

## SISTEM IDENTIFIKASI SUARA BERBASIS ANALISIS SINYAL WICARA MENGGUNAKAN ALGORITMA FAST FOURIER TRANSFORM

Khidru Fawzan Suroqoh<sup>\*)</sup> dan Joko Hariyono

Program Studi Teknik Elektro, Universitas Sebelas Maret, Surakarta, Indonesia

<sup>\*)</sup>Penulis korespondensi, E-mail: khidruadewale@gmail.com

### Abstrak

Komunikasi merupakan aspek yang sangat penting dalam kehidupan manusia, dan berbicara adalah cara paling alami serta efisien untuk menyampaikan informasi. Setiap individu memiliki karakteristik suara yang unik yang dipengaruhi oleh faktor usia dan jenis kelamin. Frekuensi suara menjadi salah satu parameter utama yang dapat membedakan suara manusia, di mana suara laki-laki memiliki frekuensi lebih rendah, perempuan lebih tinggi, dan anak-anak memiliki frekuensi tertinggi. Penelitian ini mengusulkan metode untuk menganalisis suara manusia dan mengklasifikasikannya berdasarkan frekuensi dominan menggunakan algoritma Fast Fourier Transform (FFT). Data diperoleh melalui dua jenis rekaman, yaitu rekaman langsung dalam ruangan kedap suara dan rekaman eksternal yang telah terkompresi melalui aplikasi WhatsApp, kemudian dianalisis menggunakan perangkat lunak MATLAB dengan penerapan algoritma FFT untuk mengekstraksi spektrum frekuensi suara. Hasil penelitian menunjukkan bahwa algoritma FFT memiliki tingkat akurasi sebesar 96,66% dalam mengidentifikasi dan membedakan suara manusia berdasarkan rentang frekuensi yang telah ditentukan, namun akurasi menurun pada rekaman suara yang terkompresi dari sumber eksternal karena perubahan spektrum frekuensi dan sifat akustik rekaman tersebut[1]. Temuan ini menunjukkan bahwa FFT sangat efektif untuk data berkualitas tinggi, tetapi sensitif terhadap kompresi audio, sehingga penelitian ini dapat menjadi dasar pengembangan sistem cerdas untuk aplikasi seperti pengenalan suara, asisten virtual, dan sistem interaksi manusia-mesin.

*Kata kunci: Fast Fourier Transform, Frekuensi, Sinyal wicara, Software MatLab*

### Abstract

*Communication is a crucial aspect of human life, and speech is the most natural and efficient way to convey information. Each individual has unique voice characteristics influenced by age and gender. Voice frequency is one of the main parameters that distinguish human voices, where male voices tend to have lower frequencies, female voices higher, and children's voices the highest. This study proposes a method to analyze and classify human voices based on dominant frequency using the Fast Fourier Transform (FFT) algorithm. The data were obtained from two types of recordings: direct recordings in a soundproof room and compressed external recordings from the WhatsApp application, which were then analyzed using MATLAB software with FFT implementation to extract the voice frequency spectrum. The results show that the FFT algorithm achieved an accuracy rate of 96.66% in identifying and distinguishing human voices according to predetermined frequency ranges, but accuracy decreased for compressed external recordings due to changes in frequency spectra and acoustic characteristics[1]. These findings indicate that FFT is highly effective for high-quality data but sensitive to audio compression, making this study a foundation for developing intelligent systems for applications such as voice recognition, virtual assistants, and human-machine interaction systems.*

*Keywords: Fast Fourier Transform, Frequency, Sinyal wicara, MatLab Software.*

### 1. Pendahuluan

Manusia sebagai makhluk sosial, memiliki sifat dan karakteristik yang beragam, termasuk dalam hal suara. Setiap individu menghasilkan suara yang unik, yang dipengaruhi oleh berbagai faktor seperti usia, jenis kelamin, dan anatomi tubuh[2]. Salah satu karakteristik yang membedakan suara manusia adalah frekuensi suaranya. Secara umum, suara laki-laki dewasa cenderung memiliki frekuensi lebih rendah dibandingkan dengan perempuan dewasa dan anak-anak. Rentang frekuensi

suara laki-laki biasanya berkisar antara 85 hingga 180 Hz, suara perempuan antara 165 hingga 255 Hz, dan suara anak-anak memiliki rentang frekuensi antara 225 Hz hingga 500 Hz [2]. Perbedaan ini membuka peluang untuk melakukan klasifikasi suara berdasarkan jenis kelamin dan usia, yang dapat digunakan untuk berbagai aplikasi teknologi, seperti pengenalan suara, asisten virtual, dan sistem keamanan.

