

# Implementasi Ekstraksi Gammatone-Frequency Cepstral Coefficient dan Klasifikasi Hidden Markov Model dalam Identifikasi Emosi Menggunakan Suara Jantung

Dimas Resha Putra Pamungkas<sup>1</sup>, Barlian Henryranu Prasetyo<sup>2</sup>

Program Studi Teknik Komputer, Fakultas Ilmu Komputer, Universitas Brawijaya  
Email: <sup>1</sup>dimasrsha@ub.ac.id, <sup>2</sup>barlian@ub.ac.id

## Abstrak

Kondisi emosional seseorang merupakan faktor penting dalam interaksi antar manusia, dan memengaruhi beberapa aspek di dalam komunikasi. Penelitian ini dilakukan untuk membuat sebuah sistem identifikasi emosi melalui suara jantung dengan implementasi ekstraksi Gammatone-Frequency Cepstral Coefficient (GFCC) dan klasifikasi Hidden Markov Model (HMM). Pada penelitian ini, kondisi emosi manusia akan dikelompokkan menjadi dua kelas berdasarkan nilai *Beat Per Minute* (BPM) dengan kelas tinggi untuk emosi senang, sedih, marah, takut, cemas, dan kelas rendah terkait dengan emosi santai dan bosan. Implementasi dilakukan melalui sebuah stetoskop elektronik yang terintegrasi dengan sebuah aplikasi android. Penelitian ini penting karena dapat memberikan informasi terkait efektivitas penggunaan GFCC pada pemrosesan suara jantung dalam identifikasi emosi. Pengujian *Signal-to-Noise Ratio* (SNR) yang dilakukan untuk mengetahui pengaruh ekstraksi GFCC dalam mengurangi *noise* pada suara jantung memperoleh hasil 93,33% yang membuktikan bahwa GFCC dapat mengurangi *noise* dengan baik. Selain itu, tingkat akurasi yang diperoleh sistem ini mencapai 75% pada akurasi validasi sistem dan 73,33% pada akurasi pengujian. Hasil tersebut membuktikan bahwa sistem mampu memprediksi label atau kelas dengan baik menggunakan suara jantung sebagai *input* utama sistem. Integrasi antara perangkat keras stetoskop elektronik dengan *smartphone* melalui aplikasi android membuat sistem ini mudah digunakan oleh pengguna. Penelitian ini memberikan kontribusi bagi pengguna dalam melakukan identifikasi emosi seseorang secara non-verbal melalui analisis suara jantung. Selain itu, penelitian ini juga diharapkan mampu membantu orang dengan keterbatasan komunikasi verbal serta mengurangi angka penderita gangguan kesehatan mental akibat ketidakstabilan emosional.

**Kata kunci:** *Suara Jantung, Emosi, Gammatone-Frequency Cepstral Coefficients, Hidden Markov Model, Signal-to-Noise Ratio*

## Abstract

*A person's emotional condition is an important factor in human interaction, and affects several aspects of communication. This research was conducted to create an emotion identification system using heart sounds by implementing Gammatone-Frequency Cepstral Coefficient (GFCC) extraction and Hidden Markov Model (HMM) classification. In this research, human emotional conditions will be grouped into two classes based on the Beat Per Minute (BPM) value with a high class for happy, sad, angry, afraid, anxious emotions, and a low class related to relaxed and bored emotions. The Implementation is carried out through an electronic stethoscope integrated with an Android application. This research is important because it can provide information regarding the effectiveness of using GFCC in processing heart sounds in identifying emotions. The Signal-to-Noise Ratio (SNR) test conducted to determine the effect of GFCC extraction in reducing noise in heart sounds obtained a result of 93.33% which proves that GFCC can reduce noise well. In addition, the level of accuracy obtained by this system reached 75% in system validation accuracy and 73.33% in testing accuracy. These results prove that the system is able to predict labels or classes well using heart sounds as the main input for the system. The integration of electronic stethoscope hardware with smartphones through android applications makes this system can be developed and usable. This research contributes to users in identifying someone's emotions non-verbally through heart sound analysis. In addition, this research is also expected to be*

able to help people with limited verbal communication and reduce the number of people with mental health disorders due to emotional instability.

**Keywords:** Heart Sound, Emotion, Gammatone-Frequency Cepstral Coefficients, Hidden Markov Model, Signal-to-Noise Ratio

## 1. PENDAHULUAN

Kondisi emosional seseorang merupakan faktor penting dalam interaksi antar manusia, dan memengaruhi beberapa aspek di dalam komunikasi, seperti ekspresi wajah, karakteristik suara, dan konten informasi linguistik (Aini et al., 2021). Dalam era modern ini, pemantauan dan manajemen emosi menjadi semakin penting, terutama karena meningkatnya prevalensi gangguan emosional seperti ketakutan, kecemasan, dan depresi. Menurut Robert Plutchik dalam roda emosi yang dikembangkan yang dikenal sebagai roda emosi Plutchik yaitu membagi emosi dasar menjadi delapan bagian dan setiap emosi dasar dipecah lagi menjadi tiga bagian berdasarkan intensitasnya yaitu rendah, sedang, hingga tinggi (Bimo et al., 2020).

