



**IMPLEMENTASI ALGORITMA CONVOLUTION NEURAL  
NETWORKS DENGAN EKSTRASI FITUR AUDIO DALAM  
KLASIFIKASI EMOSI SUARA**

**LAPORAN TUGAS AKHIR**

**ALIF IBRAHIM SHOLEH**

**41520010184**

**UNIVERSITAS  
MERCU BUANA**

**PROGRAM STUDI TEKNIK INFORMATIKA  
FAKULTAS ILMU KOMPUTER  
UNIVERSITAS MERCU BUANA  
JAKARTA  
2025**



**IMPLEMENTASI ALGORITMA CONVOLUTION NEURAL  
NETWORKS DENGAN EKSTRAKSI FITUR AUDIO DALAM  
KLASIFIKASI EMOSI SUARA**

**LAPORAN TUGAS AKHIR**

**ALIF IBRAHIM SHOLEH**

**41520010184**

**Diajukan sebagai salah satu syarat untuk memperoleh gelar sarjana**

**MERCU BUANA**

**PROGRAM STUDI TEKNIK INFORMATIKA**

**FAKULTAS ILMU KOMPUTER**

**UNIVERSITAS MERCU BUANA**

**JAKARTA**

**2025**

## HALAMAN PERNYATAAN KARYA SENDIRI

Saya yang bertanda tangan di bawah ini:

Nama : Alif Ibrahim Sholeh  
NIM : 41520010184  
Program Studi : Teknik Informatika  
Judul Laporan Skripsi : Implementasi Algoritma Convolution Neural Networks dengan Ekstraksi Fitur Audio dalam Klasifikasi Emosi Suara

Menyatakan bahwa Laporan Skripsi ini adalah hasil karya saya sendiri dan bukan plagiat, serta semua sumber baik yang dikutip maupun dirujuk telah saya nyatakan dengan benar. Apabila ternyata ditemukan di dalam Laporan Skripsi saya terdapat unsur plagiat, maka saya siap mendapatkan sanksi akademis yang berlaku di Universitas Mercu Buana.



Jakarta, 17 Desember 2024



Alif Ibrahim Sholeh

UNIVERSITAS  
MERCU BUANA

## HALAMAN PENGESAHAN

Laporan Skripsi ini diajukan oleh:

Nama : Alif Ibrahim Sholeh  
NIM : 41520010184  
Program Studi : Teknik Informatika  
Judul Laporan Skripsi : Implementasi Algoritma Convolution Neural Network dengan Ekstrasi Fitur Audio dalam Klasifikasi Emosi Suara

Telah berhasil dipertahankan pada sidang di hadapan Dewan Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Strata 1 pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer Universitas Mercu Buana.

Disahkan oleh:

Pembimbing : Muhaimin Hasanudin, ST., M.Kom

NIDN : 0420027508



Ketua Penguji : Dr. Hadi Santoso, S.Kom., M.Kom.

NIDN : 0225067701



Penguji 1 : Emil Robert Kaburuan, S.T., M.A.,  
Ph.D

NIDN : 0429058004



Penguji 2 : Anis Cherid, S.E., M.TI

NIDN : 0328127203



Jakarta, 19 Februari 2025

Mengetahui,

Dekan



Dr. Bambang Jekonowo, S.Si., MTI  
NIDN : 0320037002

Ketua Program Studi



Dr. Hadi Santoso, S.Kom., M.Kom  
NIDN : 0225067701

## KATA PENGANTAR

Puji syukur kehadirat Tuhan yang Maha Esa, atas segala rahmat dan ridhanya sehingga penulis dapat menyelesaikan tugas akhir yang merupakan salah satu persyaratan kelulusan Program Studi Strata Satu (S1) pada jurusan Teknik Informatika, Universitas Mercu Buana.