Seiring dengan perkembangan teknologi, analisis sinyal wicara telah menjadi bidang yang semakin penting. Salah

satu teknik yang digunakan dalam analisis suara adalah Fast Fourier Transform (FFT). FFT merupakan algoritma yang digunakan untuk mengubah sinyal dari domain waktu menjadi domain frekuensi, yang memungkinkan kita untuk menganalisis karakteristik suara, seperti frekuensi dominan [3]. Melalui analisis frekuensi ini, kita dapat mengidentifikasi dan membedakan suara berdasarkan jenis kelamin dan usia. Dalam penelitian ini, kami berfokus pada tiga kategori suara: suara laki-laki dewasa, suara perempuan dewasa, dan suara anak-anak.

Klasifikasi suara berdasarkan frekuensi merupakan pendekatan yang umum digunakan, di mana masing-masing kategori suara memiliki rentang frekuensi dan karakteristik yang berbeda. Sebagai contoh, pada suara perempuan terdapat klasifikasi lebih lanjut, yaitu suara tinggi (soprano), suara sedang (mezzo), dan suara rendah (alto). Demikian pula pada suara laki-laki, terdapat klasifikasi suara tinggi (tenor), suara sedang (baritone), dan suara rendah (bass). Perbedaan rentang frekuensi ini dapat digunakan untuk membedakan suara satu dengan yang lainnya, baik itu antara laki-laki, perempuan, atau anak-anak [4], [5], [6]

Penelitian ini bertujuan untuk menganalisis sinyal wicara menggunakan algoritma Fast Fourier Transform (FFT) untuk mengidentifikasi jenis suara berdasarkan frekuensi dominan yang dihasilkan oleh suara laki-laki dewasa, perempuan dewasa, dan anak-anak. Analisis ini dilakukan menggunakan software MATLAB sebagai platform simulasi untuk memperoleh hasil yang akurat dan efisien.

Kebaruan (*novelty*) dari penelitian ini terletak pada penggunaan dua jenis sumber data suara, yaitu rekaman langsung dan rekaman terkompresi eksternal yang berasal dari aplikasi WhatsApp, yang jarang bahkan belum digunakan dalam penelitian serupa sebelumnya. Pendekatan ini memungkinkan peneliti untuk mengevaluasi pengaruh kompresi data audio terhadap karakteristik frekuensi dominan dalam sinyal wicara, serta menguji keandalan algoritma FFT dalam kondisi data yang berbeda tingkat kualitas dan formatnya.

Sebagai langkah awal menuju pengembangan sistem identifikasi suara yang lebih cerdas, penelitian ini juga dimaksudkan sebagai fondasi bagi penelitian lanjutan yang akan mengintegrasikan metode Mel-Frequency Cepstral Coefficients (MFCC) dan pendekatan machine learning. Integrasi tersebut diharapkan dapat meningkatkan akurasi klasifikasi serta ketahanan sistem terhadap variasi kualitas rekaman, termasuk yang telah dikompresi seperti dari aplikasi WhatsApp.

Dengan memahami perbedaan frekuensi suara serta pengaruh kompresi terhadap hasil analisis, penelitian ini diharapkan dapat memberikan kontribusi baru dalam pengembangan sistem pengenalan suara yang lebih adaptif dan robust terhadap variasi kualitas rekaman. Sistem

semacam ini berpotensi meningkatkan performa aplikasi berbasis suara, seperti asisten virtual, sistem keamanan berbasis biometrik suara, serta interaksi manusia-mesin. Secara keseluruhan, penelitian ini tidak hanya memperkuat pemahaman tentang karakteristik frekuensi suara manusia, tetapi juga menghadirkan pendekatan baru dalam analisis sinyal wicara yang mempertimbangkan variasi sumber dan format rekaman audio.

### **1.1. Software MATLAB**

Matlab merupakan software yang mempunyai beberapa keunggulan dalam pengolahan grafik dan suara. Salah satu kelebihan Matlab yaitu mampu menampilkan dan mengolah grafik serta suara dengan command yang sederhana dan fleksibel. Bahkan dapat digunakan untuk perancangan simulasi GUI (Graphical User Interface) dalam mengklasifikasikan jenis suara secara otomatis berdasarkan gender, range usia dan kelas frekuensi suara[7].

### **1.2. Sinyal Wicara**

Suara adalah fenomena fisik yang dihasilkan oleh getaran benda atau getaran suatu benda yang berupa sinyal analog dengan amplitudo yang berubah secara kontinu terhadap waktu, suara berhubungan erat dengan rasa mendengar [8]. Berdasarkan bentuknya data dari sinyal wicara dapat dibedakan kedalam sinyal analog dan sinyal digital. Sinyal dikatakan analog apabila amplitudo dari data atau sinyal tersebut terus menerus ada dalam rentang waktu tertentu (kontinu) dan memiliki variasi nilai amplitudo tak terbatas. Sedangkan sinyal dikatakan digital apabila amplitudo dari data atau sinyal tersebut tidak kontinu dan memiliki variasi nilai amplitudo yang terbatas [9]. Pada umumnya, batas frekuensi bunyi yang dapat didengar oleh telinga manusia kira-kira dari 20 Hz sampai 20 kHz [10].

