langkah-langkah umum dalam penerapan DWT, yaitu memilih fungsi wavelet diskrit yang akan digunakan. Dalam penelitian ini peneliti memilih fungsi wavelet diskrit Daubechies. Fungsi wavelet diskrit Daubechies memiliki beberapa variasi yang dikenal sebagai "family" Daubechies, misalnya Daubechies-2 (db2), Daubechies-4 (db4), Daubechies-6 (db6), dan seterusnya. Angka di belakang "db" mengindikasikan jumlah koefisien dalam fungsi wavelet tersebut.

Keuntungan dari fungsi wavelet diskrit Daubechies adalah keberadaan fungsi wavelet dengan jumlah koefisien yang berbeda memungkinkan analisis dengan tingkat resolusi yang berbeda. Fungsi wavelet Daubechies juga memiliki karakteristik desain yang baik, seperti keberadaan momen nol hingga orde tertentu, penghapusan gelombang, dan pembedaan.

DWT Daubechies bekerja dengan membagi sinyal menjadi dua komponen, yaitu komponen approximation (approximation coefficient) dan komponen detail (detail coefficient). Komponen approximation merepresentasikan informasi frekuensi rendah atau gambaran umum dari sinyal, sedangkan komponen detail merepresentasikan informasi frekuensi tinggi atau perbedaan antara sinyal asli dengan komponen approximation. Melakukan downsampling terhadap hasil filter untuk mengurangi resolusi waktu sinyal. Downsampling mengambil setiap nilai kedua dari hasil filter, mengurangi jumlah sampel sebanyak setengah.

Hasil dari DWT adalah serangkaian koefisien wavelet yang merepresentasikan sinyal dalam domain waktu-frekuensi. Koefisien approximation menggambarkan komponen rendah sinyal, sementara koefisien detail menggambarkan komponen tinggi atau perbedaan antara sinyal asli dan komponen rendahnya pada setiap level transformasi.

2.4. Normalisasi

Normalisasi ekstraksi fitur digunakan sinyal wicara adalah memperbaiki kualitas dan mengoptimalkan informasi yang terkandung dalam sinyal suara sehingga dapat memfasilitasi tugas-tugas analisis dan pengenalan wicara. Normalisasi fitur adalah langkah penting dalam proses pengolahan sinyal wicara dan umumnya dilakukan sebelum fitur-fitur tersebut digunakan dalam algoritma pengenalan ucapan, pengenalan kata, atau aplikasi lainnya. Alasan mengapa normalisasi fitur diperlukan dalam ekstraksi fitur sinyal wicara adalah Menghilangkan

perbedaan skala, Sinyal wicara dapat bervariasi dalam amplitudo yang besar tergantung pada jarak dan volume ucapan. Normalisasi fitur memastikan bahwa seluruh fitur berada dalam rentang yang seragam sehingga algoritma pengenalan dapat lebih stabil dan lebih mudah dikonfigurasi. Sehingga dengan normalisasi yang dilakukan, maka inputan untuk klasifikasi mempunyai rentang nilai yang sama untuk setiap fitur yang diperoleh dari metode DWT dan FFT.

2.5. Klasifikasi menggunakan Jaringan saraf Tiruan Backpropagation

Klasifikasi sinyal wicara adalah proses mengategorikan atau mengklasifikasikan sinyal wicara ke dalam kategori atau kelas yang sesuai berdasarkan fitur atau karakteristik tertentu. Pada penelitian ini, klasifikasi sinyal wicara yang digunakan untuk menggerakkan kursi roda dibagi ke dalam lima kelas, yaitu maju, mundur, berhenti, kanan dan kiri. Metode Klasifikasi yang digunakan dalam penelitian ini yaitu Jaringan saraf tiruan backpropagation. Data Training dari sinyal wicara berjumlah 200 data sinyal wicara dan data testing sebanyak 50 data sinyal wicara. Data ini dikumpulkan kemudian dikategorikan atau diberi label sesuai dengan kelas yang diinginkan. Data ini akan digunakan sebagai data training untuk melatih model klasifikasi. Ekstraksi ciri dilakukan untuk mengonversi sinyal wicara menjadi representasi numerik yang dapat digunakan sebagai input untuk model klasifikasi. Ciri-ciri ini diperoleh dari metode FFT dan DWT.

