



## EKASAKTI JURNAL PENELITIAN & PENGABDIAN (EJPP)



Lisensi: <https://creativecommons.org/licenses/by/4.0/>

Diterima: 22 Desember 2023, Diperbaiki: 29 Maret 2024, Diterbitkan: 26 Mei 2024

### Pengenalan Suara Manusia Menggunakan Jaringan Syaraf Tiruan Dengan Metode Linear Predictive Coding dan Fast Fourier Transform

Nirwan Sinuhaji<sup>1</sup>, Benar<sup>2\*</sup>, Romulo P. Aritonang<sup>3</sup>,

<sup>1), 3)</sup> ITB Indonesia, <sup>2)</sup> Politeknik Negeri Medan, Medan Sumatera Utara.

\*Corresponding Author: [nirwansinuhaji@yahoo.co.id](mailto:nirwansinuhaji@yahoo.co.id)

**Abstract:** The voice of each person is basically different from one another, therefore voice can be used in digital technology or biometric methods used by humans to identify or certify a person's identity based on their physical characteristics or behavior and this technology is widely used in the fields of security, health and finance. The voice identification process is the process of finding out the identity of a person from the voice they speak. The voice verification process is the process of matching the inputted voice with the voice in the database. This study will use an artificial neural network using the linear predictive coding and fast methods as an initial processor. The voice signal obtained from the recording results is processed by a computer with the FFT algorithm so that a digital signal is obtained, then parsed, the results of the feature extraction with the FFT method then become the input of the LPC method and then implemented into the Artificial Neural Network.

**Keywords:** Artificial neural network, Linear Predictive Coding, Fast Fourier Transform

**Abstrak:** Suara pada setiap orang pada dasarnya memiliki perbedaan satu dengan yang lainnya oleh sebab itu suara dapat digunakan pada metode teknologi digital atau biometrik yang digunakan manusia untuk mengidentifikasi atau menyertifikasi identitas seseorang berdasarkan karakteristik fisik atau perilaku mereka dan teknologi ini banyak digunakan dalam bidang keamanan, kesehatan dan keuangan. Proses identifikasi suara yaitu proses untuk mengetahui identitas dari seseorang dari suara yang diucapkannya. Proses verifikasi suara yaitu proses untuk mencocokkan suara yang diinputkan dengan suara yang ada di database. Penelitian ini akan menggunakan jaringan syaraf tiruan menggunakan metode linear predictive coding dan fast sebagai pemroses awal. Sinyal suara yang diperoleh dari hasil perekaman, diolah oleh komputer dengan algoritma FFT sehingga diperoleh sinyal digital, kemudian diuraikan, hasil ekstraksi ciri dengan metode FFT tersebut kemudian menjadi input metode LPC dan kemudian diimplementasikan ke Jaringan Syaraf Tiruan

**Kata Kunci:** Jaringan syaraf tiruan, Linear Predictive Coding, Fast Fourier Transform

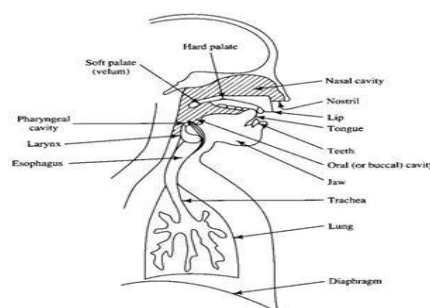
#### PENDAHULUAN

Perkembangan media komunikasi tidak lepas dari tiga hal pokok yaitu suara, penglihatan, dan sentuhan sehingga dalam mengembangkan teknologi khususnya kecerdasan buatan mengacu pada tiga hal tersebut. Banyak kemudahan yang ditawarkan untuk kepentingan interaksi manusia dan komputer, misalnya dalam pengenalan suara. Manusia bisa mengenali kondisi seseorang dengan mendengar suaranya, sebagai contoh jenis kelamin, identitas pembicara, aksen, gaya bicara, emosi, dan kondisi kesehatan dari si pembicara.

(Faradiba,2017 ). Teknik jaringan saraf tiruan telah banyak dimanfaatkan pada berbagai bidang utamanya pada sistem pengenalan pola seperti citra, suara, time series prediction, dan lain-lain. Pada penelitian ini dibuat suatu sistem memanfaatkan Jaringan Saraf Tiruan metode Propagasi Balik (Back Propagation) untuk pengenalan suara. Sistem ini dapat diterapkan pada pemberian perintah komputer, voice dialling, dan lain lain. Model jaringan saraf back propagation digunakan di sini bersama dengan metode Linear Predictive Coding (LPC) dan Fast Fourier Transform (FFT) yang dipakai sebagai pemroses awal.