### **1.3. Fast Fourier Transform (FFT)**

Fast Fourier Transform (FFT) adalah teknik perhitungan operasi matematika yang digunakan untuk mentransformasi sinyal analog menjadi sinyal digital berbasis frekuensi. FFT merupakan pengembangan dari discrete fourier transform (DFT) yang memberikan pengolahan sinyal lebih cepat. Prinsip kerja FFT adalah membagi sinyal sampling menjadi dua bagian kemudian masing-masing bagian diselesaikan dengan algoritma (DFT)[11]. Karena sinyal-sinyal dalam sistem komunikasi bersifat kontinu, sehingga hasilnya dapat digunakan untuk transformasi fourier [9]. Algoritma FFT merupakan salah satu metode untuk mentransformasi sinyal wicara (speech) dalam domain waktu menjadi sinyal domain frekuensi yang artinya proses perekaman suara disimpan dalam bentuk digital berupa gelombang spektrum suara yang berbasis frekuensi sehingga lebih mudah dalam menganalisa spektrum [3].

Algoritma Fast Fourier Transform (FFT) memiliki sejumlah kelebihan yang menjadikannya sangat efektif untuk sistem identifikasi suara berbasis analisis sinyal wicara. FFT dikenal memiliki efisiensi komputasi yang tinggi, karena mampu menguraikan sinyal waktu menjadi domain frekuensi dengan cepat dan akurat, bahkan pada dataset berukuran besar[12]. Selain itu, algoritma ini bersifat sederhana dan stabil, mudah diimplementasikan pada berbagai platform seperti MATLAB, serta menghasilkan spektrum frekuensi yang jelas tanpa memerlukan banyak parameter tambahan[13]. Dalam konteks identifikasi suara, FFT sangat sesuai untuk mengekstraksi karakteristik dasar suara, karena dapat menampilkan puncak frekuensi dominan yang berhubungan langsung dengan perbedaan antara suara laki-laki, perempuan, dan anak-anak[14].

## 2. Literature Review

Analisis sinyal wicara berbasis Fast Fourier Transform (FFT) merupakan salah satu metode klasik yang digunakan untuk mengekstraksi informasi frekuensi dominan dari sinyal wicara manusia. FFT memungkinkan konversi sinyal waktu menjadi domain frekuensi sehingga karakteristik spektral, seperti frekuensi dasar (fundamental frequency) dan harmonisa, dapat diidentifikasi dengan jelas[15]. Metode ini telah banyak diterapkan dalam sistem pengenalan dan identifikasi suara untuk membedakan jenis kelamin, usia, maupun karakteristik pembicara karena keakuratannya dalam mengungkapkan pola energi dan distribusi frekuensi[16]. Beberapa penelitian menunjukkan bahwa implementasi algoritma FFT menggunakan **MATLAB** memberikan hasil yang efisien dan mudah dianalisis, karena **MATLAB** menyediakan fungsi pengolahan sinyal yang terintegrasi dan visualisasi spektrum yang mendukung validasi hasil eksperimen.

Selain faktor algoritma, kualitas sumber data audio juga memengaruhi hasil analisis spektral. Penelitian oleh Fuchs (2016) menunjukkan bahwa kompresi audio lossy seperti MP3 dapat mengubah parameter akustik penting seperti formant dan pitch, sehingga menggeser nilai frekuensi dominan dari sinyal aslinya[17]. Hasil serupa dikemukakan oleh Niebuhr (2023) yang menemukan bahwa setiap codec kompresi (misalnya AAC, Opus, dan AMR) menghasilkan distorsi spektral yang spesifik, sehingga pengenalan suara berbasis fitur frekuensi dapat terpengaruh oleh perbedaan format audio[18]. Lebih lanjut, Oreskovic (2024) menegaskan bahwa kompresi yang digunakan pada platform komunikasi seperti *WhatsApp* dapat mengubah pola amplitudo dan distribusi spektral yang digunakan dalam proses ekstraksi fitur, yang berpotensi menurunkan akurasi sistem pengenalan suara[19].

