

# **ANALISIS LOGAT BICARA MAHASISWA UNIVERSITAS PERTAMINA MENGGUNAKAN FITUR MEL-FREQUENCY CEPSTRAL COEFFICIENTS (MFCC)**

**<sup>1</sup>Alghifari Rasyid Zola, <sup>2</sup>Bambang Istijab, <sup>3</sup>Fauzan Azhima**

Email : <sup>1</sup>[alghifarirasyidzola@gmail.com](mailto:alghifarirasyidzola@gmail.com), <sup>2</sup>[bambangistjab30@gmail.com](mailto:bambangistjab30@gmail.com),  
<sup>3</sup>[fauzannazhimaa@gmail.com](mailto:fauzannazhimaa@gmail.com)

<sup>1,2,3</sup> Mahasiswa Program Studi Ilmu Komputer Universitas Pertamina, Jakarta Selatan, Indonesia

---

## **ABSTRAK**

Mel-Frequency Cepstral Coefficients (MFCC) adalah representasi numerik dari karakteristik suara yang sering digunakan dalam pengenalan pola suara. Penelitian ini bertujuan untuk menganalisis pola suara logat bahasa Indonesia menggunakan ekstraksi MFCC, yang diambil dari data suara peserta dari lima daerah: Timur, Jawa, Medan, Sunda, dan Minang. Setiap daerah diwakili oleh dua partisipan yang masing-masing menyumbangkan 10 sampel suara, menghasilkan total 100 sampel yang dianalisis. Data suara ini kemudian diekstraksi menjadi matriks MFCC untuk menggambarkan pola frekuensi suara yang unik untuk setiap logat. Hasil ekstraksi MFCC menunjukkan variasi dimensi yang mencerminkan perbedaan dalam gaya bicara, tempo, dan intonasi setiap logat. Dimensi rata-rata MFCC dari setiap logat menunjukkan perbedaan signifikan dalam durasi dan pengucapan vokal. Visualisasi MFCC memperjelas perbedaan distribusi energi pada rentang frekuensi tertentu untuk setiap logat. Penelitian ini juga menyediakan data yang dapat digunakan untuk melatih model machine learning guna membedakan logat bahasa Indonesia berdasarkan karakteristik akustiknya, memberikan kontribusi dalam bidang pengenalan suara dan pengolahan bahasa alami.

**Kata Kunci :** *MFCC, logat bahasa Indonesia, ekstraksi suara, pengenalan pola suara, visualisasi suara*

## ***ABSTRACT***

*Mel-Frequency Cepstral Coefficients (MFCC) are numerical representations of voice characteristics that are often used in speech pattern recognition. This study aims to analyze Indonesian accent voice patterns using MFCC extraction, which is taken from the voice data of participants from five regions: Eastern, Javanese, Medan, Sundanese, and Minang. Each region was represented by two participants who each contributed 10 voice samples, resulting in a total of 100 samples being analyzed. This voice data was then extracted into MFCC matrices to describe the unique voice frequency patterns for each accent. The MFCC extraction results show dimensional variations that reflect differences in the speech style, tempo and intonation of each accent. The average MFCC dimensions of each accent show significant differences in vowel duration and pronunciation. Visualization of the MFCCs made clear the differences in energy distribution at specific frequency ranges for each accent. This research also provides data that can be used to train machine learning models to distinguish Indonesian accents based on their acoustic characteristics, contributing to the field of speech recognition and natural language processing.*

***Keywords :*** MFCC, Indonesian accent, speech extraction, speech pattern recognition, speech visualization

---

## BAB 1 PENDAHULUAN

### 1. Latar Belakang

Indonesia merupakan negara dengan tingkat keanekaragaman budaya yang tinggi, hal ini tercermin dari perbedaan bahasa dan logat yang terdapat di setiap daerah. Lebih dari 700 bahasa daerah digunakan di kepulauan Indonesia, yang masing-masing memiliki keunikan dan kekayaan budayanya sendiri, termasuk dialek dan aksen yang menjadi ciri khas masyarakat setempat. Perbedaan-perbedaan ini tidak hanya mencerminkan asal usul geografis tetapi juga aspek sosial dan budaya lainnya. Oleh karena itu, penting untuk memahami perbedaan aksen daerah dalam bidang teknologi, komunikasi, dan linguistik [18].

Keberagaman budaya ini mencetuskan semboyan Bhineka Tunggal Ika, yang berarti berbeda-beda tetapi tetap satu. Meskipun terdapat berbagai suku, budaya, agama, dan golongan, Indonesia tetap bersatu sebagai satu kesatuan [18]. Dengan lebih dari 1.340 kelompok suku bangsa dan beragam bahasa serta tradisi, keberagaman ini merupakan kekayaan yang unik dan tidak dimiliki oleh negara lain. Variasi ini tidak hanya mencerminkan asal-usul geografis, tetapi juga aspek sosial dan budaya lainnya, sehingga pemahaman tentang perbedaan logat daerah sangat penting dalam bidang teknologi, komunikasi, dan linguistik.

Salah satu tantangan dalam pengenalan suara (speech recognition) adalah variasi logat dan aksen. Sistem pengenalan suara biasanya dirancang untuk mengenali satu bentuk standar dari suatu bahasa, namun kenyataannya, suara manusia sangat bervariasi tergantung pada latar belakang geografis, sosial, dan pendidikan. Oleh karena itu, pengembangan sistem yang mampu mengenali suara dengan berbagai logat menjadi penting untuk meningkatkan akurasi dan penerimaan teknologi ini oleh masyarakat yang beragam [17].

Mel-Frequency Cepstral Coefficients (MFCC) adalah salah satu metode yang sering digunakan untuk ekstraksi fitur suara. MFCC memberikan representasi numerik dari karakteristik suara yang relevan untuk pengenalan pola dan telah terbukti efektif dalam menangkap fitur suara manusia yang beragam. Dengan mengubah sinyal suara menjadi fitur MFCC, analisis dan klasifikasi suara berdasarkan logat, aksen, atau identitas pembicara dapat dilakukan dengan lebih mudah [10][3].

Langkah penting pertama dalam penelitian ini adalah mengumpulkan data audio dari siswa dengan aksen berbeda dari berbagai daerah di Indonesia, termasuk Timur, Jawa, Sunda, Minang, dan Batak. Data audio ini diolah menggunakan teknologi MFCC untuk menganalisis perbedaan karakteristik audio yang terkait dengan aksen daerah tertentu. Penelitian sebelumnya menunjukkan bahwa penggunaan MFCC dalam identifikasi dialek dapat meningkatkan akurasi pengenalan suara[2]. Penelitian ini diharapkan dapat memberikan kontribusi terhadap pengembangan teknologi pengenalan suara yang lebih komprehensif dan dapat mengenali perubahan aksen dan aksen dengan akurasi yang tinggi.

## **2. Tujuan Penelitian**

Penelitian ini bertujuan untuk mengumpulkan, menganalisis, dan mengklasifikasikan data suara mahasiswa yang berasal dari berbagai daerah di Indonesia, dengan fokus pada variasi logat dan aksen khas dari daerah-daerah tertentu. Tujuan spesifik dari penelitian ini adalah sebagai berikut:

### **1. Mengumpulkan Data Suara dari Mahasiswa dengan Logat Khas**

Penelitian ini bertujuan untuk mengumpulkan data suara dari mahasiswa yang berasal dari berbagai daerah di Indonesia, seperti Papua, Jawa, Betawi, Sunda, dan Batak. Mahasiswa yang terpilih harus memiliki logat atau aksen yang khas, sehingga dapat merepresentasikan perbedaan logat yang ada di masing-masing daerah[18].

### **2. Mengekstraksi Fitur Suara Menggunakan Metode MFCC**

Setelah data suara terkumpul, penelitian ini bertujuan untuk melakukan ekstraksi fitur suara menggunakan teknik Mel-Frequency Cepstral Coefficients (MFCC). MFCC akan digunakan untuk mengubah sinyal suara menjadi fitur numerik yang dapat mewakili karakteristik suara manusia yang relevan dalam pengenalan logat dan aksen[18].

### **3. Menganalisis Perbedaan Logat Berdasarkan Fitur MFCC**

Penelitian ini bertujuan untuk menganalisis perbedaan logat dan aksen antara mahasiswa yang berasal dari daerah yang berbeda. Dengan menggunakan fitur MFCC, diharapkan dapat ditemukan pola atau perbedaan karakteristik suara yang mencerminkan logat daerah masing-masing [17].

### **4. Memberikan Kontribusi pada Pengembangan Sistem Pengenalan Suara**

Salah satu tujuan utama dari penelitian ini adalah untuk memberikan kontribusi terhadap pengembangan sistem pengenalan suara yang lebih inklusif dan mampu mengenali variasi logat dan aksen dengan lebih baik. Diharapkan hasil dari penelitian ini dapat digunakan untuk meningkatkan akurasi dan efektivitas teknologi pengenalan suara di Indonesia, dengan mempertimbangkan keberagaman logat dan aksen yang ada [4].

### **5. Menjadi Dasar Pengembangan Penelitian Selanjutnya**

Penelitian ini juga bertujuan untuk membuka peluang bagi penelitian lebih lanjut dalam bidang pengenalan suara berbasis logat. Dengan data suara yang telah diproses, penelitian ini akan menjadi dasar yang berguna untuk pengembangan sistem yang lebih canggih dalam memahami perbedaan logat di Indonesia [14].

## **3. Ruang Lingkup**

Penelitian ini berfokus pada pengumpulan, ekstraksi, dan analisis data suara mahasiswa yang berasal dari lima daerah di Indonesia, yaitu logat Timur, Jawa, Medan, Sunda, dan Minang. Penelitian ini bertujuan untuk mengidentifikasi perbedaan logat dengan menggunakan teknik ekstraksi fitur Mel-Frequency Cepstral Coefficients (MFCC). Adapun ruang lingkup penelitian ini adalah sebagai berikut:

### **1. Sumber Data Suara**

Penelitian ini akan mengumpulkan data suara dari mahasiswa yang berasal dari lima daerah di Indonesia, yaitu logat Timur, Jawa, Medan, Sunda, dan

Minang. Data suara akan dikumpulkan dari mahasiswa yang memiliki logat khas yang dapat merepresentasikan aksen atau dialek dari masing-masing daerah tersebut.

## 2. Jenis Data yang Dikumpulkan

Data suara yang dikumpulkan berupa rekaman suara yang diucapkan oleh mahasiswa dalam bentuk kata atau kalimat dalam bahasa Indonesia. Mahasiswa akan diminta untuk mengucapkan kalimat atau kata yang umum digunakan sehari-hari, untuk mencerminkan penggunaan bahasa yang lebih alami. Penelitian ini hanya akan mencakup data suara dalam bahasa Indonesia dan tidak akan mencakup bahasa daerah selain bahasa Indonesia.

## 3. Metode Ekstraksi Fitur

Teknik ekstraksi fitur yang digunakan adalah Mel-Frequency Cepstral Coefficients (MFCC), yang digunakan untuk mengubah sinyal suara menjadi fitur numerik yang mewakili karakteristik suara manusia yang relevan untuk analisis logat. MFCC akan dihitung untuk setiap rekaman suara yang telah dikumpulkan dari lima logat daerah yang ditentukan.

## 4. Proses Pengolahan Data

Data suara yang telah dikumpulkan akan diproses dengan memastikan kualitas suara yang memadai. Proses ini mencakup pemotongan atau padding audio sesuai panjang standar yang telah ditentukan (22050 sampel atau sekitar 1 detik). Selanjutnya, fitur MFCC akan diekstraksi dari setiap file audio, dan analisis lebih lanjut akan dilakukan berdasarkan fitur tersebut untuk menganalisis perbedaan logat.