Sinyal suara yang diperoleh dari hasil perekaman, diolah oleh komputer dengan algoritma FFT sehingga diperoleh sinyal digital, kemudian diuraikan, hasil ekstraksi ciri dengan metode FFT tersebut kemudian menjadi input metode LPC dan kemudian diimplementasikan ke Jaringan Syaraf Tiruan.

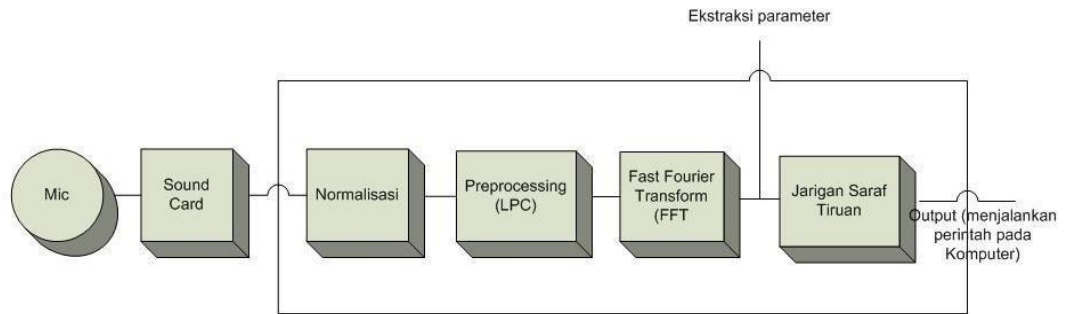
Mekanisme getaran suara sebenarnya sangat kompleks. Ketika celah suara menegang dan tekanan udara meningkat dari paru-paru, periode membuka dan menutupnya menjadi pendek dan frekuensi (pitch) sumber suara menjadi tinggi. Periode membuka dan menutup ini disebut getaran celah suara. Sebaliknya, kondisi tekanan udara yang rendah menghasilkan suara frekuensi yang rendah. Sumber suara terdiri atas komponen fundamental dan harmonik yang dimodifikasi oleh jalur vocal untuk menghasilkan suara, seperti dalam menghasilkan bunyi vokal (a) dan (o). Ucapan manusia dihasilkan oleh suatu sistem produksi ucapan yang dibentuk oleh alat alat ucap manusia. Proses tersebut dimulai dengan formulasi pesan dalam otak pembicara. Pesan tersebut akan diubah menjadi perintah-perintah yang diberikan kepada alat-alat ucap manusia, sehingga akhirnya dihasilkan ucapan yang sesuai dengan pesan yang ingin diucapkan dapat dilihat seperti gambar organ pembentikan suara manusia dibawah ini.



Gambar 2.1 Organ pembentuk suara manusia

Proses produksi **suara** pada manusia dapat dibagi menjadi tiga buah proses fisiologis, yaitu : pembentukan aliran udara dari paru-paru, perubahan aliran udara dari paru-paru menjadi suara, baik *voiced*, maupun *unvoiced* yang dikenal dengan istilah *phonation*, dan artikulasi yaitu proses modulasi/ pengaturan suara menjadi bunyi yang spesifik. Organ tubuh ini dapat dikelompokkan menjadi tiga bagian utama, yaitu : *vocal tract* (berawal di awal bukaan pita suara atau glottis, dan berakhir di bibir), *nasal tract* (dari velum sampai nostril), dan *source generator* (terdiri dari paru-paru, tenggorokan, dan *larynx*).

Secara garis besar, cara kerja sistem pengenalan suara ini adalah mula-mula sinyal suara manusia yang diterima dengan menggunakan *microphone* (sinyal analog) diubah sehingga menjadi sinyal digital dengan bantuan *sound card* pada PC. Terlihat pada gambar dibawah.