Metode backpropagation adalah teknik yang sering digunakan dalam pelatihan jaringan saraf tiruan (artificial neural network) untuk klasifikasi sinyal wicara [20]. Metode Backpropagation melibatkan penggunaan algoritma gradien turunan (gradient descent) untuk mengoptimalkan parameter-parameter jaringan saraf. Berikut adalah formula yang umum digunakan dalam metode Backpropagation untuk klasifikasi sinyal wicara:

1. Feedforward Pass:

- Hitung keluaran (output) dari setiap neuron dalam jaringan untuk setiap sampel data pelatihan. Ini melibatkan mengalikan input dengan bobot dan meneruskannya melalui fungsi aktivasi untuk setiap lapisan neuron.
- Misalkan y_j adalah output dari neuron j pada lapisan output.

2. Perhitungan Kesalahan (Error Calculation):

- Hitung kesalahan (error) antara output yang dihasilkan oleh jaringan (y_j) dan label yang seharusnya (t_j) untuk setiap neuron output. kesalahan ini dihitung dengan menggunakan fungsi kerugian seperti Mean Squared Error (MSE).

3. Backpropagation:

- Perbarui bobot dan bias jaringan untuk mengurangi kesalahan yang dihasilkan. Ini dilakukan dengan menghitung gradien dari fungsi kerugian terhadap bobot dan bias, kemudian mengubah bobot dan bias dengan mengurangi gradien tersebut.

- Gradien bobot (ΔW) dihitung menggunakan aturan rantai (chain rule) sebagai berikut:

$$\Delta W_{ij} = -\eta * \frac{\partial E}{\partial W_{ij}} \quad (5)$$

di mana η adalah tingkat pembelajaran (learning rate), E adalah fungsi kerugian, dan W_{ij} adalah bobot yang menghubungkan neuron ke- i di lapisan sebelumnya dengan neuron ke- j di lapisan saat ini.

- Gradien bias (Δb) dihitung dengan cara yang serupa:

$$\Delta b_j = -\eta * \frac{\partial E}{\partial b_j} \quad (6)$$

di mana b_j adalah bias neuron ke- j di lapisan saat ini.

4. Perbarui Bobot dan Bias:

- Perbarui bobot dan bias menggunakan gradien yang dihitung:

$$W_{ij} = W_{ij} + \Delta W_{ij} \quad (7)$$

$$b_j = b_j + \Delta b_j \quad (8)$$

5. Ulangi Langkah 1 hingga 4:

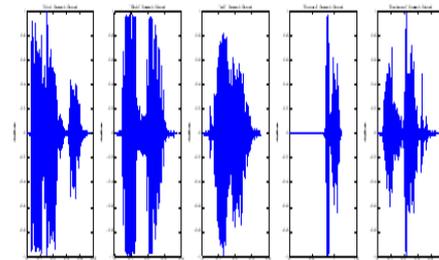
- Lakukan langkah-langkah ini untuk semua sampel data pelatihan dalam satu epoch.
- Ulangi proses pelatihan untuk sejumlah epoch yang ditentukan hingga jaringan mencapai kinerja yang memadai atau konvergensi.

Dengan melakukan iterasi langkah-langkah ini, jaringan saraf akan belajar untuk memahami fitur-fitur dalam sinyal wicara dan melakukan klasifikasi dengan lebih baik seiring berjalannya waktu.

3. Hasil dan analisis

3.1. Proses pengambilan Data

Data yang digunakan dalam penelitian ini adalah sebanyak 250 data sinyal wicara yang terdiri dari lima perintah maju, mundur, berhenti, kiri dan kanan. Berikut hasil perekaman sinyal wicara dengan lima perintah. Gambar 2 merupakan hasil perekaman yang ditampilkan dalam bentuk grafik.

