

**DETEKSI EMOSI DARI SUARA MANUSIA MENGGUNAKAN  
METODE CONVOLUTIONAL NEURAL NETWORK (CNN)**

**SKRIPSI**

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana  
Program Studi S1 Sistem Informasi



disusun oleh

**ZAIDAN RAFIF SYAUQI**

**22.12.2731**

Kepada

**FAKULTAS ILMU KOMPUTER  
UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA**

**2026**

# **DETEKSI EMOSI DARI SUARA MANUSIA MENGGUNAKAN METODE CONVOLUTIONAL NEURAL NETWORK (CNN)**

## **SKRIPSI**

untuk memenuhi salah satu syarat mencapai derajat Sarjana  
Program Studi S1 Sistem Informasi



disusun oleh

**ZAIDAN RAFIF SYAUQI**

**22.12.2731**

Kepada

**FAKULTAS ILMU KOMPUTER  
UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA**

**2026**

**HALAMAN PERSETUJUAN**

**SKRIPSI**

**DETEKSI EMOSI DARI SUARA MANUSIA MENGGUNAKAN METODE  
CONVOLUTIONAL NEURAL NETWORK (CNN)**

yang disusun dan diajukan oleh

**Zaidan Rafif Syauqi**

**22.12.2731**

telah disetujui oleh Dosen Pembimbing Skripsi  
pada tanggal <21 April 2026>

**Dosen Pembimbing,**



**Bety Wulan Sari, S.Kom., M.Kom.,  
NIK. 190302254**

HALAMAN PENGESAHAN

SKRIPSI

DETEKSI EMOSI DARI SUARA MANUSIA MENGGUNAKAN METODE  
CONVOLUTIONAL NEURAL NETWORK (CNN)

yang disusun dan diajukan oleh

Zaidan Rafif Syanqi

22.12.2731

Telah dipertahankan di depan Dewan Penguji  
pada tanggal <21 April 2026>

Susunan Dewan Penguji

Nama Penguji

Tanda-Tangan

Ali Mustina, S.Kom., M.Kom  
NIK. 190302192

Arif Nur Rohman, S.Kom., M.Kom  
NIK. 190302684

Bety Wulan Sari, S.Kom., M.Kom  
NIK. 190302254

Skrripsi ini telah diterima sebagai salah satu persyaratan  
untuk memperoleh gelar Sarjana Komputer  
Tanggal 21 April 2026

DEKAN FAKULTAS ILMU KOMPUTER



Prof. Dr. Kusrini, M.Kom.  
NIK. 190302106

## HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Zaidan Rafif Syauqi  
NIM : 22.12.2731

Menyatakan bahwa Skripsi dengan judul berikut:

### **DETEKSI EMOSI DARI SUARA MANUSIA MENGGUNAKAN METODE CONVOLUTIONAL NEURAL NETWORK (CNN)**

Dosen Pembimbing : Bety Wulan Sari, S.Kom., M.Kom.

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 21 April 2026

Yang Menyatakan,

  
Zaidan Rafif Syauqi

## HALAMAN PERSEMBAHAN

Puji syukur ke hadirat Allah SWT atas segala rahmat, hidayah dan karunia-Nya yang memberikan kekuatan serta kelancaran dalam Menyusun skripsi ini. Dengan kerendahan hati dan rasa hormat, karya ini dipersembahkan kepada:

1. Orang tua tercinta, atas doa, kasih sayang, dukungan moril dan materiil yang tak terhingga, serta kesabaran yang tak pernah pudar selama proses pendidikan dan penyusunan skripsi ini.
2. Ibu Bety Wulan sari, S.Kom., M.Kom. sebagai Dosen Pembimbing yang telah meluangkan waktu, memberikan arahan, saran, dan bimbingan yang berharga dari awal hingga akhir penyusunan skripsi ini.
3. Tim Dosen Penguji yang akan memberikan masukan konstruktif demi kesempurnaan skripsi ini.
4. Seluruh keluarga besar yang selalu memberikan dukungan dan motivasi.
5. Rekan-rekan seperjuangan di Program Studi Sistem Informasi, atas kebersamaan, semangat dan dukungannya.
6. Semua pihak yang tidak dapat penulis sebutkan satu per satu, atas segala bentuk bantuan, dukungan, dan motivasi yang telah diberikan.