## 5. Analisis Perbedaan Logat

Penelitian ini berfokus pada analisis perbedaan logat dari lima daerah yang telah disebutkan di atas, yaitu Timur, Jawa, Medan, Sunda, dan Minang. Perbedaan logat akan dianalisis berdasarkan fitur MFCC yang diekstraksi. Hasil

analisis ini akan membantu mengidentifikasi ciri khas suara dari masing-masing logat yang dimiliki oleh mahasiswa dari setiap daerah.

## 6. Batasan Geografis

Penelitian ini hanya akan melibatkan mahasiswa yang berasal dari lima daerah yang telah disebutkan, yaitu Timur, Jawa, Medan, Sunda, dan Minang. Mahasiswa dengan logat yang tidak terlalu khas atau tidak terwakili oleh kelima daerah ini tidak akan termasuk dalam cakupan penelitian.

## 7. Aplikasi Teknologi

Meskipun penelitian ini bertujuan untuk menganalisis perbedaan logat, penelitian ini tidak akan mengimplementasikan sistem pengenalan suara berbasis logat. Fokus utama penelitian ini adalah pada pengumpulan data suara, ekstraksi fitur MFCC, dan analisis perbedaan logat berdasarkan ciri suara yang dihasilkan.

## BAB 2 TINJAUAN PUSTAKA

### A. Landasan Teori

#### 1. Logat dan Aksen

Logat atau aksen merujuk pada variasi dalam pengucapan kata yang dipengaruhi oleh faktor geografis, budaya, atau sosial. Di Indonesia, setiap daerah memiliki logat yang khas, yang sering kali dipengaruhi oleh bahasa daerah yang digunakan. Logat tidak hanya melibatkan pengucapan kata, tetapi juga intonasi, irama, dan tempo berbicara yang membedakan antara satu daerah dengan daerah lainnya.

- Logat Timur, misalnya, sering dikaitkan dengan cara pengucapan yang lebih cepat dan cenderung lebih datar dalam beberapa wilayah di Indonesia bagian timur. [17]
- Logat Jawa memiliki ciri khas dalam penggunaan intonasi dan pengucapan vokal yang berbeda [11].
- Logat Medan, di sisi lain, memiliki aksen yang lebih keras dan sering kali dipengaruhi oleh bahasa Melayu.
- Logat Sunda dan Logat Minang juga memiliki ciri khas masing-masing, dengan pengucapan yang lebih lembut dan melodius untuk Sunda, serta intonasi yang agak terputus-putus untuk Minang. [18]

#### 2. Suara dan Pengolahan Sinyal Suara

Pengenalan suara adalah teknologi yang digunakan untuk mengenali dan memproses suara manusia. Proses ini melibatkan beberapa tahap penting, mulai dari akuisisi sinyal suara, pemrosesan sinyal, hingga ekstraksi fitur yang relevan untuk analisis lebih lanjut.

- **Akuisisi Sinyal:** Proses ini mencakup perekaman suara menggunakan mikrofon atau perangkat perekam suara lainnya.
- **Pemrosesan Sinyal:** Tahap ini melibatkan pengolahan sinyal suara untuk menghilangkan noise, normalisasi, dan pemotongan sinyal untuk mendapatkan potongan yang relevan.

Sinyal suara manusia yang direkam sering kali berupa gelombang kontinu yang perlu diubah menjadi representasi yang dapat dianalisis oleh komputer. Salah satu teknik

umum yang digunakan untuk mengubah sinyal suara menjadi fitur numerik adalah **Ekstraksi Fitur**, yang membantu dalam mengidentifikasi pola-pola tertentu dalam suara.

### 3. Mel Frequency Cepstral Coefficient (MFCC)

Mel Frequency Cepstrum Coefficients (MFCC), merupakan salah satu metode yang banyak digunakan dalam bidang speech technology baik speaker recognition maupun speech recognition [12]. Mel Frequency Cepstrum Coefficients (MFCC) merupakan Algoritma atau algoritma yang mengubah ciri suara menjadi sebuah koefisien yang dapat mewakili ciri suara tersebut dan didasarkan pada perhitungan frekuensi suara yang dapat didengar oleh manusia atau audiosonik [1]. Algoritma ini digunakan untuk melakukan feature extraction, sebuah proses yang mengkonversikan signal suara menjadi beberapa parameter. Keunggulan dari metode ini adalah :

- a. Mampu untuk menangkap karakteristik suara yang sangat penting bagi pengenalan suara. Atau dengan kata lain mampu menangkap informasi-informasi penting yang terkandung dalam signal suara.
- b. Menghasilkan data seminimal mungkin, tanpa menghilangkan informasi-informasi penting yang ada.
- c. Mengadaptasi organ pendengaran manusia dalam melakukan persepsi terhadap signal suara.

Perhitungan yang dilakukan dalam MFCC [5] menggunakan dasar dari perhitungan short-term analysis. Hal ini dilakukan mengingat signal suara yang bersifat quasi stationary. Pengujian yang dilakukan untuk periode waktu yang cukup pendek (sekitar 10 sampai 30 milidetik) akan menunjukkan karakteristik signal suara yang stationary. Tetapi bila dilakukan dalam periode waktu yang lebih panjang karakteristik signal suara akan terus berubah sesuai dengan kata yang diucapkan. MFCC feature extraction sebenarnya merupakan adaptasi dari sistem pendengaran manusia dimana signal suara akan difilter secara linier untuk frekuensi rendah (dibawah 1000Hz) dan secara logitmik untuk frekuensi tinggi (diatas 1000 Hz). Berikut ini adalah blok diagram untuk MFCC.

### 4. Library Librosa

Librosa adalah pustaka analisis audio dan musik Python yang sangat berguna yang membantu pengembang perangkat lunak membangun aplikasi untuk bekerja dengan

format file audio dan musik menggunakan Python. Pustaka ini sangat mudah digunakan dan dapat menangani tugas-tugas dasar maupun lanjutan yang terkait dengan pemrosesan audio dan musik. Perpustakaan adalah open source dan tersedia secara bebas di bawah Lisensi ISC [13].

Perpustakaan menyediakan banyak fleksibilitas untuk pengguna ahli dan pada saat yang sama juga mengagumkan untuk pengguna pemula yang lebih tertarik untuk memproses file audio. Pustaka telah menyertakan dukungan untuk beberapa fitur penting yang terkait dengan pemrosesan dan ekstraksi file audio seperti memuat audio dari disk, menghitung berbagai representasi spektrogram, pemisahan sumber harmonik-perkus, dekomposisi spektrogram generik, memuat dan mendekode audio, pemrosesan audio domain waktu, pemodelan sekuensial, mengintegrasikan pemisahan harmonik-perkus, sinkron ketukan, dan banyak lagi.

## 5. Library Pydub AudioSegement

PyDub adalah pustaka Python yang digunakan untuk memproses dan memanipulasi file audio. Ini menyediakan antarmuka yang mudah digunakan untuk membaca, menulis, memotong, menggabungkan, dan melakukan berbagai operasi lain pada file audio. PyDub mendukung berbagai format audio, seperti MP3, WAV, FLAC, dan banyak lagi [18].

Dalam penelitian kali ini, library pydub digunakan sebagai alat bantu untuk memproses file audio mentah yang beragam formatnya menjadi format .wav yang lebih umum digunakan untuk analisis suara. Format .wav dipilih karena lebih kompatibel dengan berbagai library pemrosesan audio seperti librosa, yang akan digunakan untuk melakukan ekstraksi fitur MFCC (Mel Frequency Cepstral Coefficient).[13]

Spesifiknya, fungsi AudioSegment dari pydub digunakan untuk membaca file audio dalam format mentah seperti .mp3, .m4a, atau format lainnya, dan mengonversinya ke format .wav. Dengan menggunakan AudioSegment, penelitian ini dapat menangani proses konversi secara otomatis dan efisien tanpa mengubah kualitas suara asli. Langkah ini penting untuk memastikan data audio dapat diolah lebih lanjut tanpa kendala format.[20]

## 6. FFmpeg

FFmpeg adalah sebuah proyek perangkat lunak gratis dan sumber terbuka yang terdiri dari serangkaian pustaka dan program untuk menangani data multimedia. Proyek ini terutama dikenal dengan kemampuannya untuk mengonversi format audio dan video, tetapi FFmpeg juga memiliki berbagai fungsi lain yang sangat berguna dalam dunia multimedia [8].

Fungsi utama FFmpeg adalah mengonversi hampir semua format audio dan video yang ada. Contohnya, mengonversi file video dari format AVI ke MP4, atau file audio dari MP3 ke WAV. FFmpeg juga memungkinkan pengguna untuk mengompres file multimedia, sehingga ukuran file menjadi lebih kecil tanpa mengurangi kualitas secara signifikan [19].

## B. Studi Terkait

Beberapa penelitian sebelumnya yang relevan dengan topik ini antara lain:

1. **Penelitian oleh T. Nasution (2012)** yang menggunakan **MFCC** untuk mengenali ucapan dalam bahasa Indonesia. Penelitian ini menunjukkan bahwa **MFCC** efektif untuk mengekstraksi fitur dari sinyal suara dalam bahasa Indonesia, meskipun terdapat tantangan dalam mengenali berbagai aksen yang ada. Penelitian ini mengungkapkan potensi MFCC dalam meningkatkan akurasi sistem pengenalan suara untuk bahasa yang memiliki banyak logat, seperti bahasa Indonesia. Namun, tantangan yang dihadapi adalah keragaman logat antar daerah yang berpengaruh pada pengenalan suara secara keseluruhan.
2. **Penelitian oleh Ajinurseto, G., & Islamuddin, N. (2023)** yang membahas penerapan **MFCC** pada sistem pengenalan suara berbasis desktop. Penelitian ini mengimplementasikan **MFCC** dalam pengembangan sistem pengenalan suara untuk aplikasi desktop yang berfokus pada **identifikasi aksen dan pengucapan** dari berbagai daerah di Indonesia. Hasil penelitian menunjukkan bahwa **MFCC** dapat memberikan hasil yang cukup akurat dalam membedakan aksen, meskipun ada beberapa faktor eksternal seperti kebisingan latar belakang dan kualitas rekaman suara yang mempengaruhi hasil sistem. Penelitian ini juga menyoroti pentingnya pemilihan parameter yang tepat dalam ekstraksi fitur **MFCC**, seperti jumlah koefisien MFCC yang digunakan, untuk meningkatkan kinerja model pengenalan suara.

Kedua penelitian di atas memberikan gambaran tentang potensi dan tantangan dalam penerapan MFCC untuk pengenalan logat di Indonesia. Penelitian Nasution (2012) mengarah pada pengenalan suara dalam konteks umum, sedangkan penelitian Ajinurseto dan Islamuddin (2023) lebih fokus pada aplikasi desktop dan pengenalan aksen. Penelitian Anda berfokus pada perbandingan logat spesifik dari daerah seperti **Papua, Jawa, Medan, Sunda, dan Minang**, yang menambah kedalaman pemahaman mengenai perbedaan-perbedaan logat yang ada di Indonesia, serta cara MFCC dapat digunakan untuk membedakan logat tersebut dalam konteks yang lebih terperinci.

## BAB 3 METODOLOGI PENELITIAN

### A. Desain Penelitian

Penelitian ini menggunakan desain penelitian kuantitatif deskriptif yang bertujuan untuk menganalisis karakteristik logat dari lima daerah di Indonesia, yaitu Papua, Jawa, Medan, Sunda, dan Minang, berdasarkan fitur akustik yang dihasilkan dari rekaman suara. Pendekatan analisis sinyal suara digunakan dalam penelitian ini untuk mengekstraksi pola logat khas dari masing-masing daerah. Proses ekstraksi fitur dilakukan dengan memanfaatkan metode Mel-Frequency Cepstral Coefficients (MFCC), yang dikenal sebagai salah satu metode yang efektif dalam pemrosesan sinyal suara untuk menggambarkan karakteristik akustik.