Deteksi emosi telah berkembang menjadi fokus penelitian yang luas, terutama dalam pengembangan teknologi berbasis kecerdasan buatan. Metode deteksi emosi konvensional umumnya mengandalkan ekspresi wajah, intonasi suara, atau teks. Namun, metode ini memiliki keterbatasan, seperti sensitivitas terhadap faktor lingkungan atau kurangnya keakuratan dalam mengenali emosi tersembunyi. Dalam konteks penelitian dan aplikasi medis, pengenalan emosi melalui ucapan atau suara sudah menjadi hal yang familiar. Beberapa penelitian terdahulu mengungkapkan bahwa terdapat parameter yang mendukung adanya hubungan yang kuat antara ucapan dengan emosi yang sedang dirasakan. Parameter tersebut diantaranya adalah *pitch*, gelombang, artikulasi dan bentuk spektral (Helmiyah et al., 2019). Meskipun berbagai solusi tersebut menjanjikan, metode berbasis ekspresi wajah atau suara masih terbatas pada konteks tertentu dan kurang mampu mendeteksi emosi internal. Selain itu, metode ini mungkin tidak selalu dapat diakses atau praktis untuk penggunaan sehari-hari. Oleh karena itu, perlu ditemukan alternatif lain dalam melakukan deteksi emosi selain menggunakan suara atau ucapan.

Sensor dan sinyal fisiologis dinilai dapat menawarkan potensi yang lebih baik, tetapi sering kali memerlukan perangkat yang mahal

atau invasif. Penelitian ini mengusulkan pendekatan baru untuk deteksi emosi menggunakan suara jantung. Suara jantung memiliki hubungan langsung dengan aktivitas sistem saraf otonom, yang berperan dalam regulasi emosi. Suara jantung merupakan bunyi yang dihasilkan oleh aktivitas mekanis jantung pada siklus pemompaan darah. suara jantung dapat dianggap sebagai indikator fisiologis yang mencerminkan keadaan emosional seseorang (Harper et al., 2022). Dalam beberapa penelitian sebelumnya, dilakukan klasifikasi emosi menggunakan suara jantung yang dinilai lebih efektif dan fleksibel untuk semua jenis pengguna. Menurut Li et al. (2019), Mel-Frequency Cepstral Coefficient (MFCC) dapat digunakan sebagai metode ekstraksi fitur dalam proses klasifikasi emosi menggunakan suara jantung. Namun metode ekstraksi fitur suara jantung dengan MFCC dinilai memiliki beberapa kelemahan seperti kurang tahan terhadap *noise* seperti suara pernapasan atau gerak tubuh, dan memiliki resolusi yang terbilang kurang pada frekuensi rendah. Selain itu penggunaan metode klasifikasi *machine learning* yang dikombinasikan dengan ekstraksi fitur dalam beberapa penelitian terhadap suara jantung juga pernah dilakukan dengan hasil yang cukup beragam.

Oleh karena itu, pada penelitian ini peneliti berencana untuk melakukan implementasi ekstraksi lain yang dinilai lebih tahan terhadap *noise* dan memiliki resolusi lebih baik pada frekuensi rendah yaitu metode ekstraksi Gammatone-Frequency Cepstral Coefficient (GFCC) dengan Klasifikasi Hidden Markov Model (HMM) dalam identifikasi emosi menggunakan suara jantung dengan membagi emosi menjadi beberapa intensitas. Penggunaan metode HMM untuk identifikasi emosi menggunakan suara jantung dilakukan karena karakteristik HMM yang cocok untuk analisis data sekuensial seperti sinyal suara jantung. Dalam pengelompokannya berdasarkan *Beat Per Minute* (BPM), emosi dasar manusia umumnya terbagi menjadi dua tingkatan yaitu tinggi dan rendah yang masing-masing memiliki keterkaitan dengan emosi seperti tinggi yang terkait dengan emosi senang, sedih, marah, takut, cemas, dan rendah yang terkait dengan

emosi santai dan bosan (Jha et al., 2018). Pengelompokan emosi yang lebih spesifik memerlukan indikator dan parameter tambahan lain agar tidak terjadi ambiguitas dalam klasifikasi kelas emosi. Penelitian-penelitian terdahulu telah melakukan deteksi emosi serupa menggunakan suara manusia maupun suara jantung dengan menggunakan berbagai metode yang berbeda. Oleh karena itu, peneliti ingin mencoba kombinasi penerapan metode ekstraksi GFCC dan klasifikasi HMM untuk mengidentifikasi emosi menggunakan suara jantung.

Berdasarkan permasalahan dan penelitian yang ada, maka penelitian ini diharapkan mampu memberikan informasi terkait pengaruh, efektifitas, dan akurasi penggunaan metode ekstraksi fitur Gammatone-Frequency Cepstral Coefficient (GFCC) dan klasifikasi Hidden Markov Model (HMM) dalam mendeteksi emosi pada manusia menggunakan suara jantung. Selain itu, pengimplementasian dari hasil penelitian ini dapat bermanfaat dalam bidang pengenalan emosi dan peningkatan kesehatan emosional untuk tujuan diagnostik atau intervensi terapi.

## 2. LANDASAN KEPUSTAKAAN

### 2.1. Heart Sound (Suara Jantung)

Suara jantung merupakan suara yang dihasilkan oleh getaran yang timbul saat katup jantung yaitu katup mitral dan triskupid membuka dan menutup serta aliran darah dalam jantung. Suara jantung pada umumnya dapat didengarkan menggunakan stetoskop atau melalui perangkat digital. Suara jantung atau *heart sound* dapat didefinisikan juga sebagai komponen utama yang dapat didengar oleh jantung yang mencakup suara jantung pertama (S1) atau juga biasa disebut lub dan suara jantung kedua (S2) atau dub (Renna et al., 2019).