Penulis menyadari bahwa tugas akhir ini masih jauh dari sempurna, karena kesempurnaan sejatinya hanya milik Tuhan yang Maha Esa. Oleh karena itu, saran dan masukan yang membangun senantiasa penulis terima dengan senang hati. Serta berkat dukungan, motivasi, bantuan, bimbingan, dan doa dari banyak pihak, penulis mengucapkan terima kasih kepada:

1. Bapak Prof. Dr. Andi Adriansyah, M.Eng. selaku Rektor Universitas Mercu Buana.
2. Bapak Dr. Bambang Jokonowo, S.Si., MTI selaku Dekan Fakultas Ilmu Komputer.
3. Bapak Dr. Hadi Santoso, S.Kom., M.Kom. selaku Ketua Program Studi Teknik Informatika Universitas Mercubuana.
4. Bapak Dosen Muhaimin Hasanudin, ST., M.Kom. selaku dosen pembimbing tugas akhir yang telah memberikan pengarahan, motivasi, menyediakan waktu, tenaga, dan pikiran sehingga selama pembuatan tugas akhir ini terjadwal dengan baik.
5. Kedua Orang Tua saya yang selalu mensupport dan mendukung saya selama menjalani masa studi sebagai mahasiswa Universitas Mercubuana.
6. Semua teman kuliah yang selalu berbagi informasi dan memberikan dukungan dalam bentuk yang berbeda-beda.

Akhir kata, penulis berharap semoga Tuhan yang Maha Esa membalas kebaikan dan selalu mencurahkan rahmat, hidayah, serta panjang umur kepada kita semua, aamiin. Terima Kasih.

Jakarta, 15 Januari 2024

Alif Ibrahim Sholeh

## HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademik Universitas Mercu Buana, saya yang bertanda tangan di bawah ini:

Nama : Alif Ibrahim Sholeh  
NIM : 41520010184  
Program Studi : Teknik Informatika  
Judul Laporan Skripsi : Implementasi Algoritma Convolution Neural Networks dengan Ekstraksi Fitur Audio dalam Klasifikasi Emosi Suara.

Demi pengembangan ilmu pengetahuan, dengan ini memberikan izin dan menyetujui untuk memberikan kepada Universitas Mercu Buana **Hak Bebas Royalti Non-Eksklusif (*Non-exclusive Royalty-Free Right*)** atas karya ilmiah saya yang berjudul di atas beserta perangkat yang ada (jika diperlukan).

Dengan Hak Bebas Royalti Non-Eksklusif ini Universitas Mercu Buana berhak menyimpan, mengalihmedia/format-kan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan mempublikasikan Laporan Magang/Skripsi/Tesis/Disertasi saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 17 Desember 2024

Yang menyatakan,

UNIVERSI  
MERCU BUANA



Alif Ibrahim Sholeh

## ABSTRAK

Nama : Alif Ibrahim Sholeh  
NIM : 41520010184  
Program Studi : Teknik Informatika  
Judul Laporan Skripsi : Implementasi Algoritma Convolution Neural Network dengan Ekstraksi Fitur Audio dalam Klasifikasi Emosi Suara  
Dosen Pembimbing : Muhaimin Hasanudin, ST., M.Kom

Ekspresi emosi dalam suara memainkan peran penting dalam komunikasi manusia, karena mampu menyampaikan maksud dan perasaan secara spontan dan alami. Penelitian ini bertujuan untuk mengenali emosi pada suara menggunakan algoritma Convolutional Neural Networks (CNN) yang didukung oleh enam metode ekstraksi fitur audio, yaitu Mel-Frequency Cepstral Coefficients (MFCC), Linear Predictive Coding (LPC), Mel-spectrogram, fitur berbasis waktu, fitur pitch dan rata-rata pitch. MFCC digunakan untuk mengekstrak informasi frekuensi utama dari sinyal suara, sedangkan LPC berfungsi untuk memodelkan spektrum akustik. Mel-spectrogram memberikan representasi visual dari suara berdasarkan skala Mel, sementara fitur temporal dan pitch mencerminkan pola durasi dan tinggi rendahnya nada suara. Dataset utama yang digunakan adalah Toronto Emotional Speech Set (TESS), yang merupakan dataset standar dalam penelitian pengenalan emosi suara. Pengujian pada dataset TESS menghasilkan performa yang sangat baik dengan akurasi 99,52%, presisi 0,9953, recall 0,9952, dan F1-score 0,9952. Penelitian ini juga menguji model pada dataset bahasa Indonesia yang dikumpulkan secara mandiri, dengan hasil yang menjanjikan berupa akurasi 84,29%, presisi 0,8477, recall 0,8429 dan F1-score 0,8408. Temuan ini menunjukkan bahwa CNN yang dipadukan dengan teknik ekstraksi fitur yang digunakan dapat secara efektif mengenali emosi dari suara, baik dalam bahasa Inggris maupun Indonesia, meskipun diperlukan penyempurnaan lebih lanjut pada dataset berbahasa Indonesia.

