

**SKRIPSI**  
**IMPLEMENTASI MFCC & CNN PADA *GENDER VOICE***  
***RECOGNITION***



**HARDI TRI NUGROHO**

**NIM : 195410093**

**PROGRAM STUDI INFORMATIKA**  
**PROGRAM SARJANA**  
**FAKULTAS TEKNOLOGI INFORMASI**  
**UNIVERSITAS TEKNOLOGI DIGITAL INDONESIA**  
**YOGYAKARTA**  
**2024**

**SKRIPSI**  
**IMPLEMENTASI MFCC & CNN PADA *GENDER VOICE***  
***RECOGNITION***

Diajukan sebagai salah satu syarat untuk menyelesaikan studi



**Program Sarjana**  
**Program Studi Informatika**  
**Fakultas Teknologi Informasi**  
**Universitas Teknologi Digital Indonesia**  
**Yogyakarta**

**Disusun Oleh**  
**HARDI TRI NUGROHO**  
**NIM : 195410093**

**PROGRAM STUDI INFORMATIKA**  
**PROGRAM SARJANA**  
**FAKULTAS TEKNOLOGI INFORMASI**  
**UNIVERSITAS TEKNOLOGI DIGITAL INDONESIA**  
**YOGYAKARTA**  
**2024**

**HALAMAN PERSETUJUAN**  
**UJIAN SKRIPSI**

Judul : Implementasi MFCC & CNN Pada *Gender Voice Recognition*  
Nama : Hardi Tri Nugroho  
NIM : 195410093  
Program Studi : Informatika  
Program : Sarjana  
Semester : Genap  
Tahun Akademik : 2023/2024

Telah diperiksa dan disetujui untuk diujikan di hadapan Dewan Penguji Skripsi

Yogyakarta, 5 September 2024

Dosen Pembimbing,



Deborah Kurniawati, S.Kom., M.Cs.

NIDN : 0511107301

# HALAMAN PENGESAHAN

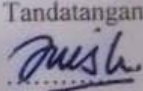

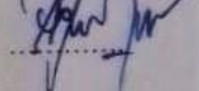
## SKRIPSI

IMPLEMENTASI MFCC & CNN PADA *GENDER VOICE RECOGNITION*

Telah dipertahankan di depan Dewan Penguji Skripsi dan dinyatakan diterima untuk memenuhi sebagian persyaratan guna memperoleh Gelar

Sarjana Komputer  
Program Studi Informatika  
Fakultas Teknologi Informasi  
Universitas Teknologi Digital Indonesia  
Yogyakarta

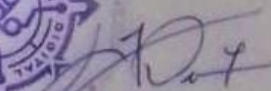
Yogyakarta, 5 September 2024

Dewan Penguji	NIDN	Tandatangan
1. Ariesta Damayanti, S. Kom., M. Cs.	0020047801	
2. Maria Mediatrix Sebatubun., S.Kom., M.Eng.	0514089101	
3. Deborah Kurniawati, S.Kom., M.Cs.	0511107301	



Mengetahui

Ketua Program Studi Informatika


  
Dini Fakta Sari, S.T., M.T.

NIDN : 0507108401

## **PERNYATAAN KEASLIAN SKRIPSI**

Dengan ini saya menyatakan bahwa naskah skripsi ini belum pernah diajukan untuk memperoleh gelar Sarjana Komputer di suatu Perguruan Tinggi, dan sepanjang pengetahuan saya tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara sah diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Yogyakarta, 6 September 2024

  
Hardi Tri Nugroho

NIM: 195410093

## HALAMAN PERSEMBAHAN

Dengan mengucap rasa syukur kehadiran Allah Subhanahu wa ta'ala Alhamdulillah Skripsi ini penulis persembahkan untuk :

- ☒ Ibu tersayang dan tercinta Sumarsihatun, A.Ma., yang selalu berdoa, sabar dan percaya kepada anak-anaknya dalam hal apapun
- ☒ Bapak tercinta Alm.Sailan yang selalu mengajarkan kesabaran, tidak menyerah menghadapi rintangan yang ada di depan mata.
- ☒ Willis Ayu Pramukti, S.Kep.Ners, dan Nafi`u Rizeqi Hidayat yang selalu mensupport adik-adiknya untuk berkembang dan berusaha meraih masa depan yang cerah.
- ☒ Ery Sigit Wahyudi, A.Md. yang selalu peduli keadaan disaat jauh dari siapapun.
- ☒ Khilda Hasni Ijati, S.pd., M.Sc. yang selalu menemani, mensupport dan percaya untuk menyelesaikan penelitian skripsi ini.
- ☒ Pedro dan Mini kucing yang menemani ketika stress dan overthinking.
- ☒ Kepada diri sendiri yang mau berusaha dan berjuang dalam menyelesaikan penelitian skripsi.