Penelitian ini dimulai dengan pengumpulan data suara dari mahasiswa yang memiliki logat khas daerahnya. Para partisipan diminta untuk membaca teks standar yang telah disiapkan sebelumnya, sehingga data suara yang diperoleh memiliki konten yang seragam. Setelah data terkumpul, proses pengolahan data dilakukan dengan menerapkan metode MFCC untuk mengekstraksi fitur suara. Ekstraksi fitur ini bertujuan untuk mendapatkan representasi numerik dari karakteristik suara yang mencerminkan logat khas masing-masing daerah.

Hasil ekstraksi MFCC kemudian dianalisis untuk menemukan pola-pola unik dari logat setiap daerah. Analisis dilakukan secara deskriptif untuk memahami perbedaan akustik antara logat daerah yang diteliti. Hasil analisis akan divisualisasikan dalam bentuk grafik atau heatmap untuk mempermudah interpretasi. Penelitian ini diharapkan mampu memberikan wawasan baru terkait perbedaan karakteristik logat daerah di Indonesia, serta menjadi dasar untuk pengembangan sistem pengenalan suara berbasis logat daerah di masa mendatang.

### B. Metode Pengumpulan Data

#### 1. Peserta Penelitian

Peserta penelitian ini terdiri dari mahasiswa Universitas Pertamina yang berasal dari berbagai daerah di Indonesia, yaitu Papua, Jawa, Medan, Sunda, dan Minang. Kriteria pemilihan peserta didasarkan pada mahasiswa yang memiliki logat khas dan representatif dari daerah masing-masing. Peserta dipilih secara sengaja (purposive sampling) untuk memastikan bahwa logat yang dihasilkan dalam rekaman suara sesuai dengan logat yang ingin diteliti. Sebagai tambahan, peserta juga diminta untuk memiliki kemampuan berbicara yang jelas dan tidak memiliki gangguan bicara yang dapat memengaruhi hasil analisis.

## **2. Proses Perekaman**

Perekaman suara dilakukan dengan menggunakan recorder bawaan dari handphine. Durasi perekaman untuk setiap peserta adalah sekitar 10 hingga 12 detik, dengan meminta peserta untuk membaca teks standar yang telah disiapkan. Teks yang dibaca adalah teks dalam bahasa Indonesia yang dapat merepresentasikan variasi bunyi yang terdapat dalam logat masing-masing daerah. Adapun teks yang dibaca adalah sebagai berikut:

*"Perkenalkan nama saya (nama), saya berasal dari (daerah), saya dari prodi Ilmu Komputer Universitas Pertamina, nim saya (nim)".*

Perekaman dilakukan dalam ruangan yang tenang untuk menghindari gangguan suara luar (noise) yang dapat mengurangi kualitas rekaman.

## **3. Instruksi kepada Peserta**

Instruksi yang diberikan kepada peserta adalah sebagai berikut:

- Peserta diminta untuk duduk dengan posisi yang nyaman dan berbicara dengan jelas dan wajar, seolah-olah sedang berbicara kepada orang lain dalam percakapan biasa.
- Peserta diberi teks bacaan yang sudah disiapkan sebelumnya. Peserta diminta untuk membaca teks tersebut dengan lancar tanpa terburu-buru, tetapi tetap mempertahankan pengucapan yang alami sesuai dengan logat daerah mereka.
- Peserta diminta untuk menghindari gangguan suara seperti berbicara terlalu keras atau terlalu pelan. Selain itu, mereka juga diminta untuk menjaga kejelasan dan ketepatan dalam pengucapan setiap kata.
- Perekaman dilakukan dalam sepuluh kali sesi untuk setiap peserta, dan tidak ada pengulangan yang dilakukan pada rekaman untuk memastikan keaslian suara yang ditangkap.
- Total suara yang dikumpulkan adalah 100 file suara.

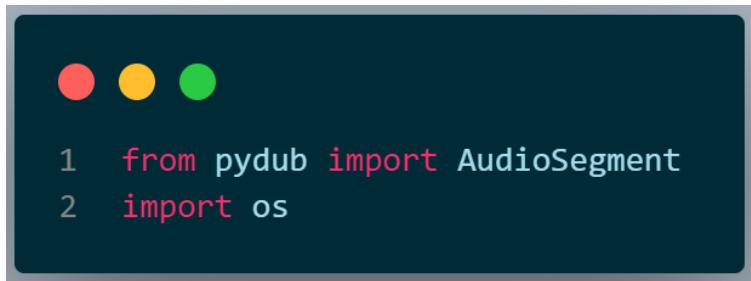
## **C. Proses Pengolahan Data**

### **1. Convert format ke WAV**

Program ini dirancang untuk mengonversi file audio dengan format audio tertentu ke format WAV secara otomatis. Program ini memanfaatkan library Pydub dan FFmpeg, program ini membaca file audio tertentu dari folder input dan mengonversinya ke format WAV, lalu menyimpan di folder output. FFmpeg berfungsi sebagai alat utama

untuk menagani konversi antar format audio. Program ini juga dapat mengoversi banyak file sekaligus dan menangani error jika terjadi masalah pada file tertentu.

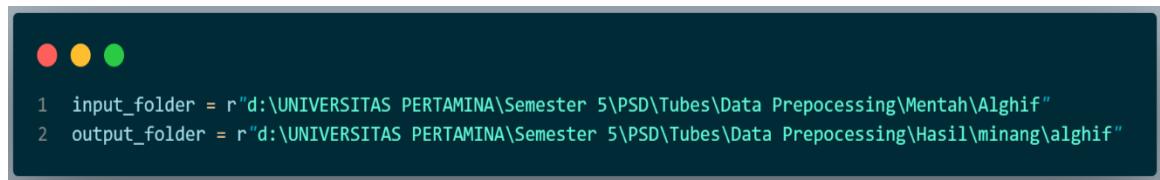
## 1. Library



```
● ● ●  
1 from pydub import AudioSegment  
2 import os
```

Gambar 1 Library

Program ini menggunakan dua library utama: **Pydub** dan **os**. **Pydub** digunakan untuk memanipulasi file audio, termasuk mengonversi file dari format M4A ke WAV dengan bantuan backend **FFmpeg**. Library **os** digunakan untuk menangani operasi file, seperti memeriksa apakah folder output sudah ada, membuat folder jika belum ada, dan mengelola path file. Kombinasi kedua library ini memungkinkan program untuk memproses banyak file audio secara otomatis dan menyimpannya di folder yang diinginkan.



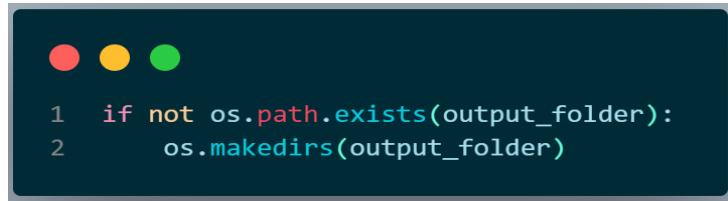
```
● ● ●  
1 input_folder = r"d:\UNIVERSITAS PERTAMINA\Semester 5\PSD\Tubes\Data Preprocessing\Mentah\Alghif"  
2 output_folder = r"d:\UNIVERSITAS PERTAMINA\Semester 5\PSD\Tubes\Data Preprocessing\Hasil\minang\alghif"
```

Gambar 2 Input dan Output Folder Path

## 2. Menetukan Lokasi Folder Input dan Output

Pada bagian awal ini, kita menentukan lokasi folder **input** dan **output**. Folder input adalah tempat di mana file audio dengan format M4A berada, dan folder output adalah tempat hasil konversi ke format WAV akan disimpan. Pengguna dapat mengganti path ini sesuai dengan folder di komputer mereka. `r` sebelum string digunakan untuk mendefinisikan raw string literal, yang memungkinkan kita menuliskan path dengan tanda backslash (\) tanpa masalah.

### 3. Membuat Folder Output Jika Belum Ada

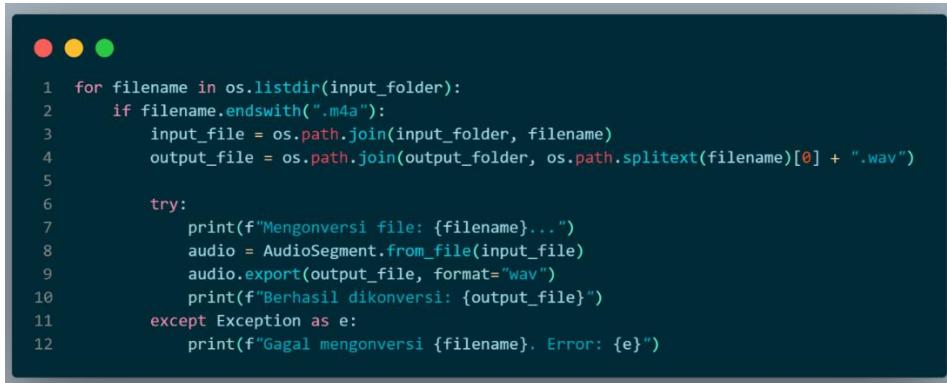


```
1 if not os.path.exists(output_folder):
2     os.makedirs(output_folder)
```

Gambar 3 Membuat Folder Output

Di sini, program memeriksa apakah folder output sudah ada atau belum. Fungsi `os.path.exists()` digunakan untuk mengecek eksistensi folder yang dimaksud. Jika folder output belum ada, maka `os.makedirs(output_folder)` akan membuatnya secara otomatis. Ini memastikan bahwa folder tujuan untuk menyimpan file hasil konversi tersedia, tanpa perlu dibuat secara manual oleh pengguna.

### 4. Konversi dan Penyimpanan File



```
1 for filename in os.listdir(input_folder):
2     if filename.endswith(".m4a"):
3         input_file = os.path.join(input_folder, filename)
4         output_file = os.path.join(output_folder, os.path.splitext(filename)[0] + ".wav")
5
6         try:
7             print(f"Mengonversi file: {filename}...")
8             audio = AudioSegment.from_file(input_file)
9             audio.export(output_file, format="wav")
10            print(f"Berhasil dikonversi: {output_file}")
11        except Exception as e:
12            print(f"Gagal mengonversi {filename}. Error: {e}")
```

Gambar 4 Konversi dan Penyimpanan File

Program ini berfungsi untuk mengonversi file audio M4A menjadi WAV secara otomatis. Program memeriksa setiap file di folder input, dan jika file tersebut berformat M4A, program akan membuat path untuk file input dan output. Menggunakan Pydub, file M4A dimuat dan dikonversi ke format WAV, kemudian disimpan di folder output. Jika terjadi error selama proses konversi, program akan mencetak pesan kesalahan dan melanjutkan ke file berikutnya. Program ini memungkinkan konversi banyak file secara otomatis dengan meminimalisir intervensi manual.

## 2. MFCC

Program ini bertujuan untuk **mengekstrak fitur MFCC (Mel-Frequency Cepstral Coefficients)** dari file audio WAV. MFCC adalah representasi dari suara yang umum digunakan dalam pemrosesan suara, seperti pengenalan suara dan analisis audio. Program ini akan membaca setiap file WAV yang ada di folder input, mengubahnya menjadi fitur MFCC menggunakan pustaka librosa, dan kemudian menyimpannya dalam format **pickle**. Pickle adalah format serialisasi data Python yang memungkinkan data disimpan dalam bentuk biner untuk digunakan kembali di masa depan. Program ini juga memastikan bahwa folder output tersedia dan menangani kemungkinan kesalahan selama proses ekstraksi dengan mencetak pesan yang sesuai.