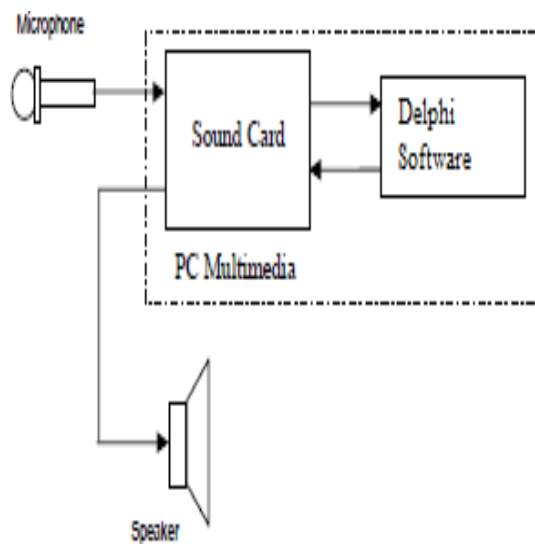
**Gambar 2.** Diagram pengenalan suara

## METODE PENELITIAN

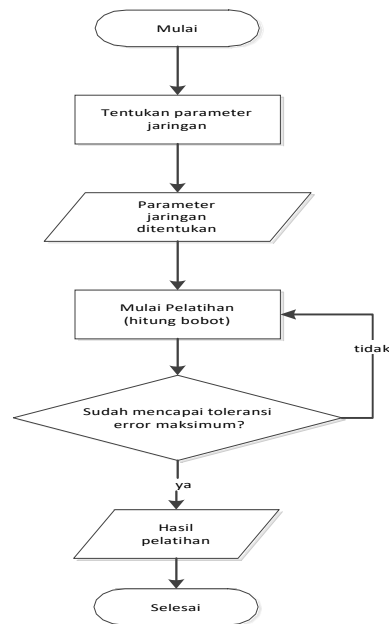
Dalam metode penelitian ini desain dan perancangan sistem metode *Neural Network* dengan algoritma *Backpropagation* untuk pengenalan suara. Desain dan perancangan sistem ini meliputi analisis sistem, perancangan sistem, dan desain antarmuka.

### Analisis Kebutuhan Sistem

Metode analisis ini merupakan tahap melakukan analisa pada sistem yang akan di bangun. Kebutuhan sistem antara lain sebuah perangkat komputer multimedia dan perangkat lunak pendukung penelitian.



**Gambar 3.** Gambaran umum kebutuhan sistem  
Flowchart jaringan syaraf tiruan dapat dilihat seperti gambar dibawah

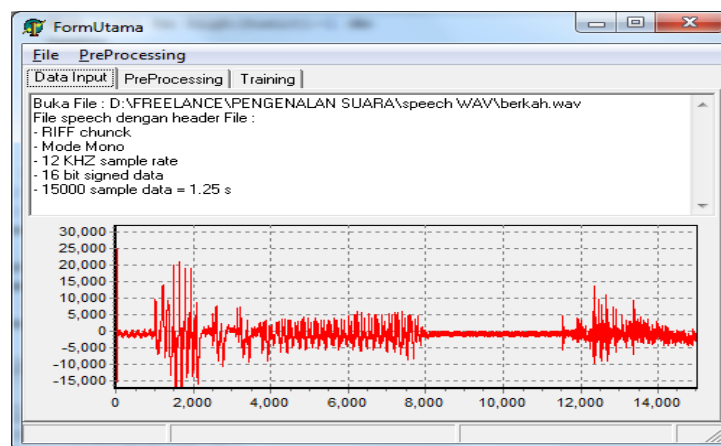


Gambar 4. Flowchart jaringan syaraf tiruan

## HASIL DAN PEMBAHASAN

**EKSTRAKSI FILE WAV** File WAV yang dibuka awalnya akan memuat parameter yang tersimpan pada *header chunk*, berikut parameter yang dimuat: *wFormatTag* berisi kata “RIFF chunk” yang menunjukkan file tersebut berupa file berformat WAV. *nChannels* menunjukkan mode stereo atau mono. *nSamplesPerSec* menunjukkan *Sample Rate*. *nBlockAlign* menunjukkan jumlah bit yang digunakan dalam satu *sample*.

Nilai ini juga ditampung ke dalam memori untuk melakukan tahapan selanjutnya yaitu *preprocessing* suara. Contoh hasil membuka file WAV terdapat pada gambar dibawah.