Segenap pihak yang tidak dapat disebutkan satu persatu. Terima kasih atas doa, motivasi dan bantuannya.

## HALAMAN MOTTO

"Just do what's in front of you; we never know what the future will be like."

-Hardi-

## KATA PENGANTAR

Segala puji syukur kami panjatkan kehadirat Tuhan Yang Maha Kuasa, karena oleh anugerah-Nya, kemurahan dan kasih setiaNya yang besar akhirnya penulis dapat menyelesaikan penulisan skripsi dengan judul: “Implementasi MFCC & CNN Pada *Gender Voice Recognition*”.

Penulis menyadari sepenuhnya bahwa skripsi ini masih jauh dari kesempurnaan karena menyadari segala keterbatasan yang ada. Untuk itu demi sempurnanya skripsi ini, penulis sangat membutuhkan dukungan dan sumbangsih pikiran yang berupa kritik dan saran yang bersifat membangun.

Skripsi ini dipersembahkan kepada kedua orang tua (Bapak Alm.Sailan dan Ibu Sumarsihatun) yang telah tulus ikhlas memberikan kasih sayang, cinta, doa, perhatian, dukungan moral dan materil yang telah diberikan selama ini.

Penyusunan skripsi ini tidak lepas dari bantuan berbagai pihak, karena itu penulis menyampaikan ucapan terima kasih kepada:

1. Ibu Sri Redjeki, S. Si., M. Kom., Ph. D. Sebagai Rektor Universitas Teknologi Digital Indonesia.
2. Ibu Dini Fakta Sari, S.T., M.T. Selaku Ketua Program Studi Informatika.
3. Ibu Deborah Kurniawati, S.Kom., M.Cs. Selaku pembimbing yang telah meluangkan waktunya untuk membimbing penulis.
4. Ibu Ariesta Damayanti, S. Kom., M. Cs., dan Ibu Maria Mediatrrix Sebatubun, S.Kom., M.Eng. selaku penguji. Terima kasih atas waktu, serta masukan dan arahnya.
5. Ibu Deborah Kurniawati, S.Kom., M.Cs. Selaku penasehat akademik yang bersedia meluangkan waktunya untuk membimbing dan mengarahkan penulis.
6. Seluruh dosen di Universitas Teknologi Digital Indonesia yang telah memberikan ilmunya kepada penulis selama perkuliahan di kampus ini.
7. Seluruh civitas akademika Universitas Teknologi Digital Indonesia yang telah banyak memberi bantuan dan dukungan selama penulis menempuh



studi dan menyelesaikan skripsi di Universitas Teknologi Digital Indonesia.

Kiranya skripsi ini dapat memberikan manfaat dan masukan bagi pembaca.  
Terima Kasih.

Yogyakarta, September 2024

Hardi Tri Nugroho

## DAFTAR ISI

	Hal
SKRIPSI.....	ii
HALAMAN PERSETUJUAN.....	iii
HALAMAN PENGESAHAN.....	iv
PERNYATAAN KEASLIAN SKRIPSI.....	v
HALAMAN PERSEMBAHAN .....	vi
HALAMAN MOTTO .....	vii
KATA PENGANTAR .....	viii
DAFTAR ISI.....	x
DAFTAR GAMBAR .....	xiii
DAFTAR TABEL.....	xv
INTISARI.....	xvi
ABSTRACT .....	xvii
BAB I PENDAHULUAN .....	1
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah .....	3
1.3 Ruang Lingkup Penelitian .....	4
1.4 Tujuan Penelitian.....	4
1.5 Manfaat Penelitian.....	5
BAB II TINJAUAN PUSTAKA DAN LANDASAN TEORI .....	6
2.1 Tujuan Pustaka .....	6
2.2 <i>Voice Recognition</i> .....	10
2.3 Sinyal Suara Manusia .....	11
2.4 <i>Mel Frequency Cepstral Coefficient (MFCC)</i> .....	11
2.4.1 <i>Pre-Emphasized</i> .....	12
2.4.2 <i>Framing</i> .....	13
2.4.3 <i>Windowing</i> .....	14
2.4.4 <i>Fast Fourier Transform (FFT)</i> .....	16
2.4.5 <i>Mel Filter Bank</i> .....	17
2.4.6 <i>Discrete Cosine Transform (DCT)</i> .....	18