Meskipun banyak penelitian sebelumnya berfokus pada analisis sinyal wicara menggunakan data rekaman langsung atau dataset standar yang tidak terkompresi, belum banyak studi yang secara eksplisit membandingkan

rekaman langsung dengan rekaman terkompresi eksternal dari aplikasi pesan instan seperti *WhatsApp*. Oleh karena itu, penelitian ini menawarkan kontribusi baru dengan menganalisis perbedaan karakteristik frekuensi dominan antara dua jenis sumber data tersebut menggunakan algoritma FFT. Pendekatan ini diharapkan dapat memberikan pemahaman yang lebih mendalam mengenai pengaruh kompresi terhadap hasil analisis sinyal wicara serta meningkatkan ketahanan (*robustness*) sistem identifikasi suara dalam menghadapi variasi kualitas rekaman di dunia nyata.

## 3. Metode

Penelitian ini menggunakan data sampel suara manusia yang direkam secara langsung pada perangkat lunak Matlab dan sampel suara yang sudah tersimpan di pesan suara *WhatsApp*. Durasi setiap sampel audio adalah 5 detik, dan seluruh data suara menggunakan sampling rate sebesar 44.100 Hz. Kategori suara yang digunakan dalam penelitian ini adalah Suara Laki-laki Dewasa, Suara Perempuan Dewasa, dan Suara Anak-anak.

### 3.1. Pengumpulan Data

Data suara dikumpulkan untuk masing-masing kategori suara (laki-laki dewasa, perempuan dewasa, dan anak-anak) dengan bertemu secara langsung dan merekam suara di perangkat lunak Matlab, sedangkan beberapa suara yang sudah tersimpan di pesan suara *WhatsApp*, file tersebut berformat MP3 dan kemudian dikonversi menjadi WAV menggunakan perangkat lunak konversi audio. Jumlah sampel suara yang digunakan dalam percobaan ini adalah 30 suara, setiap klasifikasi dengan 10 sampel suara.

```
%Kodingan pada Matlab
sampleRate = 44100;
duration = 5;
reader = audioDeviceReader('SamplesPerFrame', 1024,
'SampleRate', sampleRate);
player = audioDeviceWriter('SampleRate', sampleRate);
recordedAudio = [];
disp('Recording...');
tic
while toc < duration
    audioData = reader();
    recordedAudio = [recordedAudio; audioData];
end
```

### 3.2. Pra-Pemrosesan Data

Proses analisis diawali dengan mengonversi sinyal stereo menjadi mono untuk menyederhanakan data audio yang dianalisis. Setelah itu, dilakukan normalisasi amplitudo guna memastikan tingkat suara seragam pada setiap sampel. Selanjutnya, semua file audio di-resample ke dalam sampling rate yang konsisten, yaitu 44.100 Hz,

untuk menjaga keseragaman data. Proses dilanjutkan dengan penerapan filter band-pass yang hanya mempertahankan frekuensi dalam rentang 85 Hz hingga 255 Hz—rentang yang relevan dengan suara manusia—sekaligus mengurangi gangguan dari frekuensi di luar batas tersebut.

Setelah tahapan praproses selesai, dilakukan analisis Fast Fourier Transform (FFT) untuk mengubah sinyal dari domain waktu ke domain frekuensi. Melalui FFT, spektrum frekuensi dari setiap sinyal dapat dianalisis, dan frekuensi dominan diekstraksi untuk menentukan kategori suara berdasarkan rentang frekuensi yang telah ditentukan sebelumnya.

Kemudian, suara yang telah dianalisis kemudian diklasifikasikan berdasarkan rentang frekuensi dominan yang sesuai dengan kategori suara. Adapun rentang frekuensi untuk masing-masing kategori adalah sebagai berikut: suara laki-laki dewasa berada pada rentang 85–180 Hz, suara perempuan dewasa pada rentang 165–255 Hz, dan suara anak-anak memiliki frekuensi dominan di atas 255 Hz.

### 3.3. Implementasi

Implementasi sistem dilakukan menggunakan perangkat lunak MATLAB untuk merekam suara sampel baru dan menganalisisnya menggunakan algoritma FFT. Hasil analisis suara dibandingkan dengan sampel-sampel suara yang telah dilatih sebelumnya. MATLAB juga digunakan untuk memplot sinyal audio dan menampilkan frekuensi dominan yang dihasilkan dari sinyal yang direkam.

Sistem ini mampu mengidentifikasi suara yang direkam dan menghasilkan output sesuai dengan kategori suara yang terdeteksi berdasarkan frekuensi dominannya. Selain itu, sinyal audio yang direkam dapat divisualisasikan dalam bentuk plot untuk mempermudah interpretasi hasil. Untuk membedakan sinyal wicara berdasarkan klasifikasinya, sistem memplot sinyal dengan warna tertentu:

- Suara laki-laki ditampilkan dalam warna biru.
- Suara perempuan ditampilkan dalam warna magenta.
- Suara anak-anak ditampilkan dalam warna merah.
- Jika suara tidak terklasifikasi dalam salah satu kategori tersebut, maka waveform akan ditampilkan dalam warna hitam.