## KATA PENGANTAR

Puji syukur ke hadirat Tuhan Yang Maha Esa atas segala rahmat dan karunia-Nya, sehingga penulis dapat menyelesaikan skripsi ini dengan judul "Deteksi Emosi dari Suara Manusia Menggunakan Metode Convolutional Neural Network (CNN)" dengan baik dan tepat waktu. Skripsi ini disusun sebagai salah satu syarat untuk memperoleh gelar Sarjana pada Program Studi Sistem Informasi, Fakultas Ilmu Komputer, Universitas Amikom Yogyakarta.

Penyelesaian skripsi ini tidak lepas dari bimbingan, dukungan, dan bantuan dari berbagai pihak. Oleh karena itu, penulis ingin menyampaikan rasa terima kasih yang tulus dan mendalam kepada:

1. Allah SWT yang telah memberikan terang-Nya kepada penulis sehingga mampu menyelesaikan pengerjaan skripsi ini.
2. Ibu Bety Wulan sari, S.Kom., M.Kom. sebagai Dosen Pembimbing yang telah meluangkan waktu, memberikan arahan, saran, dan bimbingan yang berharga dari awal hingga akhir penyusunan skripsi ini.
3. Tim Dosen Penguji yang akan memberikan masukan konstruktif demi kesempurnaan skripsi ini.
4. Orang tua tercinta, Bapak Widodo atas doa, kasih sayang, dukungan moril dan materiil yang tak terhingga, serta kesabaran yang tak pernah pudar selama proses pendidikan dan penyusunan skripsi ini.
5. Seluruh keluarga besar yang selalu memberikan dukungan dan motivasi.
6. Semua pihak yang tidak dapat penulis sebutkan satu per satu, atas segala bentuk bantuan, dukungan, dan motivasi yang telah diberikan.

Penulis menyadari bahwa skripsi ini masih jauh dari kata sempurna dan masih banyak kekurangan. Oleh karena itu, penulis sangat mengharapkan kritik dan saran yang membangun demi perbaikan di masa mendatang. Semoga skripsi ini dapat memberikan manfaat bagi perkembangan ilmu pengetahuan dan pihak-pihak yang membutuhkan.