**Kata kunci:** *Analisis Suara Emosi, Fitur Ekstraksi Audio, Toronto Emotional Speech Set, Dataset Bahasa Indonesia.*

## ABSTRACT

Name : Alif Ibrahim Sholeh  
NIM : 41520010184  
Program Studi : informatic engineering  
Thesis Report Title : Implementation of Convolution Neural Network Algorithm with Audio Feature Extraction in Voice Emotion Classification  
Supervisor : Muhaimin Hasanudin, ST., M.Kom

*Emotional expressions in voice play an essential role in human communication, as they can convey intentions and feelings spontaneously and naturally. This research aims to recognize emotions in voice using a Convolutional Neural Networks (CNN) algorithm supported by six audio feature extraction methods, namely Mel-Frequency Cepstral Coefficients, Linear Predictive Coding (LPC), Mel-spectrogram, time-based features, pitch features, and pitch averaging. MFCC is used to extract the primary frequency information from a sound signal. At the same time, LPC serves to model the acoustic spectrum. The Mel-spectrogram visually represents the sound based on the Mel scale. At the same time, the temporal and pitch features reflect the duration and pitch patterns of the sound. The primary dataset used is the Toronto Emotional Speech Set, a standard voice emotion recognition research dataset. Testing on the TESS dataset resulted in an excellent performance with 99.52% accuracy, 0.9953 precision, 0.9952 recall, and 0.9952 F1-score. This research also tested the model on a self-collected Indonesian language dataset, with promising results of 84.29% accuracy, 0.8477 precision, 0.8429 recall, and 0.8408 F1-score. These findings show that CNN, combined with the feature extraction techniques, can effectively recognize emotions from voices in both English and Indonesian. However, further refinements are needed on the Indonesian dataset.*

**Keywords:** *Emotional Voice Analysis, Audio Extraction Features, Toronto Emotional Speech Set, Indonesian Language Dataset.*



## DAFTAR ISI

<b>HALAMAN JUDUL</b> .....	<b>i</b>
<b>HALAMAN PERNYATAAN</b> .....	<b>ii</b>
<b>HALAMAN PENGESAHAN</b> .....	<b>iii</b>
<b>KATA PENGANTAR</b> .....	<b>iv</b>
<b>HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI</b> .....	<b>v</b>
<b>ABSTRAK</b> .....	<b>vi</b>
<b>ABSTRACT</b> .....	<b>vii</b>
<b>DAFTAR ISI</b> .....	<b>viii</b>
<b>DAFTAR TABEL</b> .....	<b>xi</b>
<b>DAFTAR GAMBAR</b> .....	<b>xii</b>
<b>DAFTAR LAMPIRAN</b> .....	<b>xiii</b>
<b>BAB I PENDAHULUAN</b> .....	<b>1</b>
1.1 Latar Belakang .....	1
1.2 Perumusan Masalah .....	4
1.3 Tujuan Penelitian .....	4
1.4 Manfaat Penelitian .....	5
1.5 Batasan Masalah .....	5
<b>BAB II TINJAUAN PUSTAKA</b> .....	<b>7</b>
2.1 Penelitian Terdahulu .....	7
2.2 Teori Pendukung .....	28
2.2.1 Sinyal Suara Manusia .....	28
2.2.2 Teori Emosi Manusia .....	30
2.2.3 Jenis suara Manusia berdasarkan usia .....	31
2.2.4 Klasifikasi Emosi Manusia .....	31
<b>BAB III METODE PENELITIAN</b> .....	<b>34</b>
3.1 Jenis Penelitian.....	34
3.2 Tahapan Penelitian.....	34
3.2.1 Speech Collection .....	35
3.2.2 Speech Processing .....	36
3.2.3 Feature Extraction .....	37