### 3.4. Penulisan Persamaan

Beberapa rumus yang berkaitan dan sering digunakan dalam algoritma Fast Fourier Transform adalah sebagai berikut:

1) *Frekuensi Dominan*

$$f_{peak} = \frac{k_{peak}}{N} \times f_s \quad (1)$$

dimana  $f_s$  adalah frekuensi sampling dan  $N$  adalah jumlah titik data pada sinyal.

2) *Indeks Peak*

$$k_{peak} = \text{argkmax}(|X[k]|) \quad (2)$$

dimana  $|X[k]|$  adalah nilai amplitudo pada indeks  $k$

3) *Persamaan DFT*

$$X_{(k)} = \frac{1}{N} \sum_{n=0}^{N-1} x[n] e^{-i \frac{2\pi}{N} kn} \quad (3)$$

dimana  $k$  adalah indeks frekuensi dan  $N$  adalah jumlah data.

### 3.5. Alur Penelitian

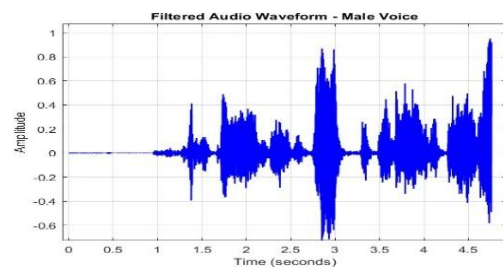


**Gambar 1. Diagram Alur penelitian**

## 4. Hasil dan Pembahasan

Setelah melakukan rekaman suara dengan mengambil sepuluh sampel untuk setiap klasifikasi yaitu wicara laki-laki dewasa, perempuan dewasa dan wicara anak-anak, berikut adalah hasil transformasi Fast Fourier Transform (FFT) yang mengubah sinyal wicara dari domain waktu menjadi domain frekuensi.

### 4.1. Wicara Laki-laki Dewasa

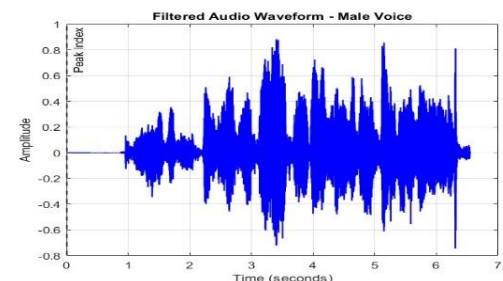


**Gambar 2. Sampel suara Laki-laki pertama**

Index of the peak: 636

Dominant Frequency: 137,423 Hz

This is likely a male voice.

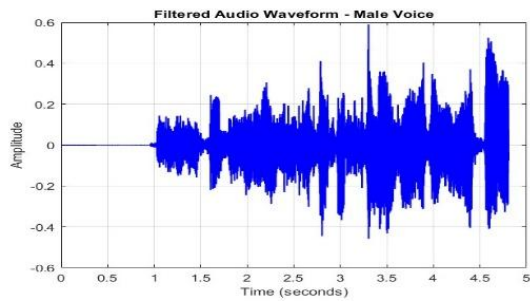


**Gambar 3. Sampel suara Laki-laki kedua**

Index of the peak: 523

Dominant Frequency: 108,0801 Hz

This is likely a male voice.



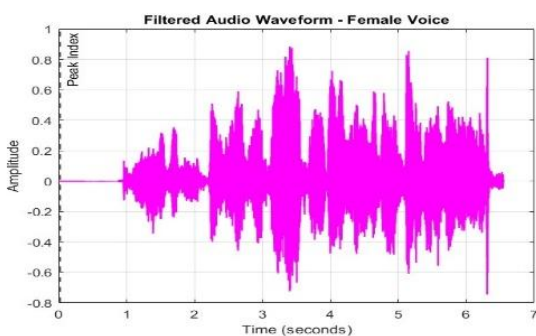
Gambar 4. Sampel suara Laki-laki ketiga

Index of the peak: 597  
Dominant Frequency: 129.1992 Hz  
This is likely a male voice.

Sepuluh sampel wicara yang berbeda digunakan untuk dalam system algoritma FFT, suara laki-laki yang direkam langsung menggunakan alat rekaman pada perangkat laptop melalui MATLAB, serta suara yang sudah direkam sebelumnya melalui aplikasi pesan suara di WhatsApp. Semua suara tersebut berhasil diklasifikasikan sebagai suara laki-laki oleh algoritma FFT.