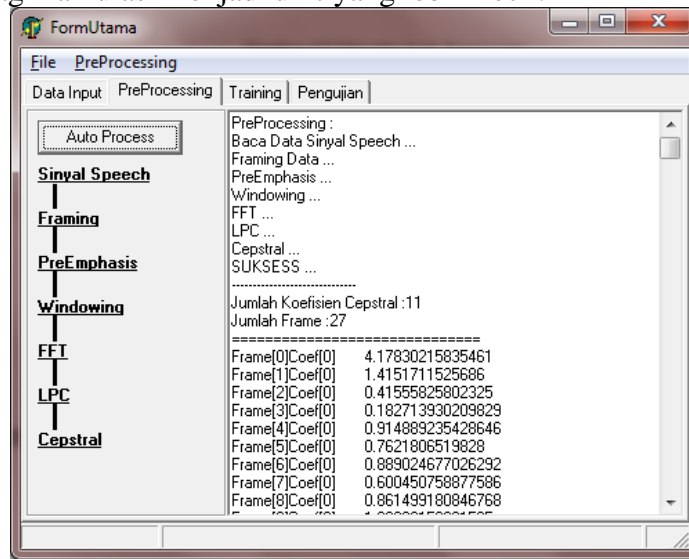


Gambar 4.1 Hasil membuka file WAV

### Propocessing Data

Didalam Propocessing Data terdapat beberapa tahapan yaitu Framing sinyal,

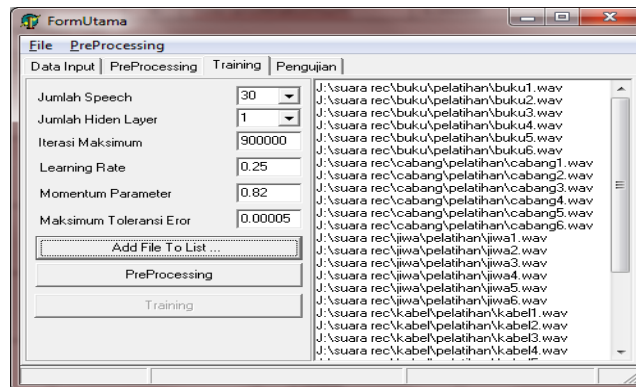
Linier Prediktif Coding (LPC), dan Fast Fourier Transform (FFT). Propocessing Data berfungsi melakukan pembagian sinyal ke dalam beberapa *Frame* dan *Cofisien Cepstral* dengan tujuan membagi kalkulasi menjadi unit yang lebih kecil.



Gambar 4. Processing data

**Training Data**

Pelatihan pada Jaringan Syaraf Tiruan *Backpropagation* ini bersifat *Supervised Learning* (Belajar Dengan Pengawasan). Parameter dari perhitungan JST ditentukan terlebih dahulu sesuai dengan keinginan pengguna. Disini data yang diinput juga harus diperkenalkan namanya. Adapaun parameter-parameter dari JST *Backpropagation* yang digunakan adalah sebagai berikut



Gambar 5 Jendela pelatihan data

**Pengujian Dan Pelatihan**

Pada pengujian pertama penulis memulai dengan melakukan pelatihan menggunakan jumlah *hidden layer* 1 dan *jumlah unit hidden* 25. Adapun parameter JST yang digunakan sebagai berikut.

**Tabel 1. Parameter pengujian pertama**

No.	Parameter JST	Nilai
1.	Jumlah speech	35
2.	Jumlah hidden layer	1

3.	Jumlah unit hidden layer	25
4.	Iterasi maksimum	900000
5.	Learning rate	0.25
6.	Momentum parameter	0.60
7.	Maksimum toleransi error	0.00005

Pengujian dilakukan terhadap sampel yang telah dilatihkan sebelumnya. Hasil pengenalan terhadap suara yang telah dilatih dapat dilihat pada tabel 2.

**Tabel 2. Persentase hasil pengujian untuk data dilatih pada pengujian pertama**

No	Kata	Perhitungan	Tingkat keberhasilan
1	Buku	$\frac{\sum \square\square\square\square\square}{\sum \square\square\square\square\square} \square 100\% = \frac{7}{7} \square 100\%$	100%
2	Cabang	$\frac{\sum \square\square\square\square\square}{\sum \square\square\square\square\square} \square 100\% = \frac{7}{7} \square 100\%$	100%
3	Jiwa	$\frac{\sum \square\square\square\square\square}{\sum \square\square\square\square\square} \square 100\% = \frac{7}{7} \square 100\%$	100%
4	Kabel	$\frac{\sum \square\square\square\square\square}{\sum \square\square\square\square\square} \square 100\% = \frac{7}{7} \square 100\%$	100%
5	Makan	$\frac{\sum \square\square\square\square\square}{\sum \square\square\square\square\square} \square 100\% = \frac{7}{7} \square 100\%$	100%
6.	Seluruh data	$\frac{\sum \square\square\square\square\square}{\sum \square\square\square\square\square} \square 100\% = \frac{35}{35} \square 100\%$	100%

Untuk pengujian yang dilakukan terhadap sampel yang belum dilatihkan sebelumnya, hasil pengenalan dapat dilihat pada tabel 3.