Yogyakarta, 21 April 2026

Zaidan Rafif Syauqi

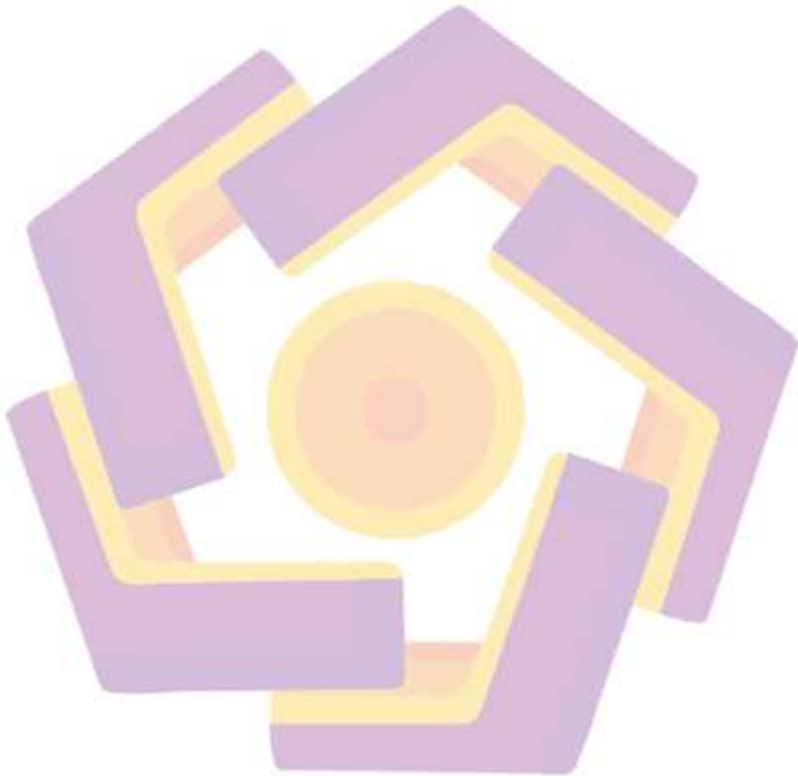
## DAFTAR ISI

|                                                       |             |
|-------------------------------------------------------|-------------|
| <b>HALAMAN JUDUL.....</b>                             | <b>i</b>    |
| <b>HALAMAN PERSETUJUAN.....</b>                       | <b>ii</b>   |
| <b>HALAMAN PENGESAHAN.....</b>                        | <b>iii</b>  |
| <b>HALAMAN PERNYATAAN KEASLIAN SKRIPSI.....</b>       | <b>iv</b>   |
| <b>HALAMAN PERSEMBAHAN.....</b>                       | <b>v</b>    |
| <b>KATA PENGANTAR.....</b>                            | <b>vi</b>   |
| <b>DAFTAR ISI.....</b>                                | <b>vii</b>  |
| <b>DAFTAR TABEL.....</b>                              | <b>ix</b>   |
| <b>DAFTAR GAMBAR.....</b>                             | <b>x</b>    |
| <b>DAFTAR LAMPIRAN.....</b>                           | <b>xi</b>   |
| <b>INTISARI.....</b>                                  | <b>xii</b>  |
| <b>ABSTRACT.....</b>                                  | <b>xiii</b> |
| <b>BAB I.....</b>                                     | <b>1</b>    |
| <b>PENDAHULUAN.....</b>                               | <b>1</b>    |
| 1.1 Latar Belakang.....                               | 1           |
| 1.2 Rumusan Masalah.....                              | 3           |
| 1.3 Batasan Masalah.....                              | 4           |
| 1.4 Tujuan Penelitian.....                            | 5           |
| 1.5 Manfaat Penelitian.....                           | 6           |
| 1.5.1 Manfaat Teoritis.....                           | 6           |
| 1.5.2 Manfaat Praktis.....                            | 7           |
| 1.6 Sistematika Penulisan.....                        | 7           |
| <b>BAB II.....</b>                                    | <b>9</b>    |
| <b>TINJAUAN PUSTAKA.....</b>                          | <b>9</b>    |
| 2.1 Studi Literatur.....                              | 9           |
| 2.2 Dasar Teori.....                                  | 11          |
| 2.2.1 Emosi Manusia.....                              | 11          |
| 2.2.2 Sinyal Suara.....                               | 13          |
| 2.2.3 Ekstraksi Fitur Audio.....                      | 14          |
| 2.2.4 Mel-Frequency Cepstral Coefficients (MFCC)..... | 15          |
| 2.2.5 Pra-pemrosesan Suara.....                       | 16          |
| 2.2.6 Machine Learning.....                           | 18          |
| 2.2.7 Support Vector Machine (SVM).....               | 20          |
| 2.2.8 Evaluasi Kinerja Model.....                     | 23          |
| 2.2.9 Speech Emotion Recognition (SER).....           | 24          |
| <b>BAB III</b>                                        |             |

|                                                    |           |
|----------------------------------------------------|-----------|
| <b>METODE PENELITIAN.....</b>                      | <b>28</b> |
| 3.1 Objek Penelitian.....                          | 28        |
| 3.2 Alur Penelitian.....                           | 29        |
| 3.3 Alat dan Bahan.....                            | 34        |
| <b>BAB IV</b>                                      |           |
| <b>HASIL DAN PEMBAHASAN.....</b>                   | <b>36</b> |
| 4.1 Hasil Penelitian.....                          | 36        |
| 4.1.1 Hasil Preprocessing Data Suara.....          | 36        |
| 4.1.2 Hasil Ekstraksi Fitur MFCC.....              | 38        |
| 4.1.3 Arsitektur Model CNN.....                    | 39        |
| 4.1.4 Hasil Pelatihan Model.....                   | 41        |
| 4.1.5 Evaluasi Model.....                          | 41        |
| 4.1.6 Implementasi Sistem Berbasis Web.....        | 42        |
| 4.2 Pembahasan.....                                | 43        |
| 4.2.1 Analisis Performa Model.....                 | 43        |
| 4.2.2 Pengaruh Penggunaan MFCC.....                | 43        |
| 4.2.3 Analisis Overfitting dan Generalisasi.....   | 44        |
| 4.2.4 Faktor yang Mempengaruhi Performa Model..... | 46        |
| 4.2.5 Implikasi Hasil Penelitian.....              | 47        |
| <b>BAB V</b>                                       |           |
| <b>PENUTUP.....</b>                                | <b>50</b> |
| 5.1 Kesimpulan.....                                | 50        |
| 5.2 Saran.....                                     | 51        |
| <b>REFERENSI.....</b>                              | <b>52</b> |
| <b>LAMPIRAN.....</b>                               | <b>54</b> |