Secara fisiologis, suara jantung sering digunakan untuk indikator kondisi kesehatan jantung. Bahkan karena suara jantung secara langsung terkait dengan aktivitas otot jantung dan aliran darah, perubahan kecil pada suara ini bisa menjadi indikator perubahan fisiologis dalam tubuh, termasuk perubahan akibat stres atau emosi. Emosi seperti ketakutan atau kecemasan biasanya cenderung meningkatkan frekuensi dari suara jantung akibat peningkatan intensitas detak jantung. Pada beberapa kondisi lain, suara jantung dapat menjadi lebih cepat dan

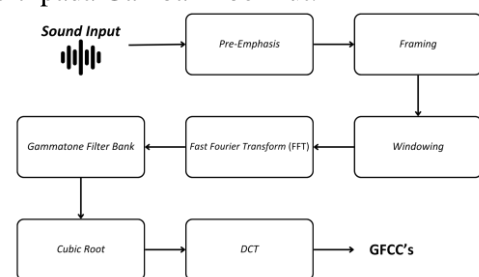
kuat yang mengartikan terjadinya peningkatan aktivitas sistem saraf simpatik sebagai respons dari rangsangan emosional seseorang.

### 2.2. Emosi

Emosi merupakan sebuah keadaan psikologis kompleks yang melibatkan pengalaman subjektif, respons fisiologis, dan perilaku. Emosi dapat mencakup berbagai kategori seperti kemarahan, kebahagiaan, kesedihan, dan ketakutan, yang dapat mempengaruhi cara seseorang berinteraksi dengan lingkungan dan orang lain (Mao et al., 2019). Emosi manusia juga memiliki kaitan erat dengan kondisi kesehatan emosional yang juga berhubungan terhadap kondisi mental, tingkat stres, dan kondisi-kondisi lain pada kejiwaan seseorang. Hal tersebut berkaitan karena informasi yang dapat diambil dari emosi manusia cukup penting untuk tujuan diagnostik atau intervensi terapi.

### 2.3. Gammatone-Frequency Cepstral Coefficient (GFCC)

Gammatone-Frequency Cepstral Coefficient (GFCC) merupakan metode yang digunakan untuk ekstraksi fitur pada penelitian ini. GFCC merupakan metode yang digunakan dalam analisis sinyal suara dan pengolahan suara, terutama dalam pengenalan suara, pemrosesan ucapan, dan tugas-tugas terkait suara lainnya. GFCC didasarkan pada transformasi *gammatone* yang merupakan pendekatan untuk merepresentasikan sinyal suara dengan cara yang mirip dengan cara manusia mendengar. Ini memungkinkan ekstraksi informasi yang berguna dari sinyal suara, terutama dalam hal representasi fitur yang lebih baik daripada representasi domain waktu atau domain frekuensi mentah. Alur ekstraksi fitur menggunakan metode GFCC dapat dilihat seperti pada Gambar 1 berikut.



Gambar 1. Diagram Alir GFCC

#### A. Pre-emphasis

Tahap ini merupakan tahapan paling awal

sebagai proses *pre processing* dan filterisasi suara jantung. Pada tahap *pre-emphasis*, dilakukan penyaringan terhadap sinyal suara dengan cara mengurangi nilai frekuensi sinyal tersebut. Akibatnya, hanya sinyal yang memiliki frekuensi tinggi yang dapat melewati proses penyaringan. Tujuannya adalah untuk mengurangi kebisingan dari suara sehingga hanya data sinyal suara yang sebenarnya yang dapat diakuisisi oleh sistem (Fitria et al., 2021).

$$p(n) = s(n) - xs(n - 1) \quad (1)$$

### B. Framing

*Framing* merupakan proses analisis sinyal ucapan yang mengubahnya menjadi bentuk *frame*. Dalam konteks ini, setiap *frame* direpresentasikan oleh fitur tunggal yang tergambar dalam spektrum rata-rata untuk rentang waktu di dalam *frame*, biasanya berkisar antara 20 hingga 40 milidetik. *Frame* diambil sebanyak mungkin untuk memperoleh resolusi frekuensi yang optimal, sementara waktu yang sependek mungkin dimaksudkan untuk mendapatkan kualitas ranah waktu yang terbaik.

### C. Windowing

*Windowing* digunakan untuk mengurangi ketidakteraturan sinyal di awal dan akhir setiap *frame* akibat pengelompokan *frame*. *Windowing* dilakukan pada setiap bagian sinyal yang terbuat dari proses *framing*. Penelitian ini menggunakan proses *hamming window* dengan persamaan sebagai berikut.

$$w(n) = 0,54 + 0,46 \cos \frac{2\pi n}{N-1}, 0 \leq n \leq N - 1 \quad (2)$$

### D. Fast Fourier Transform (FFT)

FFT merupakan metode yang digunakan untuk mengubah sinyal suara menjadi domain frekuensi atau yang biasa disebut sebagai spectrogram. Metode ini bertujuan untuk mengurangi perhitungan digital dan menyederhanakan perhitungan spektrum frekuensi pada implementasinya. Metode ini menerapkan algoritma yang beroperasi pada sinyal diskrit (Prasetio, 2023).

$$f(n) \sum_{k=0}^{N-1} w_k e^{-\frac{2\pi jkn}{N}}, 0 \leq n \leq N - 1 \quad (3)$$

### E. Gammatone Filterbanks

Tahapan ini merupakan bentuk peniruan koklea telinga manusia dalam melakukan pemrosesan suara. Pada proses ini sinyal suara yang telah diproses dalam FFT menjadi domain

frekuensi menjadi domain waktu-frekuensi (Prasetio, 2023).

$$g(t) = a^{n-1} e^{-2\pi bt} \cos(2\pi f_c t + \emptyset) \quad (4)$$

Dengan  $a$  adalah amplitudo dan  $n$  adalah urutan dari *filter*, kemudian  $b$  adalah durasi respon impuls atau lebar *filter*, lalu  $f_c$  adalah nilai frekuensi filter tengah,  $\emptyset$  merupakan fase, serta  $t$  merupakan waktu.