2.5 <i>Normalization</i> .....	19
2.6 <i>Padding</i> .....	20
2.7 <i>Convolutional Neural Network (CNN)</i> .....	21
2.8 <i>K-fold Cross Validation</i> .....	22
2.9 <i>Python</i> .....	23
2.10 <i>VoxCeleb</i> .....	24
2.11 <i>Flask</i> .....	24
<b>BAB III METODE PENELITIAN</b> .....	<b>25</b>
3.1 <i>Prosedur Penelitian</i> .....	25
3.2 <i>Bahan dan Data</i> .....	26
3.3 <i>Perangkat</i> .....	26
3.3.1 <i>Kebutuhan Input</i> .....	27
3.3.2 <i>Kebutuhan Proses</i> .....	27
3.3.3 <i>Kebutuhan Output</i> .....	27
3.3.4 <i>Kebutuhan Perangkat Keras</i> .....	27
3.3.5 <i>Kebutuhan Perangkat Lunak</i> .....	28
3.4 <i>Analisis Rancangan Sistem</i> .....	28
3.5 <i>Pre-Processing</i> .....	30
3.5.1 <i>Pre-Emphasized</i> .....	30
3.5.2 <i>Framing</i> .....	30
3.5.3 <i>Windowing</i> .....	31
3.5.4 <i>Fast Fourier Transform</i> .....	32
3.5.5 <i>Mel Filter Bank</i> .....	34
3.5.6 <i>Discrete Cosine Transform</i> .....	37
3.5.7 <i>Normalization</i> .....	39
3.5.8 <i>Padding</i> .....	40
3.6 <i>Processing</i> .....	41
3.6.1 <i>Arsitektur Model</i> .....	41
3.6.2 <i>Pembuatan Model</i> .....	44
3.6.3 <i>Evaluasi Model</i> .....	46
3.7 <i>Implementation Model</i> .....	47

3.7.1	<i>Flowchart</i> .....	48
3.7.2	User Interface .....	49
BAB IV HASIL DAN PEMBAHASAN .....		50
4.1	Hasil Penelitian.....	50
4.1.1	Persiapan Data.....	50
4.1.2	Ekstraksi Fitur MFCC .....	51
4.1.3	Klasifikasi CNN .....	70
4.1.4	Evaluasi Model.....	75
4.1.5	<i>WebApp Gender Voice Recognition</i> .....	84
BAB V KESIMPULAN .....		88
5.1	Kesimpulan.....	88
5.2	Saran .....	89
DAFTAR PUSTAKA .....		90
LAMPIRAN .....		94

## DAFTAR GAMBAR

	Hal
Gambar 2.1 Tahapan MFCC .....	12
Gambar 2.2 <i>Signal Pre-Emphasized</i> .....	13
Gambar 2.3 <i>Framing Signal Process</i> .....	14
Gambar 2.4 <i>Signal Windowing</i> .....	15
Gambar 2.5 <i>Signal Hamming Window</i> .....	15
Gambar 2.6 <i>Mel Scale Filter Bank</i> .....	17
Gambar 2. 7 <i>Process CNN</i> .....	22
Gambar 2. 8 <i>K-fold cross validation</i> .....	23
Gambar 3.1 Rancangan Sistem .....	25
Gambar 3.2 Rancangan Kebutuhan Sistem.....	29
Gambar 3.3 Arsitektur CNN .....	43
Gambar 3.4 Alur Pembuatan Model.....	45
Gambar 3.5 Proses Evaluasi Model .....	46
Gambar 3.6 <i>Flowchart WebApp</i> .....	48
Gambar 3.7 <i>Interface WebApp</i> .....	49
Gambar 4.1 <i>Convert Format Audio Code</i> .....	51
Gambar 4.2 Visualisasi <i>Wavefrom Original Signal</i> .....	52
Gambar 4.3 <i>Python Pre-Emphasized Code</i> .....	53
Gambar 4.4 Visualisasi <i>Wavefrom Signal Pre-Emphasized</i> .....	54
Gambar 4.5 <i>Python Framing Code</i> .....	55
Gambar 4.6 Visualisasi <i>Wavefrom Signal Framing</i> .....	56