### 1. Library

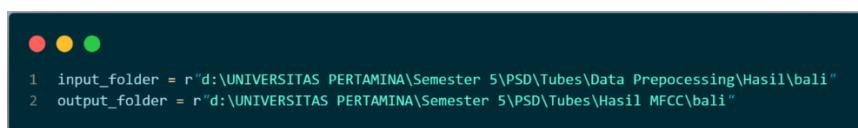


```
● ● ●
1 import os
2 import librosa
3 import numpy as np
4 import pickle
5
```

Gambar 5 Library Librosa

Program ini menggunakan empat library utama: os untuk mengelola file dan folder, seperti memeriksa dan membuat folder; librosa untuk memproses file audio WAV dan mengekstrak fitur MFCC (Mel-Frequency Cepstral Coefficients) dari audio; numpy yang digunakan oleh librosa untuk operasi numerik pada data audio; dan pickle untuk menyimpan hasil ekstraksi MFCC dalam format file biner yang dapat digunakan kembali. Keempat library ini bekerja bersama untuk mempermudah pemrosesan dan penyimpanan data audio.

### 2. Path Folder input WAV dan folder output MFCC

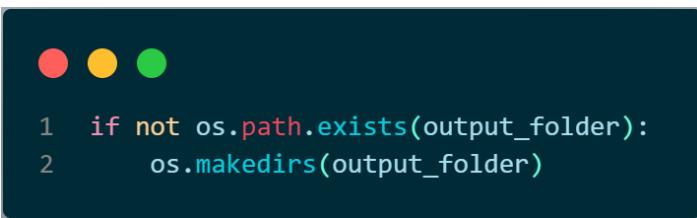


```
● ● ●
1 input_folder = r"d:\UNIVERSITAS PERTAMINA\Semester 5\PSD\Tubes\Data Preprocessing\Hasil\bali"
2 output_folder = r"d:\UNIVERSITAS PERTAMINA\Semester 5\PSD\Tubes\Hasil MFCC\bali"
```

**Gambar 6** Path Folder input WAV dan folder output MFCC

Pada bagian ini, program mendefinisikan path untuk folder input dan output. `input_folder` adalah folder yang berisi file audio WAV yang akan diproses, sedangkan `output_folder` adalah folder tujuan untuk menyimpan hasil ekstraksi MFCC dalam format file pickle. Kedua path ini ditentukan dengan menggunakan string raw (dengan `r` di depan) untuk menghindari masalah karakter khusus dalam path seperti tanda backslash (`\`) pada Windows.

### 3. Membuat folder output jika belum ada

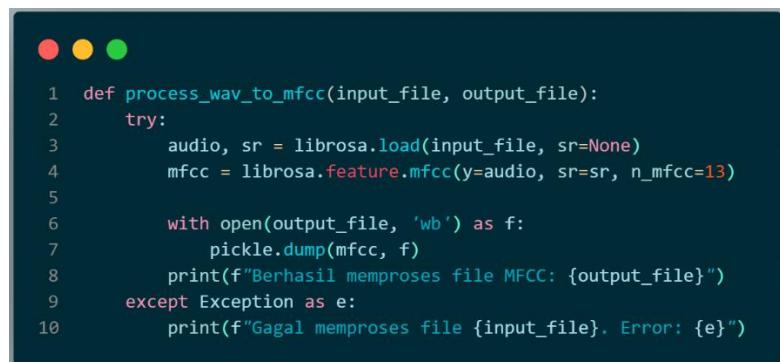


```
● ● ●
1 if not os.path.exists(output_folder):
2     os.makedirs(output_folder)
```

**Gambar 7** Folder Output

Bagian kode ini memeriksa apakah folder tujuan output sudah ada atau belum. Fungsi `os.path.exists(output_folder)` digunakan untuk mengecek keberadaan folder output. Jika folder tersebut belum ada, maka perintah `os.makedirs(output_folder)` akan membuat folder baru dengan path yang telah ditentukan. Hal ini memastikan bahwa folder output tersedia sebelum proses menyimpan hasil ekstraksi MFCC dimulai.

### 4. Fungsi untuk mengubah WAV menjadi MFCC



```
● ● ●
1 def process_wav_to_mfcc(input_file, output_file):
2     try:
3         audio, sr = librosa.load(input_file, sr=None)
4         mfcc = librosa.feature.mfcc(y=audio, sr=sr, n_mfcc=13)
5
6         with open(output_file, 'wb') as f:
7             pickle.dump(mfcc, f)
8             print(f"Berhasil memproses file MFCC: {output_file}")
9     except Exception as e:
10         print(f"Gagal memproses file {input_file}. Error: {e}")
```

**Gambar 8** Fungsi untuk mengubah WAV menjadi MFCC

Fungsi `process_wav_to_mfcc` bertanggung jawab untuk memproses file audio WAV menjadi fitur MFCC dan menyimpannya dalam file pickle. Pertama, file audio dibaca menggunakan `librosa.load()` untuk memuat audio dan memperoleh sample rate (sr). Kemudian, `librosa.feature.mfcc()` digunakan untuk mengekstrak 13 koefisien MFCC dari audio yang dimuat. Setelah itu, hasil ekstraksi MFCC disimpan dalam file pickle menggunakan `pickle.dump()`, dengan mode penulisan biner ('wb'). Jika proses berjalan lancar, fungsi ini mencetak pesan bahwa file berhasil diproses. Jika terjadi kesalahan, misalnya file tidak dapat diproses, maka pesan error akan dicetak menggunakan `except` dan menampilkan deskripsi kesalahan yang terjadi.

## 5. Iterasi untuk semua file WAV dalam folder input

```
● ● ●
1 for filename in os.listdir(input_folder):
2     if filename.endswith(".wav"):
3         input_file = os.path.join(input_folder, filename)
4         output_file = os.path.join(output_folder, os.path.splitext(filename)[0] + ".pkl") # Simpan sebagai file pickle
5         process_wav_to_mfcc(input_file, output_file)
6
7 print("Proses ekstraksi MFCC selesai!")
```

Gambar 9 Iterasi untuk semua file WAV dalam folder input

Bagian kode ini melakukan iterasi untuk memproses semua file WAV dalam folder input. Program akan memeriksa setiap file di dalam folder menggunakan `os.listdir(input_folder)` dan memastikan hanya file dengan ekstensi .wav yang diproses menggunakan `filename.endswith(".wav")`. Untuk setiap file WAV, path file input dan output ditentukan menggunakan `os.path.join()`. Nama file output diubah dengan mengganti ekstensi menjadi .pkl agar hasil ekstraksi MFCC disimpan dalam format pickle. Fungsi `process_wav_to_mfcc(input_file, output_file)` kemudian dipanggil untuk memproses file audio dan menyimpan hasilnya. Setelah semua file diproses, program mencetak pesan "Proses ekstraksi MFCC selesai!" untuk menandakan bahwa seluruh proses telah selesai.

## BAB 4 HASIL DAN PEMBAHASAN

### A. Hasil

#### 1. Deskripsi Dataset

Penelitian ini menggunakan data suara yang diambil dari lima daerah berbeda, yaitu Timur, Jawa, Medan, Sunda, dan Minang. Setiap daerah diwakili oleh dua partisipan, masing-masing menyumbangkan 10 sampel suara, sehingga total dataset terdiri dari 100 sampel suara. Dataset ini telah diekstraksi menggunakan metode Mel Frequency Cepstral Coefficients (MFCC), menghasilkan representasi numerik dari pola suara dengan dimensi yang bervariasi, tergantung pada durasi masing-masing rekaman. Dimensi matriks MFCC untuk setiap file suara tidak seragam, melainkan bergantung pada panjang rekaman. Hal ini dikarenakan proses ekstraksi MFCC menghasilkan jumlah frame yang proporsional terhadap durasi audio, dengan jumlah koefisien tetap per frame.

#### 2. Hasil Ekstraksi Fitur

Hasil ekstraksi menggunakan MFCC menghasilkan pola spektral yang unik untuk setiap logat. Tabel berikut menunjukkan dimensi rata-rata matriks MFCC untuk setiap logat per daerahnya:

Logat	Partisipan	Jumlah Sampel	Avg Dimensi MFCC
Timur	2	20	(13 x 1171.2)
Jawa	2	20	(13 x 1200)
Medan	1	10	(13 x 1227.5)
Sunda	2	20	(13 x 865.3)
Minang	2	20	(13 x 1005.6)
Bali	1	10	(13 x 876.1)

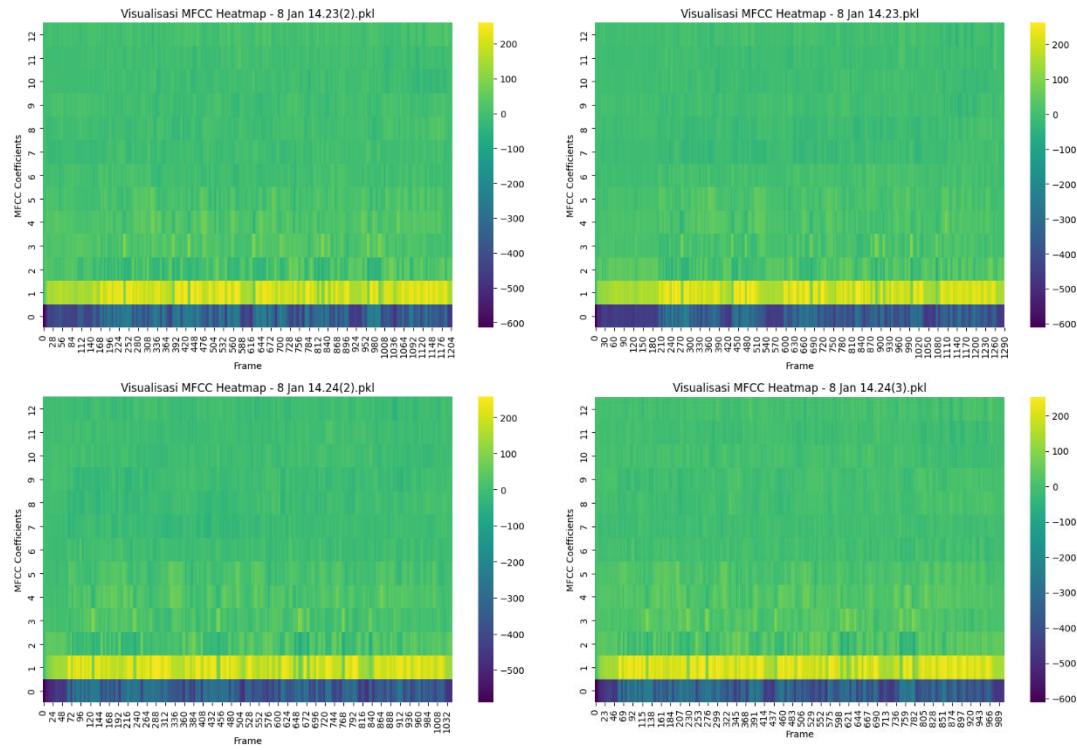
Hasil Matriks MFCC (Mel-Frequency Cepstral Coefficients) ini disimpan dalam format pickle (pkl) dengan tujuan untuk memudahkan penyimpanan dan pengambilan data pada sesi yang berbeda. Tujuan utamanya adalah untuk mempermudah proses training data pada machine learning. Format pickle memungkinkan untuk menyimpan data yang telah diproses, seperti fitur MFCC dari file audio, sehingga tidak perlu mengulang proses ekstraksi fitur setiap kali ingin melatih model. Ini menghemat waktu yang berharga, terutama ketika bekerja dengan dataset besar. Selain itu, dengan menggunakan pickle, data