### F. Cubic Root

Sinyal *cepstral* hasil transformasi biasanya memiliki banyak dimensi yang dapat menyebabkan *overhead* pada sistem. Operasi *cubic root* dapat membantu mengurangi dimensi sinyal *cepstral* dengan cara mengambil akar tiga dari setiap elemen sinyal tersebut. Dengan demikian, dimensi sinyal *cepstral* akan berkurang sehingga proses pengolahan sinyal menjadi lebih cepat dan sistem menjadi lebih ringan.

$$G_m[i] = |g_{downSampled}[i, m]|^{\frac{1}{3}}, i = 0 \dots N - 1, m = 0 \dots M - 1 \quad (5)$$

Operasi *cubic root* biasanya dilakukan dengan tujuan untuk mengurangi dimensi sinyal *cepstral* hasil transformasi yang berdampak terhadap kecepatan proses olah sinyal serta mengurangi *overhead*.

### G. Discrete Cosine Transform

Tahapan terakhir yaitu tahapan DCT hasil yang telah diproses sebelumnya. Hal ini bertujuan untuk melakukan dekorelasi dan juga mengurangi dimensi fitur yang telah dihasilkan (Helmiyah et al., 2019).

$$C_j = \sqrt{\frac{2}{N}} \sum_{i=0}^{M-1} G_m[i] \cos(j (i - \frac{1}{2}) \frac{\pi}{M}), j = 0, 1, 2, \dots, N - 1 \quad (6)$$

## 2.4 Hidden Markov Model (HMM)

Hidden Markov Model (HMM) adalah model probabilistik yang digunakan untuk menggambarkan sistem yang memiliki keadaan tersembunyi dan menghasilkan urutan observasi. Dalam konteks pengenalan emosi, HMM digunakan untuk memodelkan proses di mana urutan emosi berdasarkan BPM (tinggi dan rendah) dapat diwakili oleh dua keadaan tersembunyi. HMM beroperasi dengan mendefinisikan probabilitas transisi antara keadaan dan probabilitas emisi dari simbol



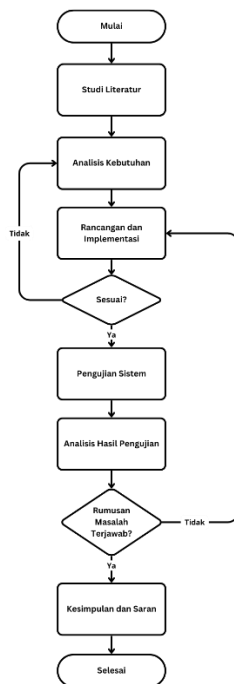
observasi yang terkait dengan setiap keadaan. Model ini memungkinkan analisis urutan data yang kompleks, seperti sinyal fisiologis, dengan memanfaatkan struktur probabilistik untuk meningkatkan akurasi klasifikasi emosi (Zamani et al., 2024).

### 3. METODOLOGI PENELITIAN

#### 3.1. Tipe Penelitian

Penelitian ini merupakan penelitian implementatif pengembangan. Penelitian implementatif merupakan sebuah penelitian yang tidak hanya melakukan analisis saja, namun terdapat penerapan yang dilakukan menggunakan perangkat lunak atau perangkat keras dalam mencapai solusi permasalahannya. Sedangkan penelitian pengembangan merupakan penelitian yang bertujuan untuk menghasilkan sesuatu yang baru yang dapat dikembangkan dari penelitian terdahulu.

#### 3.2. Tahapan Penelitian



Gambar 2. Diagram Alir Penelitian

Tahapan pertama merupakan studi literatur yang dilakukan untuk menggali secara mendalam terkait permasalahan, metode, dan teori ilmiah yang sesuai dengan penelitian. Berikutnya adalah tahap analisis kebutuhan yang dilakukan setelah pemahaman teori dari studi literatur.

Setelah melakukan analisis kebutuhan, peneliti akan melakukan perancangan sistem dan

implementasi sistem. Apabila telah sesuai, tahapan berikutnya adalah pengujian terhadap sistem, namun jika belum sesuai, maka akan kembali dilakukan proses perancangan dan implementasi sistem. Setelah tahap rancangan dan implementasi selesai, diperlukan tahap pengujian yang dilakukan dengan proses uji coba alat untuk mengukur berapa nilai akurasi deteksi yang didapatkan. Pada tahap ini juga akan dilakukan analisis terhadap hasil uji yang didapatkan. Penulis akan menganalisa hasil pengujian apakah sudah sesuai dan menjawab rumusan masalah yang ada.

Tahap terakhir adalah tahap penulisan kesimpulan yang berisi kesimpulan dari hasil penelitian yang dilakukan. Dalam kesimpulan ini pula penulis akan menuliskan apakah penelitian ini berhasil atau tidak. Setelah hasil dari penelitian telah disimpulkan, berikutnya penulis akan menuliskan saran yang dapat dilakukan untuk pengembangan penelitian berikutnya.

### 4. REKAYASA KEBUTUHAN

#### 4.1. Kebutuhan Fungsional

1. Sistem dapat menerima masukan dalam bentuk suara jantung.
2. Sistem dapat mengambil fitur melalui suara jantung menggunakan algoritma GFCC dan dapat melakukan klasifikasi dengan masukan tersebut menggunakan algoritma Hidden Markov Model.
3. Sistem terintegrasi dengan baik antara perangkat keras dengan perangkat lunak.
4. Sistem dapat menampilkan hasil klasifikasi emosi pada layar *smartphone*.