## DAFTAR TABEL

|                                                    |    |
|----------------------------------------------------|----|
| Tabel 2.1 Keaslian Penelitian                      | 28 |
| Tabel 4.1 Ringkasan Hasil Preprocessing Audio      | 54 |
| Tabel 4.2 Distribusi Dataset Setelah Preprocessing | 54 |
| Tabel 4.3 Parameter Ekstraksi MFCC                 | 55 |
| Tabel 4.4 Arsitektur Model CNN                     | 56 |
| Tabel 4.4 Arsitektur Model CNN                     | 56 |
| Tabel 4.6 Hasil Classification Report              | 58 |



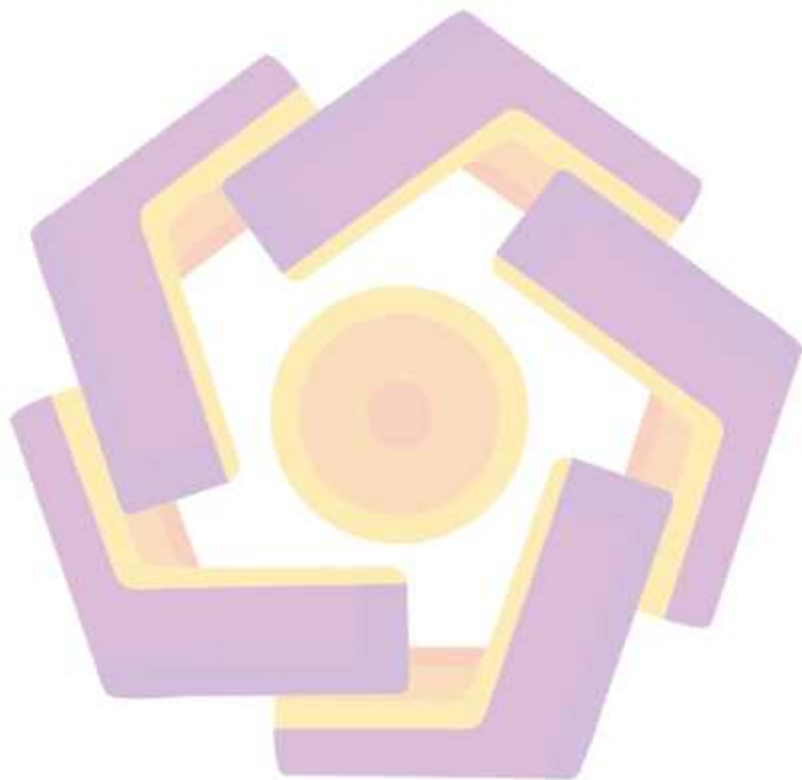
## DAFTAR GAMBAR

|                                                         |    |
|---------------------------------------------------------|----|
| Gambar 2.1 Model Valence–Arousal                        | 29 |
| Gambar 2.2 Gelombang Sinyal Suara (Waveform)            | 31 |
| Gambar 2.3 Diagram Alur Proses MFCC                     | 33 |
| Gambar 2.5 Hyperplane dan Margin pada SVM               | 39 |
| Gambar 2.4 Diagram Umum SER                             | 43 |
| Gambar 3.1 Alur Penelitian                              | 48 |
| Gambar 4.1 Sistem Berbasis Web                          | 59 |
| Gambar 4.2 Visualisasi Fitur MFCC + Delta + Delta-Delta | 61 |
| Gambar 4.3 Grafik Akurasi Model CNN                     | 62 |
| Gambar 4.4 Grafik Loss Model CNN                        | 63 |
| Gambar 4.5 Implementasi Website                         | 65 |