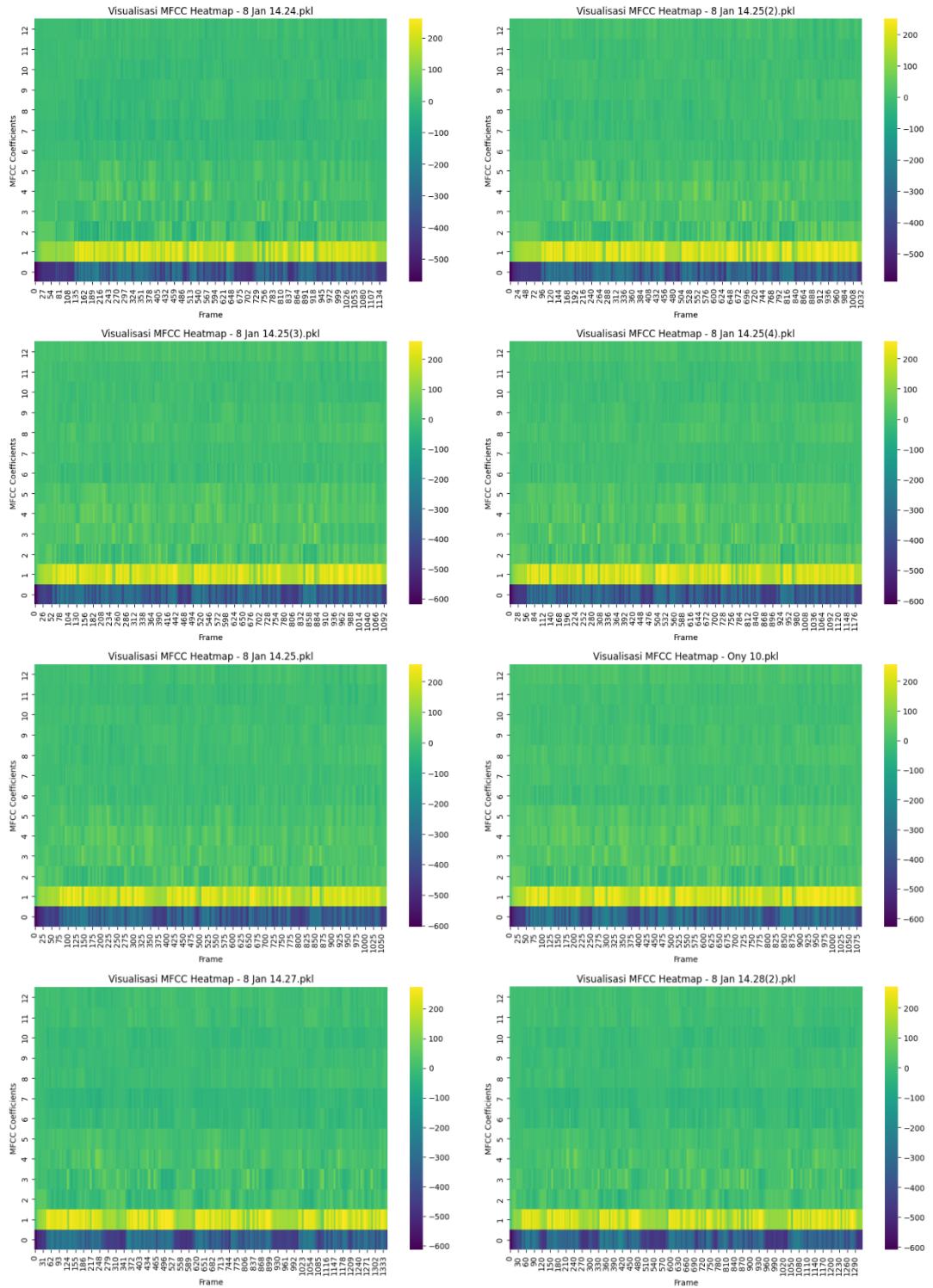
yang digunakan untuk pelatihan model tetap konsisten dan tidak mengalami perubahan yang tidak diinginkan. Hal ini juga membantu dalam mengelola dataset besar dengan cara yang lebih efisien, karena dapat menyimpan data dalam file terpisah untuk setiap kategori atau label, dan memuatnya sesuai kebutuhan saat melatih model. Dengan demikian, dapat langsung memuat data MFCC yang telah disimpan dan menggunakannya untuk melatih model machine learning tanpa perlu pengolahan lebih lanjut, meningkatkan efisiensi dan mempercepat proses eksperimen model.

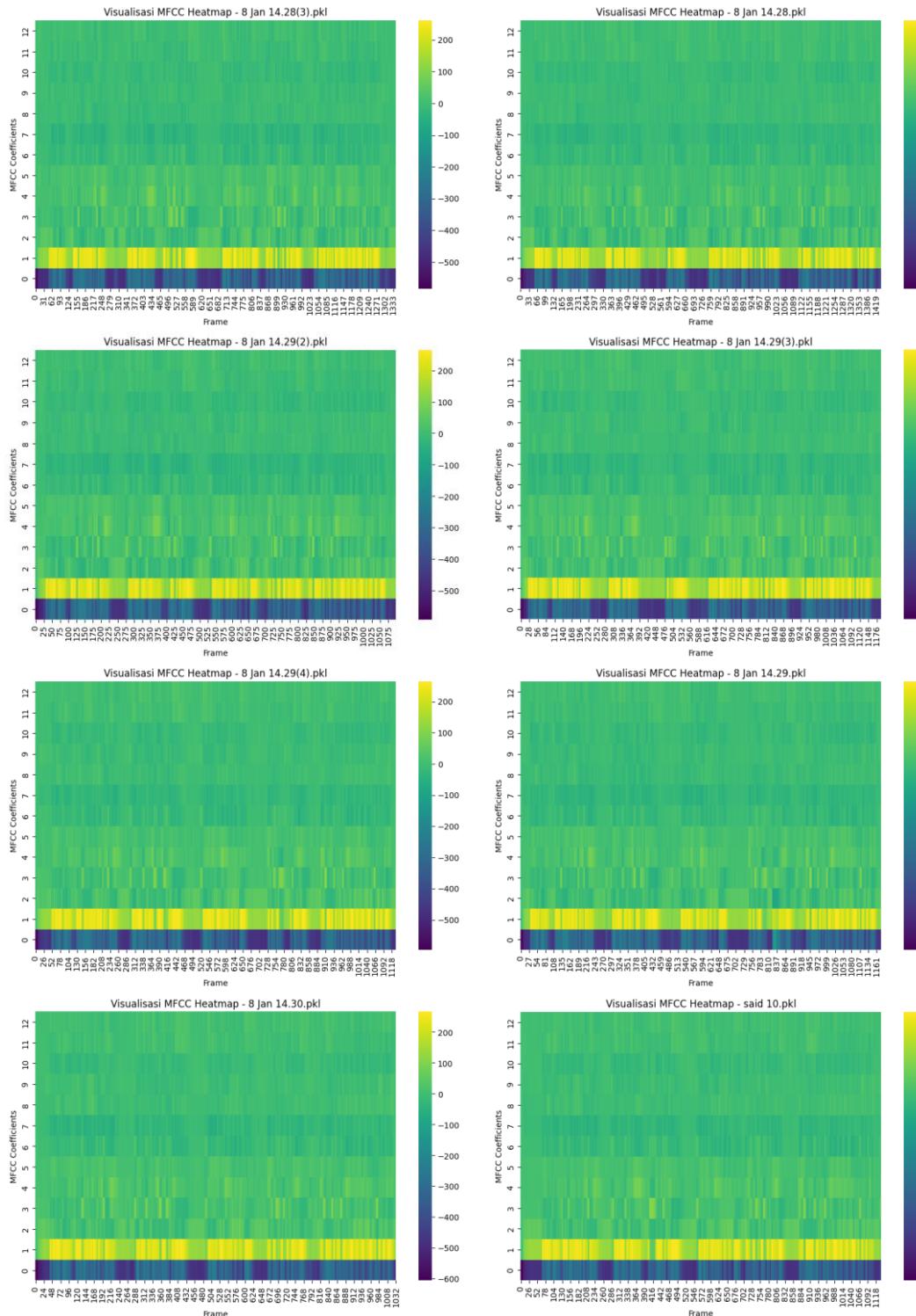
### 3. Visualisasi MFCC

Visualisasi MFCC dilakukan untuk menunjukkan pola spektral khas dari setiap logat. Meskipun dimensi MFCC berbeda pada setiap file suara, pola yang unik untuk setiap logat dapat diamati pada visualisasi tersebut.