#### 4.2. Spesifikasi Sistem

1. Sistem membutuhkan integrasi perangkat keras dan perangkat lunak dalam penggunaan. Perangkat keras yang digunakan adalah stetoskop elektronik dan android *smartphone*.
2. Perangkat android yang digunakan harus memiliki port audio jack atau menggunakan adapter jika tidak tersedia untuk terhubung dengan stetoskop elektronik yang merekam suara jantung.

3. Perangkat android yang digunakan minimum sudah menggunakan android 9.0 untuk mendukung *library* pemrosesan audio modern.
4. Aplikasi yang dibuat harus mampu menangkap dan merekam masukan yang diterima oleh stetoskop elektronik.
5. Menggunakan bahasa pemrograman yang mendukung pengembangan *mobile* seperti java atau kotlin dalam pembuatan aplikasi.

digunakan untuk proses pengembangan dalam membuat aplikasi android.

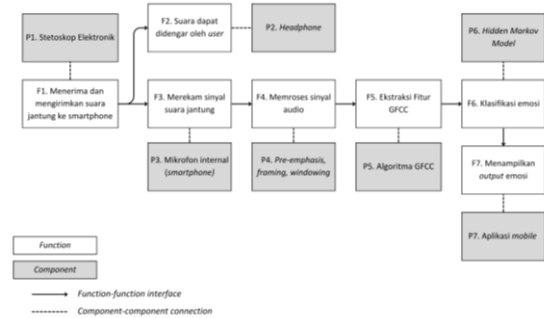
### 4.3. Analisis Kebutuhan Perangkat Keras dan Perangkat Lunak

Tabel 1. Analisis Kebutuhan Perangkat Keras dan Perangkat Lunak

Jenis Perangkat	Kebutuhan Sistem	Spesifikasi yang Memenuhi
Perangkat Keras	Sistem membutuhkan perangkat android yang memiliki <i>port</i> audio jack untuk terhubung dengan stetoskop elektronik dan minimum sudah menggunakan android 9.0 untuk mendukung pemrosesan audio modern.	<b>Android Smartphone:</b> OS: Android 14 (Untuk hasil optimal) CPU: Octa-core (2x2.0 GHz Cortex-A75 & 6x1.8 GHz Cortex-A55) GPU: Mali-G52 MC2 <i>Memory</i> : 6 GB <i>Storage</i> : 128 GB <i>Port</i> : Tipe-C 2.0, USB On-The-Go, audio jack.
	Sistem membutuhkan perangkat yang dapat menerima dan merekam suara jantung dan dapat terhubung ke <i>smartphone</i> untuk pemrosesan.	<b>Stetoskop elektronik:</b> Stetoskop elektronik dengan kabel audio jack <b>Headphone:</b> <i>Headphone</i> yang <i>compatible</i> dengan <i>iport</i> audio jack untuk dapat disambungkan pada stetoskop.
Perangkat Lunak	Sistem menggunakan bahasa pemrograman yang menyediakan <i>library</i> untuk program identifikasi emosi. Sistem menggunakan IDE yang	<b>Python</b>  <b>Android Studio</b>

## 5. PERANCANGAN SISTEM

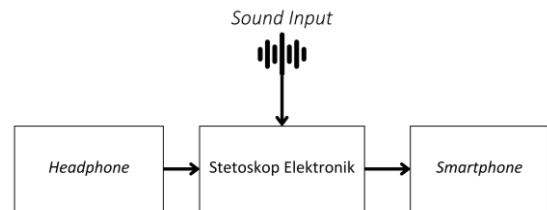
Perancangan sistem ini meliputi proses perancangan perangkat lunak hingga integrasi pada perangkat keras. Proses perancangan sistem menggambarkan alur kerja sistem secara keseluruhan yang digambarkan dalam diagram blok arsitektur alokasi sebagai berikut.



Gambar 3. Diagram Blok Arsitektur Alokasi

### 5.1. Perancangan Perangkat Keras

Pada perancangan sistem ini diperlukan 3 perangkat keras utama yaitu stetoskop elektronik, *headphone*, dan *smartphone*. Suara yang masuk melalui stetoskop digital akan terdengar melalui *headphone*. Suara yang masuk dan kemudian diterima oleh *smartphone* akan terekam untuk selanjutnya diproses pada aplikasi deteksi emosi. Proses perancangan perangkat keras dapat dilihat pada diagram blok berikut.

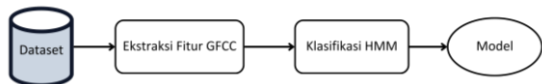


Gambar 4. Diagram Blok Perancangan Perangkat Keras

### 5.2. Perancangan Ekstraksi dan Klasifikasi

Perancangan Ekstraksi dan Klasifikasi Suara ini mencakup langkah-langkah teknis untuk mengolah suara jantung dan mengidentifikasi emosi yang terkandung di dalamnya. Dari dataset yang ada, proses dimulai dengan mengambil fitur yang diperlukan dengan metode GFCC kemudian dilakukan klasifikasi dengan

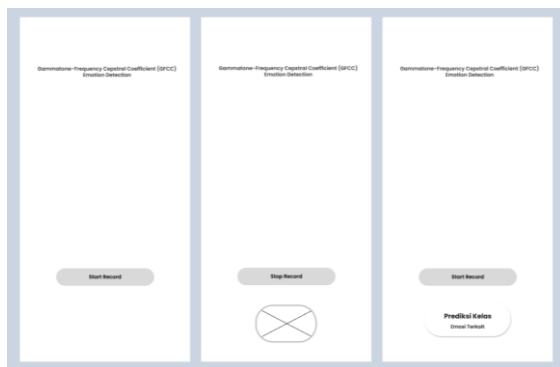
HMM. Hasilnya kemudian akan masuk ke dalam sistem untuk ditampilkan.