**Tabel 3. Persentase hasil pengujian untuk data luar pelatihan pada pengujian pertama**

No	Kata	Perhitungan	Tingkat keberhasilan
1	Buku	$\frac{\sum \square\square\square\square\square}{\sum \square\square\square\square\square} \square 100\% = \frac{3}{4} \square 100\%$	75%
2	Cabang	$\frac{\sum \square\square\square\square\square}{\sum \square\square\square\square\square} \square 100\% = \frac{2}{4} \square 100\%$	50%
3	Jiwa	$\frac{\sum \square\square\square\square\square}{\sum \square\square\square\square\square} \square 100\% = \frac{1}{4} \square 100\%$	25%
4	Kabel	$\frac{\sum \square\square\square\square\square}{\sum \square\square\square\square\square} \square 100\% = \frac{2}{4} \square 100\%$	50%

5	Makan	$\frac{\sum \square\square\square\square\square}{100\%} = \frac{3}{4} = 75\%$	75%
6.	Seluruh data	$\frac{\sum \square\square\square\square\square}{100\%} = \frac{11}{20} = 55\%$	55%

Sistem memiliki tingkat keberhasilan sebesar 100% untuk mengenali kata yang telah dilatih. Sedangkan pengujian terhadap kata diluar pelatihan yang ditunjukkan oleh tabel 3. memiliki tingkat keberhasilan lebih kecil yaitu sebesar 55%.

## KESIMPULAN

### Simpulan

Kesimpulan dari penelitian ini menunjukkan bahwa Jaringan Syaraf Tiruan dengan metode Back Propagation merupakan alat yang efektif dalam pengenalan suara. Pengujian pada sampel yang telah dilatih dengan sistem ini berhasil mencapai tingkat pengenalan hingga 100%, menandakan keakuratan yang sangat tinggi dalam kondisi pelatihan. Namun, performa sistem ketika diuji dengan sampel yang tidak termasuk dalam pelatihan menunjukkan variasi tingkat pengenalan yang lebih rendah, berkisar antara 30% hingga 55%. Temuan ini mengindikasikan adanya tantangan dalam generalisasi model terhadap data baru. Selain itu, tingkat pengenalan tertinggi tercatat saat menggunakan satu hidden layer dengan dua puluh lima unit, menunjukkan bahwa konfigurasi ini paling optimal untuk tugas pengenalan suara dalam penelitian ini.

### Saran

Beberapa hal yang dapat diperhatikan untuk penelitian ini adalah sebagai berikut.

1. Penelitian dapat dilanjutkan secara waktu nyata (*real time*), yang dapat dilakukan dengan cara menghubungkan alat input suara (*microphone*) dengan perangkat lunak secara langsung.
2. Perlu dilakukan proses pra pengolahan (*preprocessing*) yang lebih kompleks pada suara sehingga ciri-ciri suara yang lebih detail.

## REFERENSI

- Ahmad Arifin (2018) Aplikasi jaringan syaraf tian metode perceptron pada pengenalan suara
- Andi Sri Irtawat,dkk (2019)Implementasi Metode Fast Fourier Transform (FFT) Dalam Mengklasifikasikan Suara Pria dan Wanita di Laboratorium Jurusan Teknik Elektro Politeknik Negeri Balikpapan
- Arif Syaifuddin, Suryono (2014) FAST FOURIER TRANSFORM (FFT) UNTUK ANALISIS SINYAL SUARADOPPLER ULTRASONIK
- Dine Tiara Kusuma (2021) Fast Fourier Transform (FFT) Dalam Transformasi Sinyal Frekuensi Suara Sebagai Upaya Perolehan Average Energy (AE) Musik
- FaradibaJuli (2017 ) Pengenalan Pola Sinyal Suara Manusia Menggunakan Metode Back Propagation Neural Network

- Harun Sujad,dkk (2017)SISTEM PENGOLAHAN SUARA MENGGUNAKAN ALGORITMA FFT (FAST FOURIER TRANSFORM)
- Irmawan S.Si,M.T (2014) Pengenalan kata dengan metode linear predictive coding dan jaringan syaraf tiruan pada mobile robot jaringgan
- Irsam Rahmat Yusuf, - (2018PEMBANGUNAN MODEL PENGENALAN SUARA MENGGUNAKAN METODE DEEP LEARNING
- Lisa Setiawati (dkk) (2017) Penerapa Algoritma genetika untuk pengenalan pola suara  
Mada Sanjaya W.S dkk, (2014) IMPLEMENTASI PENGENALAN POLA SUARA MENGGUNAKAN MEL-FREQUENCY CEPSTRUM COEFFICIENTS (MFCC) DAN ADAPTIVE NEURO-FUZZY INFERENCE SYSTEM (ANFIS) SEBAGAI KONTROL LAMPU OTOMATIS
- M. Ma'ruf Idris (2010) pengenalan suara manusia dengan menggunakan jaringan syaraf tiruan model propagasi