## DAFTAR LAMPIRAN

|                                     |    |
|-------------------------------------|----|
| Lampiran 1. Profil obyek Penelitian | 10 |
| Lampiran 2. Dokumentasi Penelitian  | 11 |



## INTISARI

Emosi merupakan salah satu aspek penting dalam komunikasi manusia yang dapat disampaikan melalui berbagai media, salah satunya adalah suara. Seiring dengan perkembangan teknologi, kemampuan sistem komputer dalam mengenali emosi manusia menjadi kebutuhan yang semakin penting, terutama pada bidang interaksi manusia dan komputer, layanan pelanggan, serta kesehatan mental. Penelitian ini bertujuan untuk mengembangkan sistem deteksi emosi berbasis suara menggunakan metode Convolutional Neural Network (CNN) dengan memanfaatkan fitur Mel-Frequency Cepstral Coefficients (MFCC). Dataset yang digunakan dalam penelitian ini adalah RAVDESS (Ryerson Audio-Visual Database of Emotional Speech and Song) dengan empat kategori emosi, yaitu marah (angry), senang (happy), sedih (sad), dan netral (neutral). Tahapan penelitian meliputi preprocessing data suara, ekstraksi fitur menggunakan MFCC yang dikombinasikan dengan fitur delta dan delta-delta, serta proses klasifikasi menggunakan model CNN. Proses preprocessing dilakukan untuk meningkatkan kualitas sinyal suara melalui normalisasi, penghapusan bagian hening, dan penyesuaian durasi. Model CNN yang dibangun terdiri dari beberapa lapisan konvolusi, pooling, dropout, dan fully connected layer yang berfungsi untuk mengekstraksi dan mengklasifikasikan pola emosi dari data suara. Evaluasi model dilakukan menggunakan metrik akurasi, precision, recall, dan F1-score, serta analisis confusion matrix. Hasil penelitian menunjukkan bahwa model CNN mampu mencapai akurasi sebesar 76% dalam mengklasifikasikan emosi suara. Selain itu, model yang telah dilatih diimplementasikan dalam bentuk aplikasi berbasis web menggunakan Streamlit sehingga dapat digunakan untuk mendeteksi emosi dari input suara secara langsung. Hasil penelitian ini menunjukkan bahwa kombinasi fitur MFCC dan metode CNN cukup efektif dalam sistem Speech Emotion Recognition (SER), meskipun masih terdapat keterbatasan dalam membedakan emosi dengan karakteristik yang mirip, terutama pada kelas netral.

**Kata kunci:** Speech Emotion Recognition, MFCC, Convolutional Neural Network, Klasifikasi Emosi, Audio Processing.

## ABSTRACT

Emotion is an essential component of human communication that can be conveyed through various modalities, one of which is speech. With the rapid advancement of technology, the ability of computer systems to recognize human emotions has become increasingly important, particularly in areas such as human-computer interaction, customer service, and mental health applications. This study aims to develop a speech-based emotion detection system using the Convolutional Neural Network (CNN) method with Mel-Frequency Cepstral Coefficients (MFCC) as the main feature. The dataset used in this research is the RAVDESS (Ryerson Audio-Visual Database of Emotional Speech and Song), consisting of four emotion categories: angry, happy, sad, and neutral. The research process includes audio preprocessing, feature extraction using MFCC combined with delta and delta-delta features, and classification using a CNN model. The preprocessing stage is performed to improve audio quality through normalization, silence removal, and duration adjustment. The CNN model is designed with several layers, including convolutional, pooling, dropout, and fully connected layers, to automatically extract and classify emotional patterns from speech data. Model evaluation is conducted using accuracy, precision, recall, and F1-score metrics, along with confusion matrix analysis. The experimental results show that the proposed CNN model achieves an accuracy of 76% in classifying speech emotions. Furthermore, the trained model is implemented into a web-based application using Streamlit, allowing users to input audio and obtain emotion predictions in real time. The results indicate that the combination of MFCC features and CNN is effective for Speech Emotion Recognition (SER), although challenges remain in distinguishing emotions with similar characteristics, particularly the neutral class.

**Keywords:** Speech Emotion Recognition, MFCC, Convolutional Neural Network, Emotion Classification, Audio Processing.