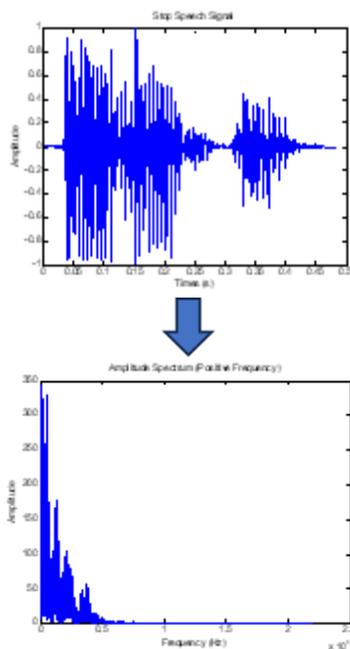


Gambar 2. Sinyal Wicara dengan perintah maju, mundur, berhenti, kiri dan kanan

Pada gambar 2, merupakan gambar grafik yang diperoleh dari perekaman untuk bunyi sinyal wicara maju, mundur, berhenti, kiri dan kanan. Perekaman sinyal wicara ini dilakukan oleh lima subyek. Masing-masing subyek merekam sinyal wicara sebanyak 10 kali. Sehingga data yang digunakan sebanyak 250 data sinyal wicara.

3.2. Ekstraksi Fitur menggunakan FFT

Ekstraksi fitur menggunakan metode Fast Fourier Transform (FFT) adalah salah satu pendekatan yang sering digunakan dalam pemrosesan sinyal dan analisis data, terutama untuk sinyal yang berubah seiring dengan waktu seperti sinyal audio atau sinyal dari sensor. FFT mengkonversi sinyal dari domain waktu ke domain frekuensi. Di domain frekuensi, sinyal dapat dianalisis berdasarkan komponen frekuensinya. Informasi ini dapat sangat berguna, misalnya, untuk mengidentifikasi frekuensi dominan dalam sinyal atau untuk menghapus noise dari sinyal. Algoritma FFT untuk mengubah sampel sinyal dari domain waktu ke domain frekuensi. Hasil dari FFT adalah serangkaian koefisien kompleks yang mewakili amplitudo dan fase dari masing-masing komponen frekuensi. Untuk keperluan analisis, pada magnitude dari hasil FFT. Magnitude ini dapat dihitung dengan mengambil norma dari masing-masing koefisien kompleks. Magnitude menunjukkan seberapa kuat komponen frekuensi tertentu dalam sinyal. Pada gambar 3, merupakan salah satu contoh hasil dari FFT dengan perintah berhenti.



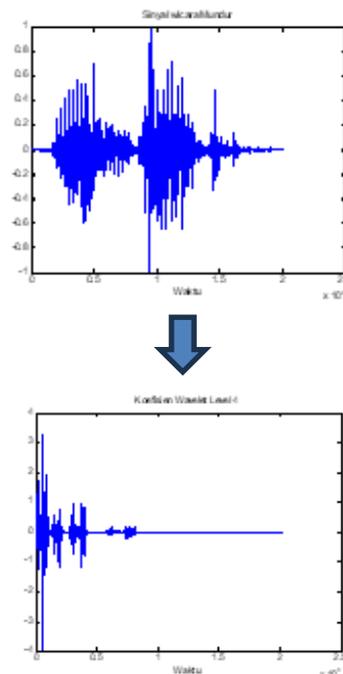
Gambar 3. Hasil Proses FFT dengan perintah berhenti

Pada gambar 3 merupakan hasil proses sinyal wicara yang direkam, kemudian di normalisasi dan diproses FFT. Dari hasil proses FFT ini, proses selanjutnya mengambil nilai rata-rata energi dari masing-masing perintah (maju, mundur, berhenti, kiri dan kanan).

3.3. Ekstraksi Fitur menggunakan metode DWT

DWT adalah metode lain yang digunakan untuk ekstraksi fitur, terutama pada sinyal. DWT menyediakan analisis

multi-resolusi, yang berarti sinyal dapat dianalisis pada berbagai skala atau tingkat detail. Dibandingkan dengan FFT yang mengubah sinyal ke domain frekuensi, DWT mengubah sinyal menjadi koefisien wavelet, yang merepresentasikan detail dan pendekatan dari sinyal di berbagai skala. DWT pada sinyal untuk mendapatkan koefisien pendekatan dan detail. Koefisien pendekatan menggambarkan komponen sinyal berskala besar (rendah frekuensi), sedangkan koefisien detail merepresentasikan komponen sinyal berskala kecil (frekuensi tinggi). Setelah mendapatkan koefisien, maka dapat dilakukan analisis untuk ekstraksi fitur. Fitur-fitur ini mungkin termasuk statistik dasar dari koefisien (seperti rata-rata, varians, dll.), energi, dan lain-lain. Pada penelitian ini, dari proses DWT, nilai subband energy dan zero crossing threshold dipilih sebagai nilai ekstraksi fitur.