3.2.4 Supervised Learning.....	40
3.2.5 Recognition Task.....	44
3.2.6 Evaluasi Model.....	44
3.2.7 Visualisasi .....	46
<b>BAB IV PEMBAHASAN .....</b>	<b>49</b>
4.1. implementasi CNN dengan ekstraksi fitur audio .....	49
4.1.1 Cara implementasi .....	50
4.2 Hasil Dataset TESS ((Toronto Emotional Speech Set).....	51
4.2.1 Waveplot .....	51
4.2.2 Akurasi dan Loss Model.....	53
4.2.3 Confusion Matrix .....	54
4.2.4 Performance Metrics .....	56
4.3 Hasil dataset bahasa indonesia.....	56
4.3.1 Waveplot Bahasa Indonesia .....	56
Akurasi dan Loss Model .....	59
Confusion Matrix .....	60
Distribution Of Predicted Classes.....	61
Performance Metrics .....	62
4.3 Hasil analisis klasifikasi suara emosi CNN dengan ekstraksi fitur Audio.63	
4.4.1 hasil analisis ekstrasi fitur ekstrasi audio dalam klasifikasi emosi.....	63
hasil analisis ekstrasi fitur MFCC dalam klasifikasi emosi Menggunakan Bahasa Indonesia .....	71
4.2.3 hasil analisis algoritma CNN dalam klasifikasi emosi .....	77
4.5 Implikasi Industri hasil klasifikasi emosi .....	79
4.6 Perhitungan Manual .....	80
4.7 Prediksi Label Dataset TESS menggunakan model prediksi Dataset Bahasa Indonesia.....	83
4.7.1 Confusion Matrix .....	83
4.7.2 Distribusi Prediksi Emosi .....	84
4.7.3 performance Metrics .....	85
4.8 Prediksi Label Dataset Indonesia menggunakan model prediksi Dataset TESS.....	87

4.8.1 Confusion Matrix .....	87
4.8.2 Distribusi Prediksi Emosi .....	88
4.8.3 performance Metrics .....	89
BAB V KESIMPULAN DAN SARAN .....	91
5.1 Kesimpulan .....	91
5.2 Saran .....	92
DAFTAR PUSTAKA .....	94
LAMPIRAN.....	100



## DAFTAR TABEL

Table 1 Representasi dataset .....	2
Tabel 2. 1 Penelitian Terkait .....	7
Tabel 4. 1 Performance Metrics Tess .....	56
Tabel 4. 2 Performance Metrics .....	62



## DAFTAR GAMBAR

Gambar 2. 1sinyal suara manusia .....	29
Gambar 3. 1Diagram Alir .....	35
Gambar 3. 2 Confusion Matriks.....	45
Gambar 4. 1Waveplot .....	52
Gambar 4. 2 Akurasi dan Loss Model .....	53
Gambar 4. 3 Confusion Matriks.....	54
Gambar 4. 4 Distribution of predicted class .....	55
Gambar 4. 5 Waveplot bahasa indonesia .....	58
Gambar 4. 6 Akurasi dan loss model Bahasa indonesia .....	59
Gambar 4. 7 Confusion matriks Bahasa indonesia .....	60
Gambar 4. 8 Distribution of predicted classes Bahasa indonesia .....	61
Gambar 4. 9 Hasil ekstraksi fitur MFCC .....	70



## DAFTAR LAMPIRAN

Lampiran 1 Kartu Asistensi .....	18
Lampiran 2 Lampiran Halaman Pernyataan Luaran Tugas Akhir (Khusus Publikasi Ilmiah) .....	19
Lampiran 3 Lampiran Naskah Artikel Jurnal (Khusus Publikasi Ilmiah) .....	20
Lampiran 4 Curriculum Vitae .....	21
Lampiran 5 Surat Pernyataan HAKI .....	22
Lampiran 6 Sertifikat BNSP .....	24
Lampiran 7 Hasil Cek Turnitin .....	25