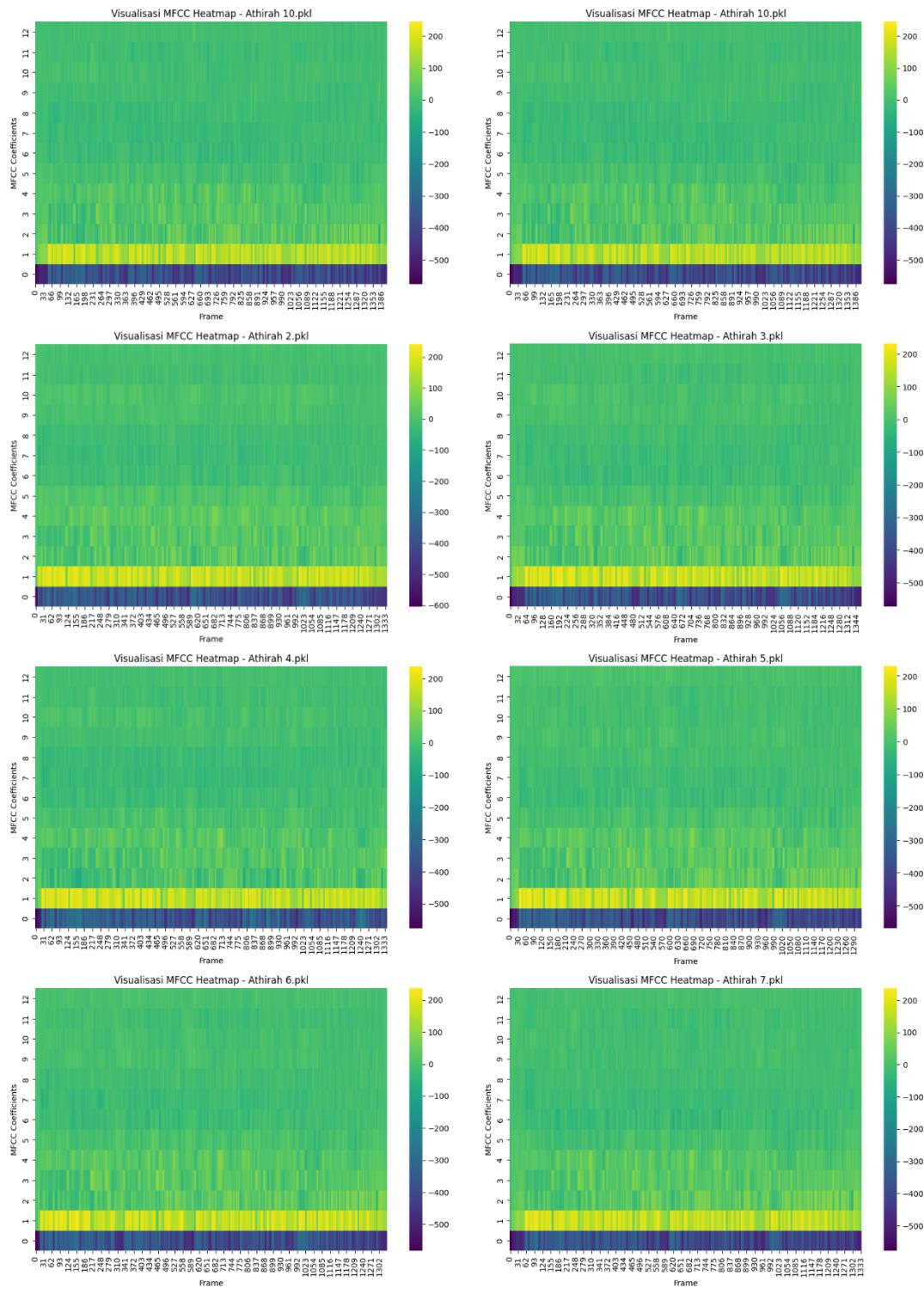
- Visualisasi MFCC logat Timur

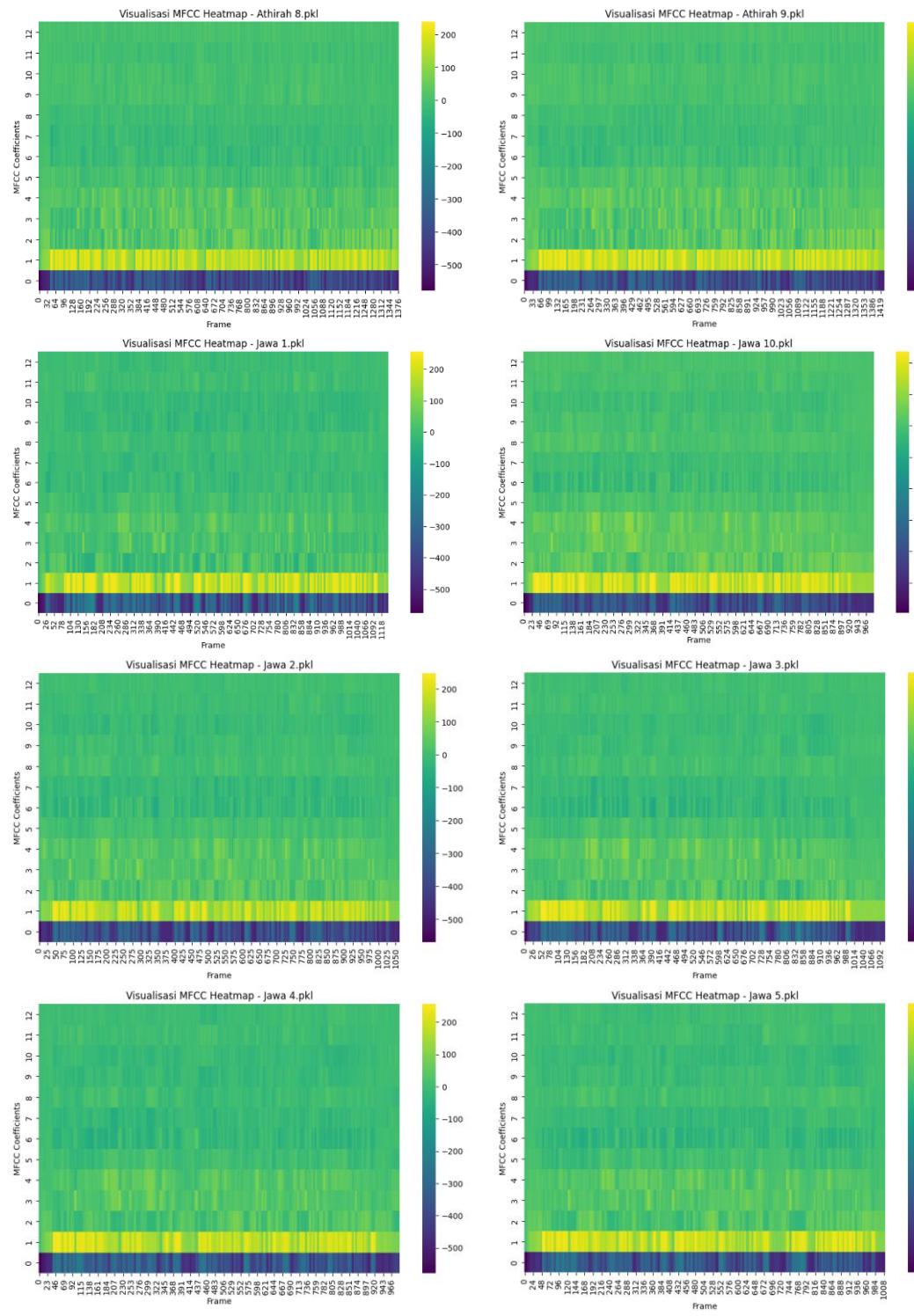


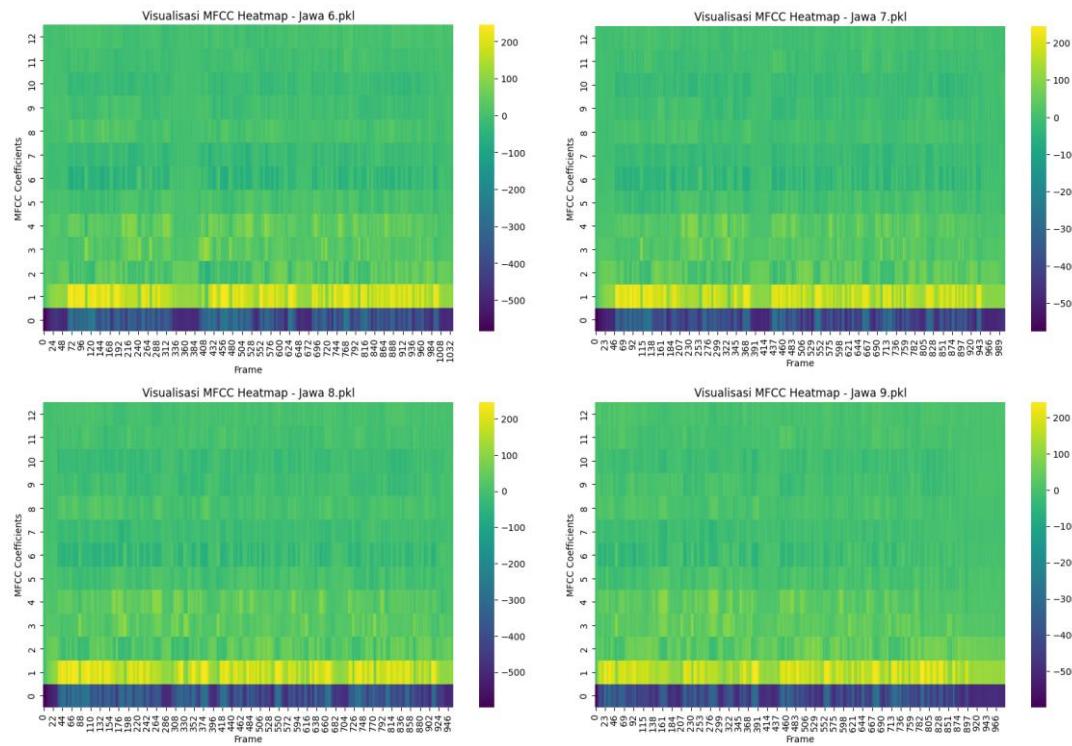




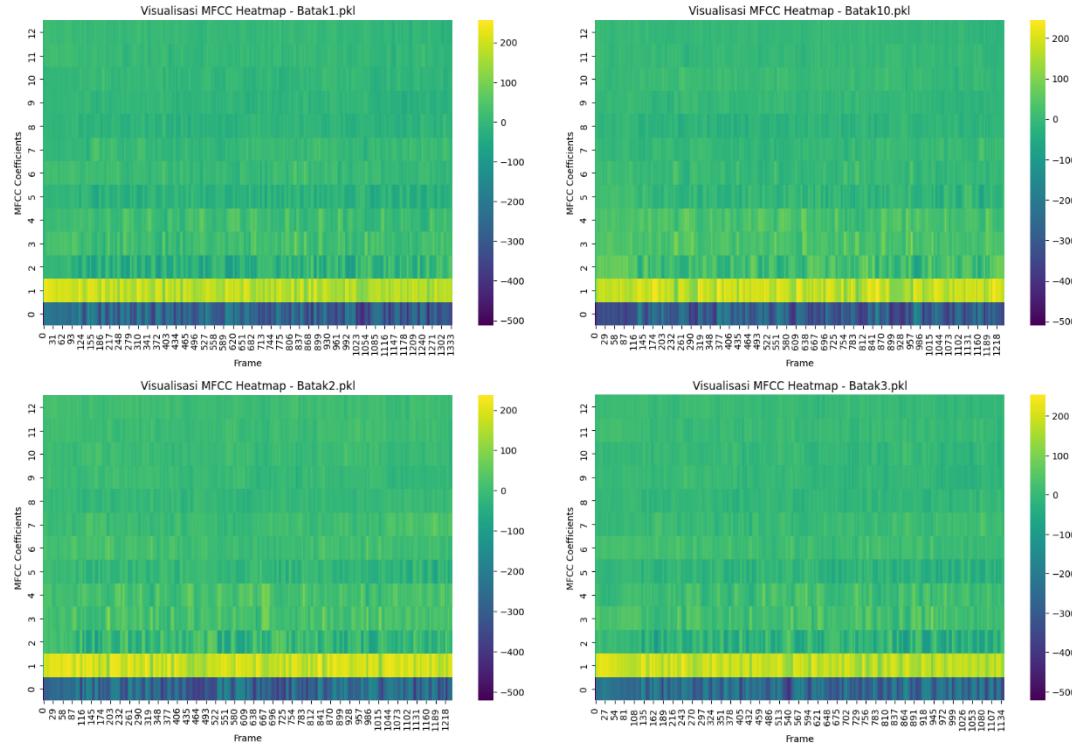