Gambar 4. Hasil Proses DWT Level 4 dengan perintah Mundur

Pada gambar 4, merupakan hasil proses sinyal wicara yang direkam, kemudian di normalisasi dan diproses DWT. Dari hasil proses DWT ini, proses selanjutnya mengambil nilai subband energy dan zero crossing threshold dipilih sebagai nilai ekstraksi fitur dari masing-masing perintah (maju, mundur, berhenti, kiri dan kanan).

3.4. Klasifikasi menggunakan metode Jaringan saraf Backpropagation

Pada proses klasifikasi data dibagi menjadi set pelatihan (80%) dan set pengujian (20%). Data telah dibagi menjadi set pelatihan dengan 200 sampel dan set pengujian dengan 50 sampel.

Untuk mengklasifikasikan data ke dalam lima kelas berbeda yang mewakili "kiri", "kanan", "maju", "mundur", dan "berhenti", perlu memodifikasi pendekatan. Setiap file yang diunggah berhubungan dengan salah satu kelas. Oleh karena itu, diberikan label pada data yang sesuai dan kemudian melatih pengklasifikasi menggunakan Jaringan saraf Tiruan Backpropagation untuk membedakan kelima kelas.

Langkah-langkah untuk proses adalah sebagai berikut:

1. Memberi label pada Data: Menetapkan setiap kumpulan data label unik yang sesuai dengan kelasnya (kiri, kanan, maju, mundur, berhenti).
2. Menggabungkan dan Praproses Data: Menggabungkan kumpulan data menjadi satu, menangani nilai yang hilang atau tidak terbatas, dan standarisasi fitur.
3. Melatih klasifikasi: menggunakan data berlabel untuk melatih pengklasifikasi.
4. Evaluasi Pengklasifikasi: Menilai kinerja model menggunakan metrik seperti akurasi, matriks konfusi, dan laporan klasifikasi.

Hasil dari klasifikasi Backpropagation yang dilatih pada dataset yang diberi label lima kelas (kiri, kanan, maju, mundur, berhenti) adalah sebagai berikut:

Confusion Matrix: Matriksnya sangat besar, menunjukkan potensi masalah pada proses pelabelan atau klasifikasi. Tampaknya pengklasifikasi tidak secara efektif membedakan kelas-kelas sebagaimana dimaksud.

Classification Report: Laporan ini menunjukkan presisi, perolehan, dan skor F1 yang sangat rendah untuk sebagian besar kelas, kecuali satu kelas yang tampaknya mendominasi klasifikasi (diberi label 33509286.0 dalam laporan). Hal ini menunjukkan adanya ketidakseimbangan yang signifikan dalam representasi kelas atau masalah pada cara kelas dikodekan, Sehingga diperlukan perbaikan.

Berdasarkan hasil ini, nampaknya ada masalah dengan cara kelas yang ditugaskan atau dengan data itu sendiri. Ada kemungkinan variabel target tidak teridentifikasi dengan benar, sehingga menyebabkan klasifikasi salah.

Untuk mengatasinya, perlu Pastikan variabel target klasifikasi diidentifikasi dengan benar sebagai label kelas (0 untuk berhenti, 1 untuk kanan, 2 untuk kiri, 3 untuk maju, 4 untuk mundur).

Memeriksa kembali langkah pemrosesan data untuk memastikan fitur dan label yang digunakan benar.

Setelah memberi label dan memproses kumpulan data dengan benar, pengklasifikasi Backpropagation telah dilatih ulang dan dievaluasi. Berikut hasilnya:

1. Confusion Matrix:

10	0	0	0	0
0	10	0	0	0
0	0	10	0	0
0	0	0	10	0
0	0	0	0	10

Dari tabel confusion matrix diatas, maka diperoleh analisa hasil prediksi untuk lima kelas, yaitu:

Kelas 0: TP = 10, FP = 0, FN = 0

Kelas 1: TP = 10, FP = 0, FN = 0

Kelas 2: TP = 10, FP = 0, FN = 0

Kelas 3: TP = 10, FP = 0, FN = 0

Kelas 4: TP = 10, FP = 0, FN = 0

Hal ini menunjukkan klasifikasi sempurna dan tidak ada kesalahan klasifikasi. Setiap kelas (kiri, kanan, maju, mundur, berhenti) telah diprediksi secara akurat.