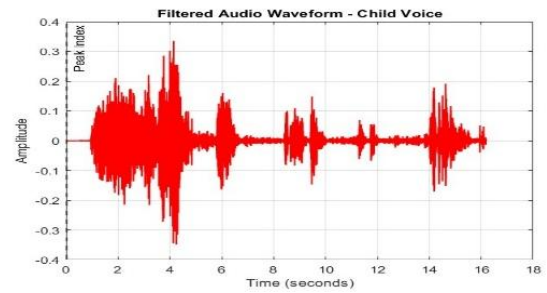
Pada suara pertama hasil analisis FFT menunjukkan Indeks peak: 636 dan frekuensi dominan: 137.423 Hz, algoritma mengklasifikasikan suara ini sebagai suara laki-laki dengan benar. Pada suara kedua hasil analisis FFT menunjukkan Indeks peak: 523 dan frekuensi dominan: 108.0801 Hz, suara ini juga diklasifikasikan dengan benar sebagai suara laki-laki. Kemudian pada suara yang ketiga meskipun berasal dari rekaman yang sudah ada, hasil analisis FFT menunjukkan Indeks peak: 597 dan frekuensi dominan: 129.1992 Hz, algoritma tetap berhasil mengklasifikasikan suara ini sebagai suara laki-laki.

#### 4.2. Wicara Perempuan Dewasa



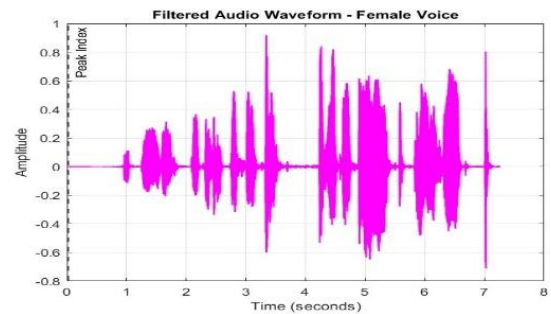
Gambar 5. Sampel suara perempuan pertama

Index of the peak: 915  
Dominant Frequency: 195.8343 Hz  
This is likely a female voice.



Gambar 6. Sampel suara perempuan kedua

Index of the peak: 1479  
Dominant Frequency: 307.4983 Hz  
This is likely a child's voice.

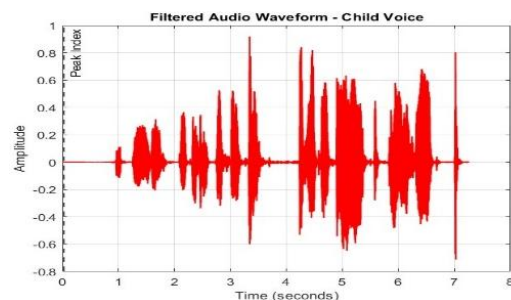


Gambar 7. Sampel suara perempuan ketiga

Index of the peak: 1004  
Dominant Frequency: 214.9035 Hz  
This is likely a female voice.

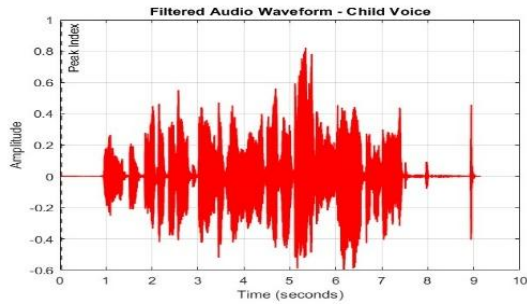
Dari sepuluh rekaman suara perempuan yang dianalisis, algoritma FFT berhasil mengklasifikasikan sembilan sampel suara dengan benar sebagai suara perempuan. Namun, satu sampel suara lainnya salah diklasifikasikan sebagai suara anak-anak, kesalahan klasifikasi tersebut kemungkinan disebabkan oleh nada suara yang sangat tinggi serta artefak akibat kompresi WhatsApp, yang mengubah karakteristik spektrum frekuensi.

#### 4.3. Wicara Anak-anak



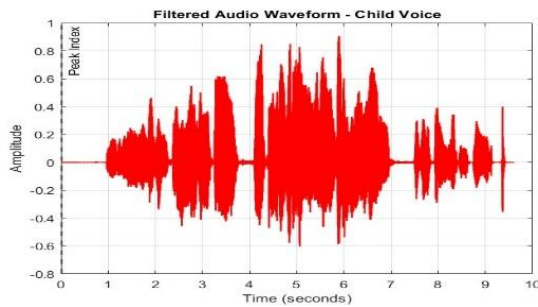
Gambar 8. Sampel suara Anak pertama

Index of the peak: 1293  
 Dominant Frequency: 272,7539 Hz  
 This is likely a child's voice.