Gambar 5. Diagram Alir Perancangan Ekstraksi dan Klasifikasi

### 5.3. Perancangan Aplikasi

Perancangan aplikasi dimulai dengan perancangan alur penggunaan dan antarmuka pengguna aplikasi sederhana. Aplikasi ini dirancang khusus untuk mengidentifikasi emosi sehingga pada penggunaannya, pengguna langsung diarahkan pada halaman perekaman suara untuk proses identifikasi atau klasifikasi emosi.



Gambar 6. Wireframe Aplikasi

## 6. IMPLEMENTASI SISTEM

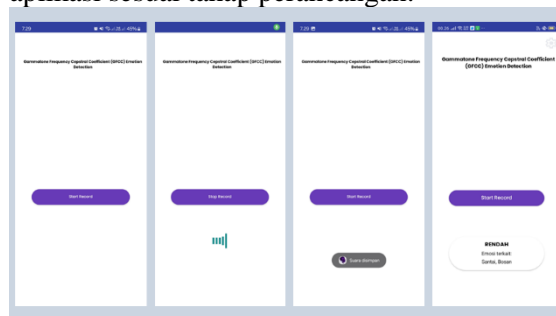
Implementasi sistem dilakukan berdasarkan perancangan yang telah dijelaskan sebelumnya. Berdasarkan perancangan perangkat keras untuk perekaman suara jantung, *headphone* dan stetoskop elektronik yang telah saling tersambung berikutnya disambungkan ke perangkat android atau *smartphone* melalui *port* audio.



Gambar 7. Implementasi Perangkat Keras

Berikutnya adalah implementasi ekstraksi dan klasifikasi yang dilakukan dengan proses ekstraksi GFCC untuk dataset *training*, kemudian hasilnya akan dilakukan pelatihan model menggunakan HMM. Hasil dari model HMM inilah yang selanjutnya akan dipanggil pada aplikasi android untuk diimplementasikan.

Proses implementasi aplikasi meliputi implementasi server dan implementasi antarmuka pengguna aplikasi. Implementasi server pada aplikasi ini digunakan sebagai pengolahan data atau komunikasi dengan *database* serta penyediaan API untuk aplikasi *frontend*. Implementasi antarmuka pengguna didesain dan dibuat sesuai dengan kebutuhan aplikasi sesuai tahap perancangan.



Gambar 8. Implementasi Antarmuka Pengguna

## 7. PENGUJIAN

Setelah penjelasan mengenai perancangan dan implementasi, pada bagian ini akan dijelaskan mengenai hasil pengujian sistem berdasarkan implementasi yang telah dilakukan.

### 7.1. Hasil Pengujian

#### A. Pengujian *Signal-to-Noise Ratio* (SNR)

Pengujian *Signal-to-Noise Ratio* (SNR) dilakukan untuk menguji sejauh mana ekstraksi GFCC dapat meredam dan mengurangi *noise* dalam sinyal dan mempertahankan informasi penting. SNR merupakan rasio yang digunakan untuk mengukur perbandingan antara *signal* yang diinginkan dan *noise* yang tidak diinginkan.

$$SNR (dB) = 10 \log_{10} \left( \frac{S}{N} \right) \tag{7}$$

Tabel 2. Hasil Uji SNR

Sampel (wav)	<i>Signal-to-Noise Ratio</i> (dB)	
	Asli	GFCC
RNDH_66_A_TEST	21.94	30.28

RNDH_66_B_TEST	20.87	30.23
RNDH_78_A_TEST	25.82	40.84
RNDH_78_B_TEST	30.42	36.64
TGG_93_A_TEST	23.72	32.66
TGG_93_B_TEST	18.93	29.53
TGG_105_TEST	12.29	16.90
TGG_117_TEST	25.51	31.26
TGG_120_TEST	10.82	22.00
TGG_123_A_TEST	24.09	30.04
TGG_123_B_TEST	23.38	20.18
TGG_126_A_TEST	15.77	27.38
TGG_129_TEST	15.65	25.37
TGG_132_TEST	13.05	24.75
TGG_156_TEST	33.98	40.10

TGG_93_B_TES T	TINGGI	RENDA H	SALAH
TGG_105_TEST	TINGGI	TINGGI	BENAR
TGG_117_TEST	TINGGI	RENDA H	SALAH
TGG_120_TEST	TINGGI	TINGGI	BENAR
TGG_123_A_TE ST	TINGGI	TINGGI	BENAR
TGG_123_B_TE ST	TINGGI	RENDA H	SALAH
TGG_126_A_TE ST	TINGGI	TINGGI	BENAR
TGG_129_TEST	TINGGI	TINGGI	BENAR
TGG_132_TEST	TINGGI	TINGGI	BENAR
TGG_156_TEST	TINGGI	TINGGI	BENAR

Berdasarkan hasil pengujian, dapat diketahui bahwa pengaruh GFCC dalam mengurangi atau meredam *noise* dalam sinyal adalah baik. Dari 15 sampel yang diuji, 14 diantaranya memiliki SNR yang lebih tinggi setelah melalui ekstraksi GFCC. Artinya, dalam pengujian ini GFCC 93.33% efektif dalam mengurangi *noise* pada sinyal suara jantung.