2. Classification Report:

Presisi, perolehan, dan skor F1 semuanya 1,00 untuk setiap kelas. Akurasi keseluruhan model adalah 1,00. Hasil ini menunjukkan bahwa pengklasifikasi Backpropagation memiliki kinerja yang sangat baik pada kumpulan data, dengan akurasi sempurna dalam membedakan kelima kelas. Tingkat akurasi ini sangat baik, namun penting untuk mempertimbangkan kemungkinan overfitting, terutama jika kumpulan data tidak cukup beragam atau jika kumpulan pengujian tidak mewakili skenario dunia nyata. Classification Report dan Confusion Matrix yang diberikan sebelumnya menunjukkan hasil set pengujian. Untuk mengevaluasi keakuratan set pelatihan dan pengujian, akan menghitungnya secara terpisah. Melanjutkan dengan menghitung akurasi pelatihan dan pengujian untuk pengklasifikasi Backpropagation. Keakuratan pengklasifikasi Backpropagation adalah: Akurasi Pelatihan: 100% (1.0) dan Akurasi Pengujian: 100% (1.0). Akurasi pelatihan dan pengujian keduanya sempurna, yang menunjukkan bahwa model berperforma sangat baik pada kumpulan data ini. Meskipun akurasi tinggi seperti itu diinginkan, penting juga untuk memastikan bahwa model tidak melakukan overfitting pada data pelatihan dan bahwa data pengujian mewakili skenario dunia nyata.

Akurasi pengklasifikasi Backpropagation dengan jumlah lapisan tersembunyi yang berbeda pada set pengujian adalah sebagai berikut: 1 Lapisan Tersembunyi: 95,95%, 2 Lapisan Tersembunyi: 100% dan 3 Lapisan Tersembunyi: 100%

Model dengan dua dan tiga lapisan tersembunyi mencapai akurasi sempurna, sedangkan model dengan satu lapisan tersembunyi menunjukkan akurasi yang sedikit lebih rendah. Hal ini menunjukkan bahwa kompleksitas tambahan (lebih banyak lapisan tersembunyi) membantu model menangkap nuansa kumpulan data tertentu dengan lebih efektif.

Tabel 1. Komparasi hasil pengujian dengan penelitian lain

Peneliti	Judul	Metode	Hasil
Mohsen Bakouri dkk, 2022	Steering a Robotic Wheelchair Based on Voice Recognition System Using Convolutional Neural Networks	A convolutional neural network (CNN)-based network-in-network (NIN)	87,2%
Sutikno dkk, 2021	Voice Controlled Wheelchair for Disabled Patients based on CNN and LSTM	Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM)	97.80 %
Mohsen Bakouri dkk, 2022	Development of Voice Control Algorithm for Robotic Wheelchair Using NIN and LSTM Models	a network in network (NIN) and long shortterm memory (LSTM) st	98.2%
Arief Wisaksono dkk, 2024	Klasifikasi Sinyal Wicara untuk Gerakan Kursi Roda menggunakan metode Jaringan Saraf Tiruan Backpropagation	Backpropagation, FFT dan DWT	100%

4. Kesimpulan

Dalam penelitian ini, metode klasifikasi telah diterapkan untuk mengklasifikasikan fitur ekstraksi dari sinyal wicara yang bertujuan untuk mengendalikan gerakan kursi roda, dengan menggunakan kombinasi DWT dan FFT. Dari metode yang diuji, yakni metode jaringan saraf tiruan backpropagation terbukti memberikan hasil terbaik. Dengan menggunakan metode jaringan saraf tiruan backpropagation, pada penelitian ini berhasil mencapai akurasi akurasi pengujian sebesar 100% untuk hidden later 2 dan hidden layer 5. Untuk penelitian mendatang, disarankan untuk mengeksplorasi metode ekstraksi fitur lain yang mungkin lebih sesuai untuk lima kelas yang berbeda, dengan harapan meningkatkan akurasi yang jauh lebih baik.