Gambar 9. Sampel suara Anak kedua

Index of the peak: 1278  
 Dominant Frequency: 286,4365 Hz  
 This is likely a child's voice.



Gambar 10. Sampel suara Laki-laki ketiga

Index of the peak: 613  
 Dominant Frequency: 280.3898 Hz  
 This is likely a child's voice.

Dari sepuluh rekaman suara anak-anak yang dianalisis, algoritma FFT berhasil mengklasifikasikan semua sampel suara tersebut dengan benar sebagai suara anak-anak.

Tabel 1. Hasil Analisis Suara

Sampel Suara	Frekuensi Dominasi (Hz)	Indeks peak	Hasil Klasifikasi	Evaluasi	Sumber
Laki-laki 1	137.42	636	Laki-laki	Betul	WhatsApp
Laki-laki 2	108.08	523	Laki-laki	Betul	Rekaman langsung
Laki-laki 3	129.1992	597	Laki-laki	Betul	Rekaman langsung
Perempuan 1	195.8343	915	Perempuan	Betul	Rekaman langsung
Perempuan 2	307.4983	1479	Anak	Salah	WhatsApp
Perempuan 3	214.9035	1004	Perempuan	Betul	Rekaman langsung
Anak 1	272.7539	1293	Anak	Betul	WhatsApp
Anak 2	286.4365	1278	Anak	Betul	Rekaman langsung
Anak 3	280.3898	613	Anak	Betul	Rekaman langsung

Persentase keakurat Algoritma FFT dalam klasifikasi suara wicara;

$$\text{Persentase} = \frac{\text{Klasifikasi yang benar}}{\text{Jumlah sampel suara}} \times 100$$

$$\text{Persentase} = 29/30 \times 100$$

$$\text{Persentase} = 29/30 \times 100$$

$$\text{Persentase} = 96,66\%$$

Keakurat FFT dalam sampel yang diambil adalah 96,66%

## 5. Kesimpulan

Penelitian ini menegaskan peran penting algoritma Fast Fourier Transform (FFT) dalam menganalisis dan mengklasifikasikan sinyal wicara manusia berdasarkan jenis kelamin dan rentang usia. Berdasarkan hasil pengujian, FFT menunjukkan tingkat akurasi tinggi sebesar 96,66% dalam mengidentifikasi suara laki-laki, perempuan, dan anak-anak ketika digunakan pada rekaman langsung berkualitas tinggi yang diambil di lingkungan terkontrol dan kedap suara. Hal ini menunjukkan bahwa FFT sangat efektif dalam menampilkan perbedaan karakteristik spektrum frekuensi antar kategori suara.

Namun demikian, hasil juga memperlihatkan penurunan akurasi pada analisis rekaman terkompresi dari aplikasi WhatsApp, di mana satu sampel tidak teridentifikasi dengan benar. Hal ini disebabkan oleh proses kompresi audio yang dilakukan oleh aplikasi tersebut, yang dapat mengubah struktur spektral suara dengan menghilangkan komponen frekuensi tertentu atau menambahkan artefak digital, sehingga menyebabkan pergeseran puncak frekuensi dominan yang menjadi dasar analisis FFT. Kondisi ini menunjukkan bahwa meskipun FFT unggul dalam kecepatan dan ketepatan pada sinyal berkualitas tinggi, algoritma ini kurang adaptif terhadap degradasi sinyal akibat kompresi.

Untuk meningkatkan kinerja sistem, penelitian selanjutnya disarankan mengombinasikan FFT dengan metode Mel-Frequency Cepstral Coefficients (MFCC) yang lebih sesuai dengan persepsi pendengaran manusia, serta Support Vector Classification (SVC) atau model machine learning lainnya untuk meningkatkan kemampuan klasifikasi terutama pada kategori suara perempuan dan anak-anak. Integrasi pendekatan tersebut diharapkan dapat menghasilkan sistem identifikasi suara yang lebih akurat, adaptif, dan robust terhadap variasi sumber maupun format rekaman.

## Ucapan Terima Kasih

Penulis mengucapkan terima kasih kepada Dosen Mata Kuliah Isyarat dan Sistem Ir. Joko Hariyono, S.T., M.Eng., Ph.D. atas bimbingan dan arahnya selama proses penyusunan artikel ini. Penghargaan juga disampaikan kepada semua pihak yang telah memberikan dukungan pendanaan dan fasilitas, sehingga penelitian ini dapat diselesaikan dengan baik.