- Visualisasi MFCC logat Jawa

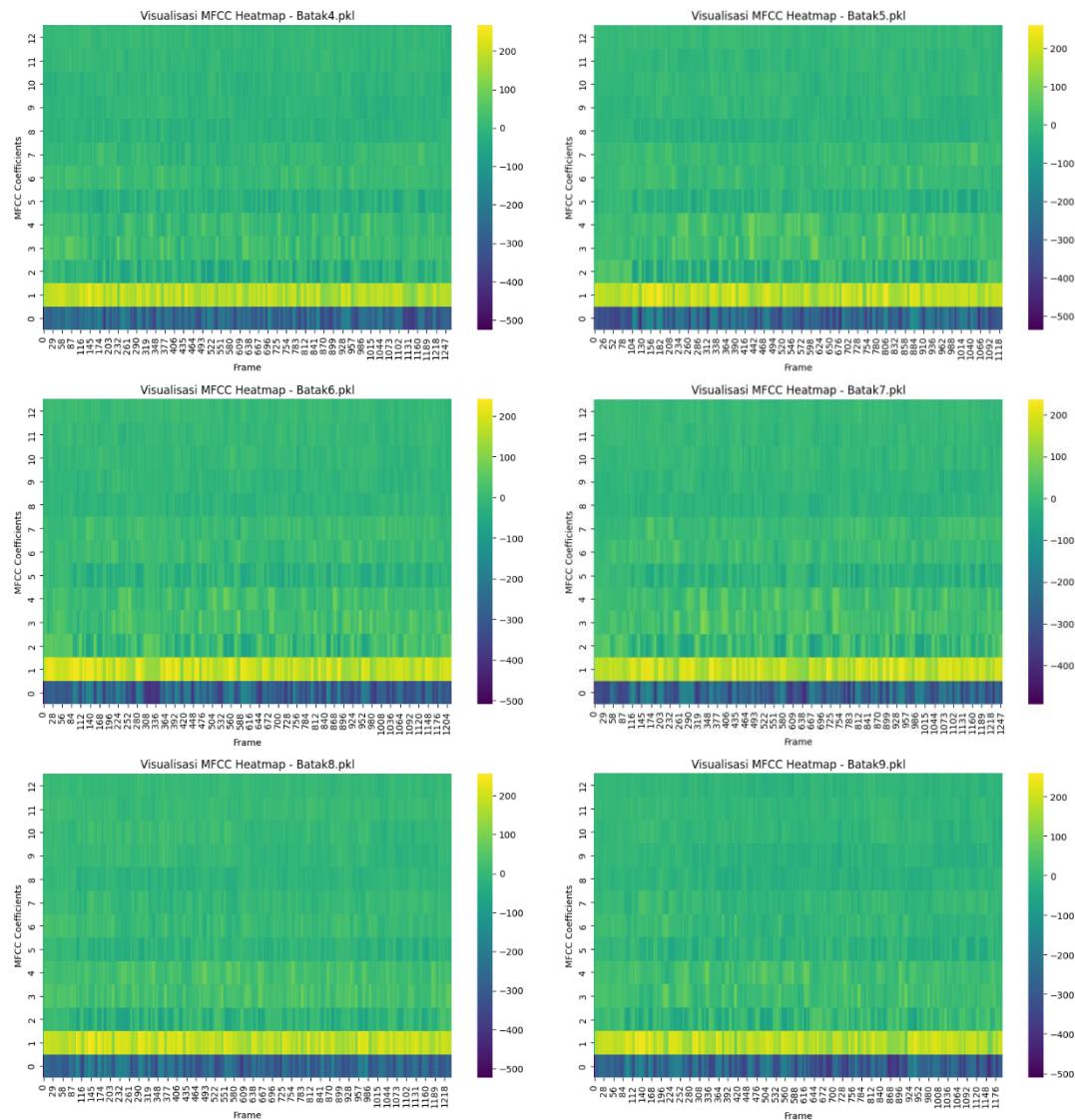




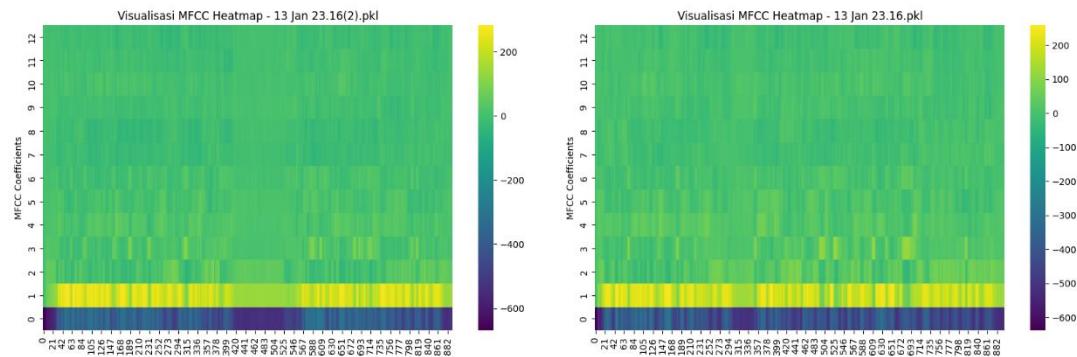


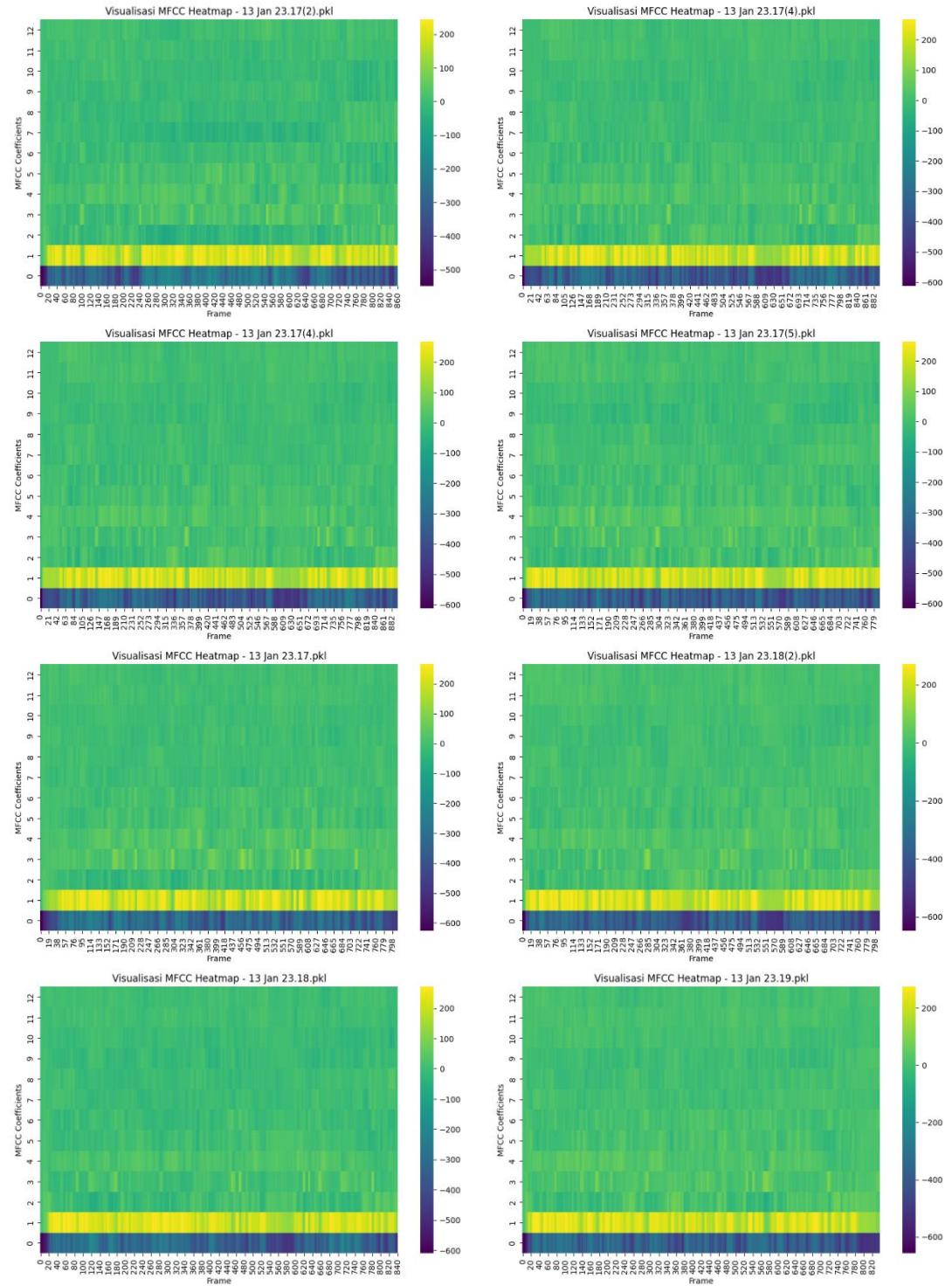
- Visualisasi MFCC logat Medan

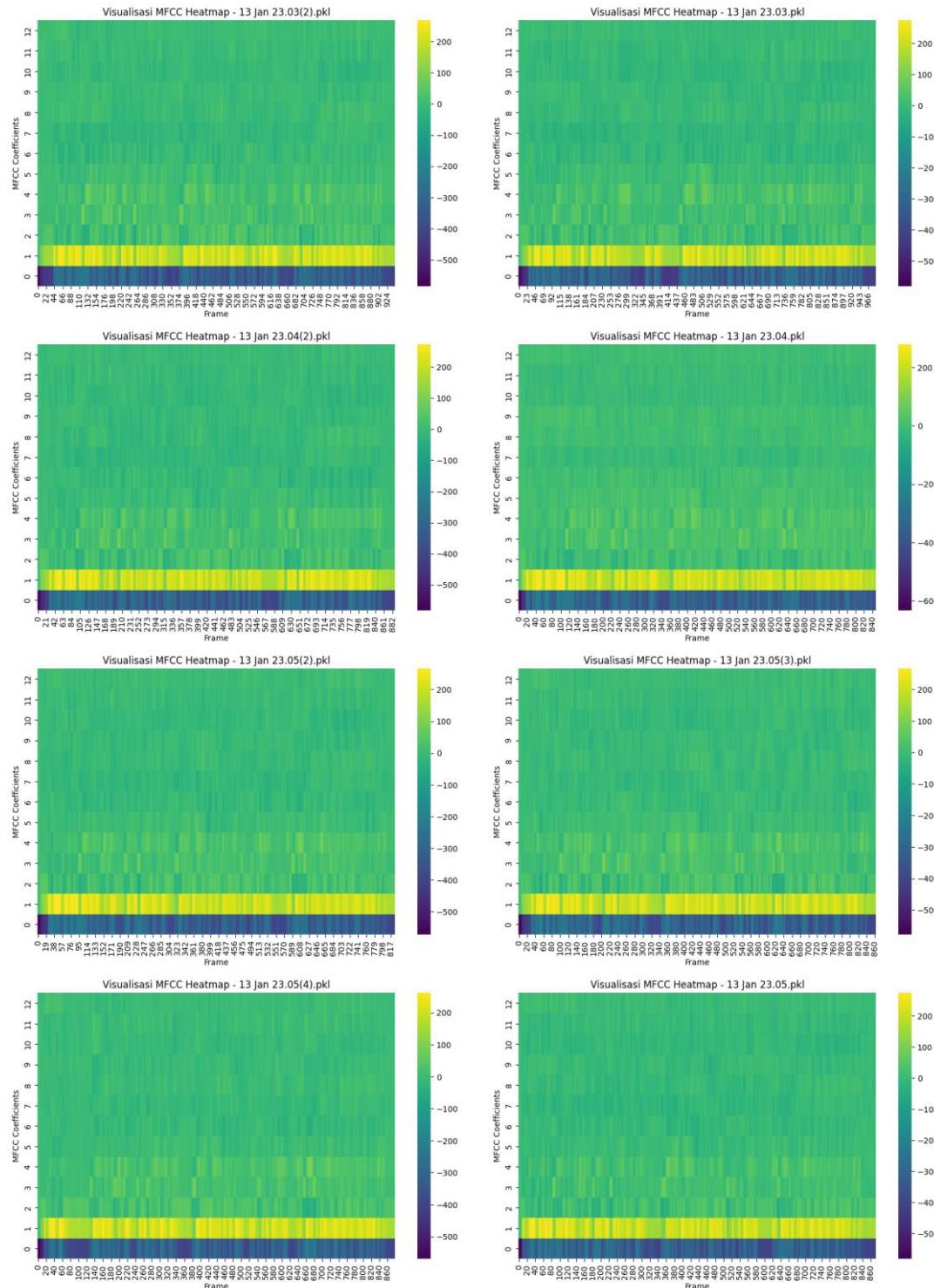