Referensi

- [1]. D. S. Kulkarni, "A Review of Speech Signal Enhancement Techniques," no. August, 2019, doi: 10.5120/ijca2016909507.
- [2]. E. Science, "Signal acquisition system based on wireless transmission for environmental sound monitoring system Signal acquisition system based on wireless transmission for environmental sound monitoring system", doi: 10.1088/1755-1315/969/1/012015.
- [3]. J. Kim, G. Bailly, C. Davis, J. Kim, G. Bailly, and C. Davis, "Introduction to the special issue on auditory-visual expressive speech and gesture in humans and machines To cite this version : HAL Id : hal-01821001," 2018.
- [4]. T. Hueber *et al.*, "Statistical Mapping between Articulatory and Acoustic Data , Application to Silent Speech Interface and Visual Articulatory Feedback".
- [5]. D. Ahmad, "Voice based virtual assistant," no. July, 2023, doi: 10.13140/RG.2.2.32116.12163.
- [6]. A. Vasilateanu and R. Ene, "Call-Center Virtual Assistant Using Natural Language Processing and Speech Recognition Call-Center Virtual Assistant Using Natural Language Processing and Speech Recognition," no. December 2018, 2020, doi: 10.33150/JITDETS-2.2.3.
- [7]. H. B. Pasandi and H. B. Pasandi, "Evaluation of Automated Speech Recognition Systems for Conversational Speech : A Linguistic Perspective".
- [8]. M. I. Malik, T. Bashir, and O. F. Khan, "Voice Controlled Wheel Chair System," vol. 6, no. 6, pp. 411–419, 2017.
- [9]. H. Krishna, S. Emani, R. Manapaka, and S. P. V. S. Rao, "VOICE CONTROLLED AND JOYSTICK BASED WHEEL CHAIR FOR DIFFERENTLY ABLED PEOPLE," no. 06, pp. 3111–3114, 2021.
- [10]. C. Bartolein and A. Wagner, "Easing Wheelchair Control by Gaze-based Estimation of Intended Motion," pp. 9162–9167, 2008.
- [11]. M. S. Arsha, A. R. Raj, S. R. Pooja, R. Manoj, S. A. Sabitha, and S. Mohan, "Voice Controlled Wheelchair," no. 9, pp. 1–4, 2020.
- [12]. M. Bakouri *et al.*, "Steering a Robotic Wheelchair Based on Voice Recognition System Using Convolutional Neural Networks," *Electron.*, vol. 11, no. 1, pp. 1–17, 2022, doi: 10.3390/electronics11010168.
- [13]. M. Bakouri, "Development of Voice Control Algorithm for Robotic Wheelchair Using MIN and LSTM Models," *Comput. Mater. Contin.*, vol. 73, no. 2, pp. 2441–2456, 2022, doi: 10.32604/cmc.2022.025106.
- [14]. M. S. I. Sharifuddin, S. Nordin, and A. M. Ali, "Comparison of CNNs and SVM for voice control wheelchair," *IAES Int. J. Artif. Intell.*, vol. 9, no. 3, pp. 387–393, 2020, doi: 10.11591/ijai.v9.i3.pp387-393.
- [15]. M. M. Abdulghani, K. M. Al-Aubidy, M. M. Ali, and Q. J. Hamarsheh, "Wheelchair neuro fuzzy control and tracking system based on voice recognition," *Sensors (Switzerland)*, vol. 20, no. 10, 2020, doi: 10.3390/s20102872.
- [16]. G. Kaur, M. Srivastava, and A. Kumar, "Integrated Speaker and Speech Recognition for Wheel Chair Movement Using Integrated Speaker and Speech Recognition for Wheel Chair Movement using Artificial Intelligence," no. November 2019, 2018, doi: 10.31449/inf.v42i4.2003.
- [17]. Q. P. Ha, T. H. Tran, and G. Dissanayake, "A wavelet- and neural network-based voice interface system for wheelchair control A wavelet- and neural network-based voice interface system for wheelchair control," no. December, pp. 48–65, 2013, doi: 10.1504/IJISTA.2005.007307.
- [18]. D. T. Kusuma, "Fast Fourier Transform (FFT) Dalam Transformasi Sinyal Frekuensi Suara Sebagai Upaya Perolehan Average Energy (AE) Musik," vol. 14, no. 1, pp. 28–35, 2021.
- [19]. S. Pittner and S. Kamarthi, "Coefficients for Pattern Recognition Tasks," no. January 1999, 2015, doi: 10.1109/34.745739.
- [20]. O. Bilokon and I. Denkov, "Backpropagation artificial neural network learning algorithm process impact based on hyperparameters," vol. 14, no. 8, pp. 1–15, 2021.