**B. Pengujian Akurasi Ekstraksi GFCC dan Klasifikasi HMM**

Pengujian akurasi dilakukan untuk menguji tingkat akurasi dari kinerja aplikasi identifikasi emosi. Pengujian dilakukan dengan menghitung data benar dan data salah dalam pengujian klasifikasi. Perhitungan pengujian akurasi dihitung dengan persamaan sebagai berikut.

$$Akurasi = \frac{TP+TN}{TP+TN+FP+FN} \tag{8}$$

Tabel 3. Hasil Uji Akurasi Identifikasi Emosi

Sampel	Label	Hasil Deteksi	Keterangan
RNDH_66_A_TE ST	RENDA H	RENDA H	BENAR
RNDH_66_B_TE ST	RENDA H	RENDA H	BENAR
RNDH_78_A_TE ST	RENDA H	RENDA H	BENAR
RNDH_78_B_TE ST	RENDA H	RENDA H	BENAR
TGG_93_A_TES T	TINGGI	RENDA H	SALAH

Berdasarkan data yang didapatkan diketahui bahwa sistem dapat mengidentifikasi secara benar 11 dari 15 sampel yang ada sehingga berdasarkan perhitungan dapat diperoleh akurasi ekstraksi dan klasifikasi sebesar 73.33%.

**8. KESIMPULAN DAN SARAN**

**8.1. Kesimpulan**

Kesimpulan yang ditulis akan menjawab rumusan masalah berdasarkan implementasi dan pengujian yang telah dilakukan.

1. Ekstraksi Gammatone-frequency Cepstral Coefficient (GFCC) memiliki pengaruh yang baik dan menjadi metode yang efektif dengan 93.33% menghasilkan SNR lebih tinggi dibandingkan sinyal suara asli dalam mengurangi *noise* pada sinyal suara jantung. Proses ekstraksi dengan GFCC terbukti mampu menerima ciri khas dari suara jantung seperti frekuensi rendah. Namun, efektivitas metode ini sangat bergantung pada kualitas dataset suara jantung yang digunakan, pengaturan parameter dalam ekstraksi, dan jumlah maupun variasi data pelatihan.
2. Hasil pengujian ekstraksi dan klasifikasi menunjukkan bahwa akurasi dari penggunaan ekstraksi GFCC dan klasifikasi HMM pada penelitian memiliki akurasi yang cukup tinggi yaitu 75% pada pelatihannya dan 73.33% pada pengujian. Dengan hasil tersebut dapat disimpulkan bahwa model sudah cukup mampu memprediksi dan mempelajari data, namun



masih memerlukan perbaikan untuk meningkatkan generalisasi dan kemampuan prediksi terutama untuk kategori dengan kesalahan prediksi yang lebih tinggi.

## 8.2. Saran

Saran merupakan masukan atau evaluasi yang perlu dilakukan terkait sistem untuk penelitian berikutnya.

1. Menambahkan langkah/tahapan atau algoritma lain pada tahap *preprocessing* supaya fitur yang diambil dapat lebih optimal untuk dilatih.
2. Melakukan penambahan data latih agar dapat menghasilkan model yang lebih baik atau melakukan augmentasi data yang lebih bervariasi untuk meningkatkan generalisasi model dan menghindari kesalahan prediksi yang tinggi pada salah satu kelas.

## 9. DAFTAR PUSTAKA

- Aini, Y. K., Santoso, T. B., & Dutono, T. (2021). Pemodelan CNN Untuk Deteksi Emosi Berbasis Speech Bahasa Indonesia. *Jurnal Komputer Terapan*, 7(1), 143–152. <https://doi.org/10.35143/jkt.v7i1.4623>
- Aljufri, M. N., & Prasetyo, B. H. (2022). Sistem Deteksi Tingkat Stress Menggunakan Suara dengan Metode Jaringan Saraf Tiruan dan Ekstraksi Fitur MFCC berbasis Raspberry Pi. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 6(11), 5278–5285. <https://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/11842>
- Akour, M. (2020). Mobile Voice Recognition Based for Smart Home Automation Control. *International Journal of Advanced Trends in Computer Science and Engineering*, 9(3), 3788–3792. <https://doi.org/10.30534/ijatcse/2020/196932020>
- Bimo, A., Kombinasi, G., Multi, F., Bimo Gumelar, A., Yuniarno, E. M., Anggraeni, W., Sugiarto, I., Kristanto, A. A., & Purnomo, M. H. (2020). Kombinasi Fitur Multispektrum Hilbert dan Cochleagram untuk Identifikasi Emosi Wicara (Spectrum Features Combination of Hilbert and Cochleagram for Speech Emotions Identification). *Jurnal Nasional Teknik Elektro Dan Teknologi Informasi*, 9(2), 180–189.
- Deng, M., Meng, T., Cao, J., Wang, S., Zhang, J., & Fan, H. (2020). Heart sound classification based on improved MFCC features and convolutional recurrent neural networks. *Neural Networks*, 130, 22–32. <https://doi.org/10.1016/j.neunet.2020.06.015>
- Egger, M., Ley, M., & Hanke, S. (2019). Emotion Recognition from Physiological Signal Analysis: A Review. *Electronic Notes in Theoretical Computer Science*, 343, 35–55. <https://doi.org/10.1016/j.entcs.2019.04.009>
- Fitria, L., Muttaqin, K., & Nasution, M. S. (2021). Implementasi Speech Recognition Pada Kata Kerja Dasar Menggunakan Metode MFCC. *J-ICOM-Jurnal Informatika Dan Teknologi Komputer*, 02(01), 43–50. <https://ejournalunsam.id/index.php/jicom/article/view/4076%0Ahttps://ejournalunsam.id/index.php/jicom/article/download/4076/2715>
- Harper, R., & Southern, J. (2022). A Bayesian Deep Learning Framework for End-To-End Prediction of Emotion from Heartbeat. *IEEE Transactions on Affective Computing*, 13(2), 985–991. <https://doi.org/10.1109/TAFFC.2020.2981610>
- Helmiyah, S., Fadlil, A., & Yudhana, A. (2019). Pengenalan Pola Emosi Manusia Berdasarkan Ucapan Menggunakan Ekstraksi Fitur Mel-Frequency Cepstral Coefficients (MFCC). *CogITO Smart Journal*, 4(2), 372–381. <https://doi.org/10.31154/cogito.v4i2.129.372-381>
- Jha, V., Prakash, N., & Sagar, S. (2018). Wearable anger-monitoring system. *ICT Express*, 4(4), 194–198. <https://doi.org/10.1016/j.ict.2017.07.002>
- Julian, T. S., Utaminingrum, F., & Syauqy, D. (2022). Sistem Voice Command pada Kursi Roda Pintar menggunakan MFCC dan CNN berbasis Jetson TX2. ... *Teknologi Informasi Dan Ilmu ...*, 6(11), 5505–5510. <https://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/11917%0Ahttp://j-ptiik.ub.ac.id/index.php/j-ptiik/article/download/11917/5290>
- Li, F., Liu, M., Zhao, Y., Kong, L., Dong, L., Liu, X., & Hui, M. (2019). Feature