## Referensi

- [1]. J. C. Cavalcanti, M. Englert, M. Oliveira, and A. C. Constantini, "Microphone and Audio Compression Effects on Acoustic Voice Analysis: A Pilot Study," *Journal of Voice*, vol. 37, no. 2, pp. 162–172, Mar. 2023, doi: 10.1016/j.jvoice.2020.12.005.
- [2]. M. F. , Erinsyah, V. , Karenina, and D. S. Wibowo, "Klasifikasi Rentang Usia Dan Gender Dengan Deteksi Suara Menggunakan Metode Deep Learning Algoritma CNN (Convolutional Neural Network)," *Komputika: Jurnal Sistem Komputer*, vol. 12, no. 2, 2023.
- [3]. S. Y. , Hartono, H. B. , Ashad, and B. P. Asmara, "Analisis Sinyal Wicara (Speech) Perekaman Dalam Ruang Kedap Suara Dengan Algoritma Fast Fourier Transform (FFT)," *Prosiding Seminar Ilmiah Sistem Informasi dan Teknologi Informasi*, vol. 11, no. 1, 2022.
- [4]. A. S. , Irtawaty, M. , Ulfah, and R. S. Fathmala, "Pengembangan Aplikasi pembeda suara Laki-laki dan perempuan berdasarkan gender, Range usia, kelas Frekuensinya berbasis FFT dan K-Means," *Jurnal ECOTIPE*, vol. 9, no. 1, 2022.
- [5]. Andreas, "Klasifikasi Suara Manusia ke Dalam Sopran, Mezzo Sopran, Alto, Tenor, Bariton, Bass dengan Self Organizing Map," *Jurnal Informatika*, vol. 11, no. 1, 2015.
- [6]. S. B. Bhaskoro, "Aplikasi Pengenalan Gender Menggunakan Suara," *Seminar Nasional Aplikasi Teknologi Informasi (SNATI), Jogjakarta*, 2012.
- [7]. J. F. Gunawan, *Pengolahan Sinyal Digital dengan Pemrograman Matlab*, 1st Edition. Jogjakarta: Graha Ilmu, 2012.
- [8]. L. S. , Kristina, G. F. , Fitriana, and A. Prasetyadi, "Pemisahan Suara Manusia Berdasarkan Jenis Kelamin Menggunakan Fast Fourier Transform (FFT)," *Jurnal Teknik Informatika dan Sistem Informasi*, vol. 7, no. 3, 2020.
- [9]. D. T. Kusuma, "Fast Fourier Transform (FFT) Dalam Transformasi Sinyal Frekuensi Suara Sebagai Upaya Perolehan Average Energy (AE) Musik," *PETIR : Jurnal Pengkajian dan Penerapan Teknik Informatika*, vol. 14, no. 1, 2021.
- [10]. M. M. Effendi, "Pengenalan Pengaruh Suara Konsonan Terhadap Vokal Dengan Nilai Koefisien Korelasi," *Journal Pelita Teknol*, vol. 12, no. 4, pp. 35–41, 2017.
- [11]. M. I. , Siddiq, I. D. , Karyatanti, B. Y. , Dewantara, A. S. , Syahputrib, and M. D. Bahtiar, "Monitoring Bearing Motor Induksi Pada Variasi Noise Menggunakan Pengolahan Sinyal FFT," *ALINIER JURNAL*, vol. 5, no. 1, 2024.
- [12]. J. W. , Cooley and J. W. Tukey, "An algorithm for the machine calculation of complex Fourier series," *Math Comput*, vol. 19, no. 90, pp. 297–301, 1965.
- [13]. J. G. , Proakis and D. G. Manolakis, *Digital Signal Processing: Principles, Algorithms, and Applications*. Pearson Prentice Hall, 2007.
- [14]. Nilu Singh, "Speaker Recognition and Fast Fourier Transform," 2015.
- [15]. R. , Togneri and D. Püllella, "An Overview of Speaker Identification: Accuracy, Robustness, and Security," *IEEE Circuits and Systems Magazine*, vol. 11, no. 2, pp. 23–25, 2011.
- [16]. Al-Qaderi M., M. G. , Zamil, and M. F. Al-Mistarihi, "Two-Level Speaker Identification System Using Frequency and Time Domain Features," *Int J Speech Technol*, vol. 23, no. 2, pp. 189–201, 2020.
- [17]. R. Fuchs, "The Effects of MP3 Compression on Acoustic Measurements of Speech," *Proceedings of the 17th Annual Conference of the International Speech Communication Association (INTERSPEECH)*, 2016.
- [18]. O. , Niebuhr, C. , Kohrs, and J. Neitsch, "Popular Speech Compression Codecs Produce Codec-Specific Distortion Effects on Prosody," *Front Commun (Lausanne)*, vol. 8, pp. 1–4, 2023.
- [19]. J. Oreskovic, "Impact of Audio Data Compression on Feature Extraction and Acoustic Biomarkers," *Applied Acoustics*, pp. 109–117, 2024.