Gambar 4.7 <i>Python Windowing Code</i> .....	57
Gambar 4.8 <i>Visualisasi Waveform Signal Windowing</i> .....	57
Gambar 4.9 <i>Python FFT Code</i> .....	58
Gambar 4.10 <i>Visualisasi Waveform Signal FFT</i> .....	59
Gambar 4.11 <i>Python Power Spectrum Code</i> .....	60
Gambar 4.12 <i>Visualisasi Waveform Signal Power Spectrum</i> .....	61
Gambar 4.13 <i>Python Mel-Filter Bank Code</i> .....	63
Gambar 4.14 <i>Visualisasi Triangles &amp; Spectrogram Signal Mel-Filter Bank</i> .....	63
Gambar 4.15 <i>Python DCT Code</i> .....	65
Gambar 4.16 <i>Visualisasi Spectrogram Signal DCT</i> .....	65
Gambar 4.17 <i>Python Normalization Code</i> .....	67
Gambar 4.18 <i>Visualisasi Spectrogram Signal Normalization</i> .....	67
Gambar 4.19 <i>Python Padding Code</i> .....	68
Gambar 4.20 <i>Visualisasi Spectrogram Signal Padding</i> .....	69
Gambar 4.21 <i>Python LoadData Code</i> .....	71
Gambar 4.22 <i>Python Build CNN Model</i> .....	72
Gambar 4.23 <i>Python Training Code</i> .....	74
Gambar 4.24 <i>Interface Before Process</i> .....	84
Gambar 4.25 <i>Interface After Process I</i> .....	85
Gambar 4.26 <i>Interface After Process II</i> .....	85
Gambar 4.27 <i>Interface After Process III</i> .....	86
Gambar 4.28 <i>Interface After Process IV</i> .....	86
Gambar 4.29 <i>Interface After Process V</i> .....	87

## DAFTAR TABEL

	Hal
Tabel 2.1 Tabel Tinjauan Pustaka.....	8
Tabel 3.1 <i>Hyperparameter</i> model .....	41
Tabel 4.1 Tabel Data <i>VoxCeleb</i> .....	50
Tabel 4.2 Tabel Parameter Ekstraksi Fitur .....	51
Tabel 4.3 <i>Array Original Signal</i> .....	53
Tabel 4. 4 <i>Implementation Pre-emphasized</i> .....	54
Tabel 4.5 <i>Implementation Framing</i> .....	56
Tabel 4.6 <i>Implementation Windowing</i> .....	58
Tabel 4.7 <i>Implementation FFT</i> .....	59
Tabel 4.8 <i>Implementation Power Spectrum</i> .....	61
Tabel 4.9 <i>Implementation Mel-Filter Bank</i> .....	64
Tabel 4.10 <i>Implementation DCT</i> .....	66
Tabel 4.11 <i>Implementation Normalization</i> .....	67
Tabel 4.12 <i>Implementation Padding</i> .....	69
Tabel 4.13 <i>Output Model Training</i> .....	75
Tabel 4.14 <i>Evaluation Model 1</i> .....	76
Tabel 4.15 <i>Evaluation Model 2</i> .....	77
Tabel 4.16 <i>Evaluation Model 3</i> .....	79
Tabel 4.17 <i>Evaluation Model 4</i> .....	80
Tabel 4.18 <i>Evaluation Model 5</i> .....	82

## INTISARI

Penelitian mengembangkan sistem Gender Voice Recognition menggunakan kombinasi Mel-Frequency Cepstral Coefficients (MFCC) untuk ekstraksi fitur dan Convolutional Neural Networks (CNN) untuk klasifikasi suara berdasarkan gender. MFCC dipilih karena kemampuannya dalam menangkap karakteristik penting dari sinyal suara, terutama frekuensi rendah yang lebih relevan dengan persepsi manusia. CNN digunakan untuk klasifikasi karena kemampuannya mengenali pola kompleks secara efektif. Hasil pengujian menunjukkan bahwa sistem yang dikembangkan memiliki akurasi tinggi dengan nilai F1-score sebesar 0.97 untuk suara laki-laki dan 0.95 untuk suara perempuan. Selain itu, validasi menggunakan K-Fold Cross-Validation memperlihatkan akurasi tertinggi sebesar 94.74% pada fold ke-4, membuktikan bahwa model memiliki performa yang stabil dan mampu melakukan generalisasi dengan baik. Dengan demikian, kombinasi MFCC dan CNN terbukti efektif dalam membangun sistem pengenalan gender suara yang akurat.

**Kata kunci:** MFCC, CNN, Gender, Voice recognition, K-Fold



## **ABSTRACT**

The study developed a Gender Voice Recognition system using a combination of Mel-Frequency Cepstral Coefficients (MFCC) for feature extraction and Convolutional Neural Networks (CNN) for gender classification based on voice. MFCC was chosen for its ability to capture important characteristics of the audio signal, especially low frequencies that are more relevant to human perception. CNN was used for classification due to its effectiveness in recognizing complex patterns. The test results show that the developed system achieved high accuracy, with an F1-score of 0.97 for male voices and 0.95 for female voices. Additionally, validation using K-Fold Cross-Validation demonstrated the highest accuracy of 94.74% on the 4th fold, proving that the model has stable performance and strong generalization capabilities. Thus, the combination of MFCC and CNN is proven to be effective in building an accurate gender voice recognition system.

**Keywords:** MFCC, CNN, Gender, Voice recognition