- extraction and classification of heart sound using 1D convolutional neural networks. *Eurasip Journal on Advances in Signal Processing*, 2019(1). <https://doi.org/10.1186/s13634-019-0651-3>
- Muttaqin, D., & Suyanto, S. (2020). Speech Emotion Detection Using Mel-Frequency Cepstral Coefficient and Hidden Markov Model. *2020 3rd International Seminar on Research of Information Technology and Intelligent Systems, ISRITI 2020*, 463–466. <https://doi.org/10.1109/ISRITI51436.2020.9315433>
- Prasetyo, B. H. (2023). ... Melalui Ucapan menggunakan Ekstraksi Gammatone-Frequency Cepstral Coefficients dan Klasifikasi Random Forest Classifier berbasis Raspberry Pi 4. *Jurnal Pengembangan Teknologi Informasi Dan ...*, 7(2), 1003–1011. <https://jptiik.multi.web.id/index.php/jptiik/article/view/12364%0Ahttps://jptiik.ub.ac.id/index.php/jptiik/article/download/12364/5620>
- Renna, F., Oliveira, J., & Coimbra, M. T. (2019). Deep Convolutional Neural Networks for Heart Sound Segmentation. *IEEE Journal of Biomedical and Health Informatics*, 23(6), 2435–2445. <https://doi.org/10.1109/JBHI.2019.2894222>
- REVISITING HIDDEN MARKOV MODELS FOR SPEECH EMOTION RECOGNITION Shuiyang Mao , Dehua Tao , Guangyan Zhang , P . C . Ching and Tan Lee. (2019). 6715–6719.
- Shing-Tai Pan, Zong-Hong Huang, Sheng-Syun Yuan, Xu-Yu Li, Yu-De Su, & Jia-Hua Li. (2020). Application of Hidden Markov Models in Speech Command Recognition. *Journal of Mechanics Engineering and Automation*, 10(2), 41–45. <https://doi.org/10.17265/2159-5275/2020.02.001>
- Singh, P., Nayak, P., Datta, A., Sani, D., Raghav, G., & Tejpal, R. (2019). Voice Control Device using Raspberry Pi. *Proceedings - 2019 Amity International Conference on Artificial Intelligence, AICAI 2019, February*, 723–728. <https://doi.org/10.1109/AICAI.2019.8701409>
- Torad, M. A., Bouallegue, B., & Ahmed, A. M. (2022). A voice controlled smart home automation system using artificial intelligent and internet of things. *Telkomnika (Telecommunication Computing Electronics and Control)*, 20(4), 808–816. <https://doi.org/10.12928/TELKOMNIKA.v20i4.23763>
- Ullah, I., Ahmad, R., & Kim, D. H. (2018). A prediction mechanism of energy consumption in residential buildings using hidden markov model. *Energies*, 11(2), 1–20. <https://doi.org/10.3390/en11020358>
- Uma, S., Eswari, R., Bhuvanya, R., & Kumar, G. S. (2019). IoT based Voice/Text Controlled Home Appliances. *Procedia Computer Science*, 165(2019), 232–238. <https://doi.org/10.1016/j.procs.2020.01.085>
- Xiefeng, C., Wang, Y., Dai, S., Zhao, P., & Liu, Q. (2019). Heart sound signals can be used for emotion recognition. *Scientific Reports*, 9(1), 1–11. <https://doi.org/10.1038/s41598-019-42826-2>
- Xiyun Liu, Yang Yang, Lijun Yang, Xiaohui Yang, C. Z. (2024). *Heart Sound Classification Based on Improved GFCC Features and Multi-attention Fusion Convolutional NEural Network*.
- Zamani Khanghah, S., & Maghooli, K. (2024). Emotion Recognition From Heart Rate Variability Using a Hybrid System Combined With a Hidden Markov Model and Poincare Plot. *Applied Computer Science*, 20(1), 106–121. <https://doi.org/10.35784/acs-2024-07>