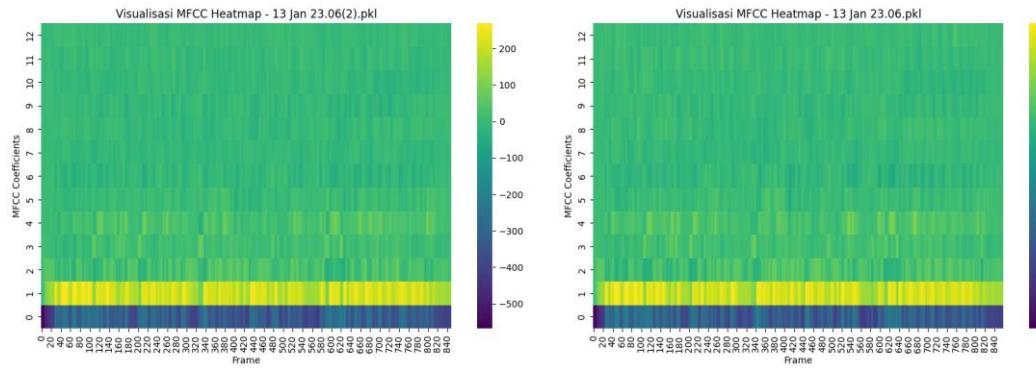


- Visualisasi MFCC logat Sunda

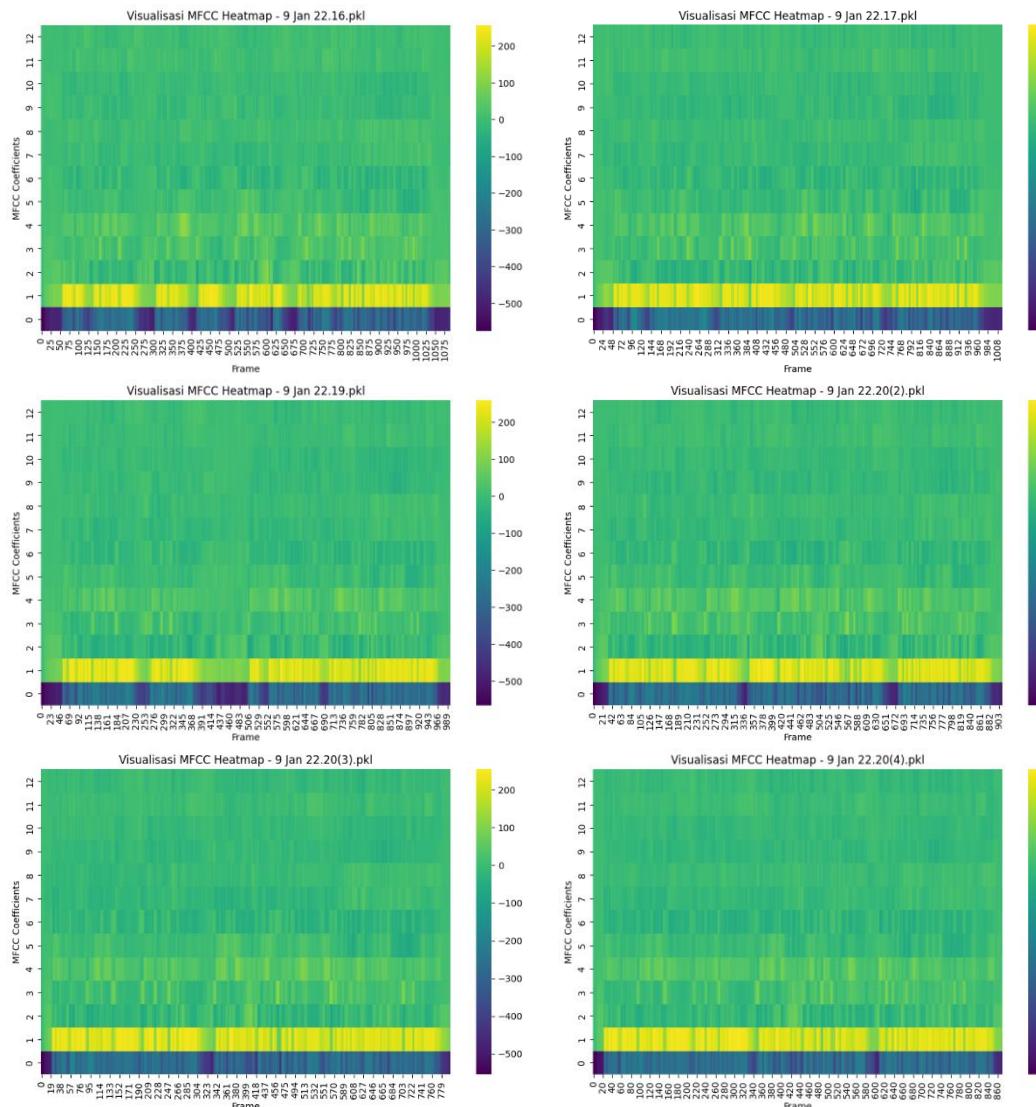


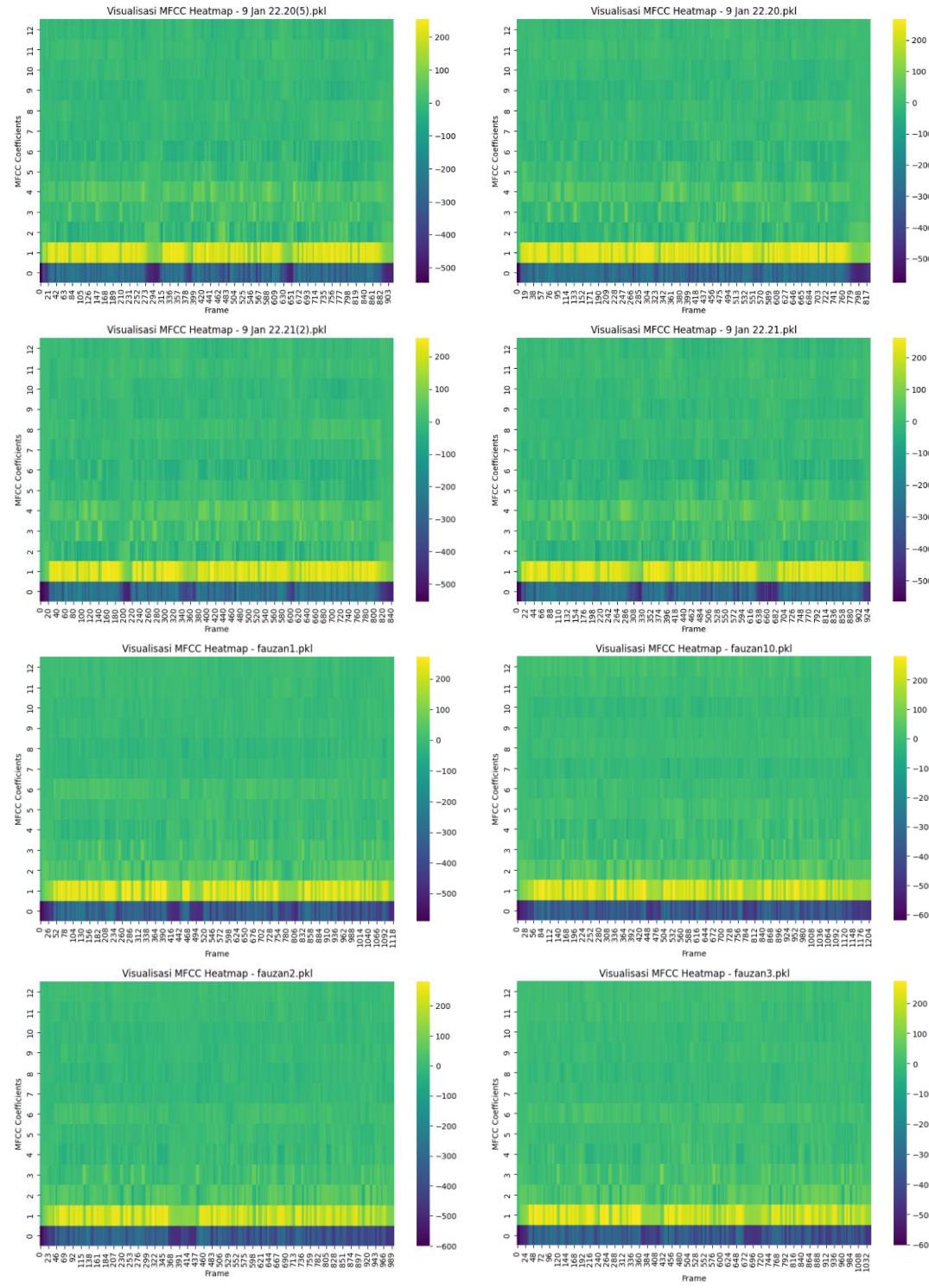


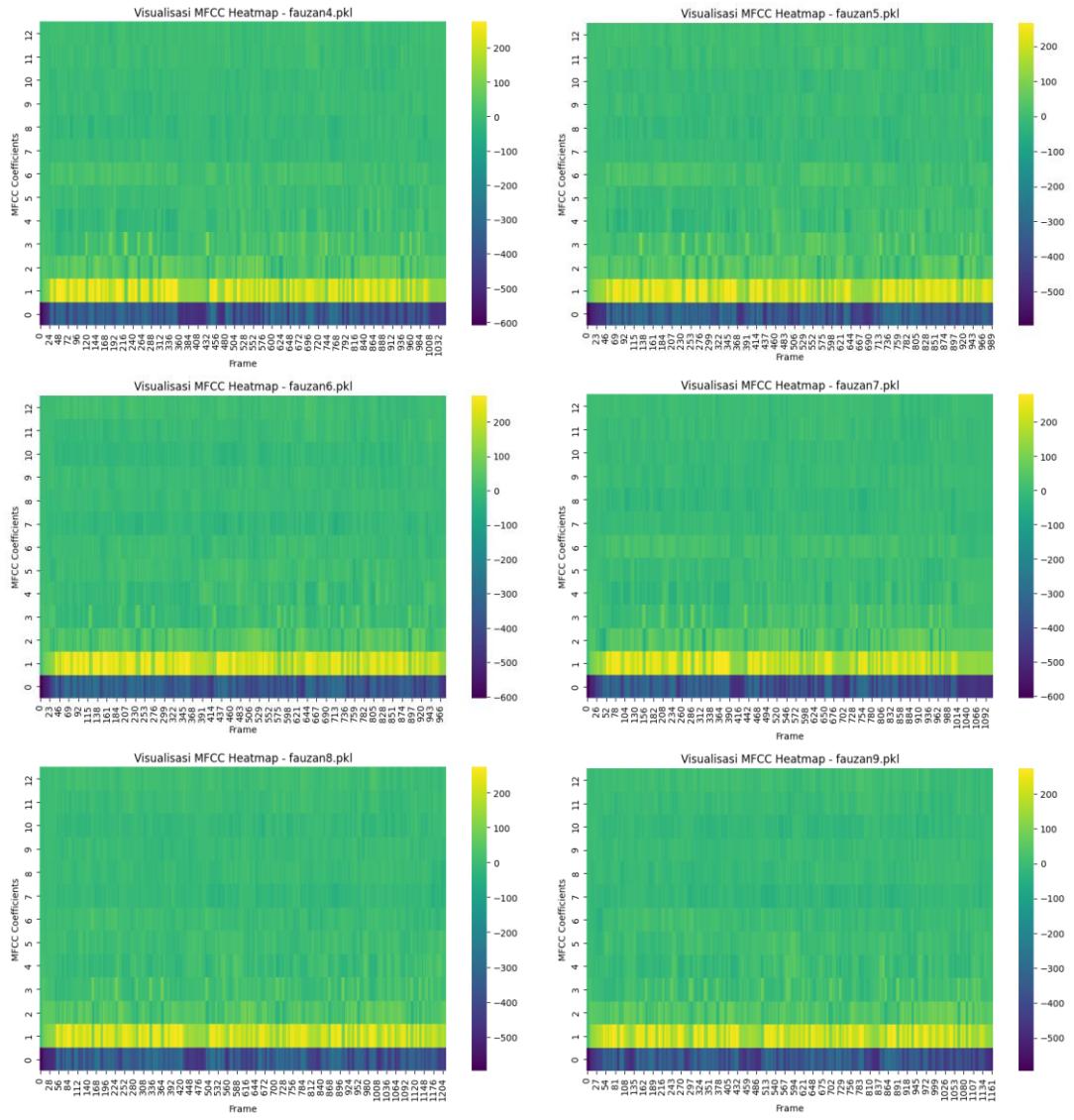




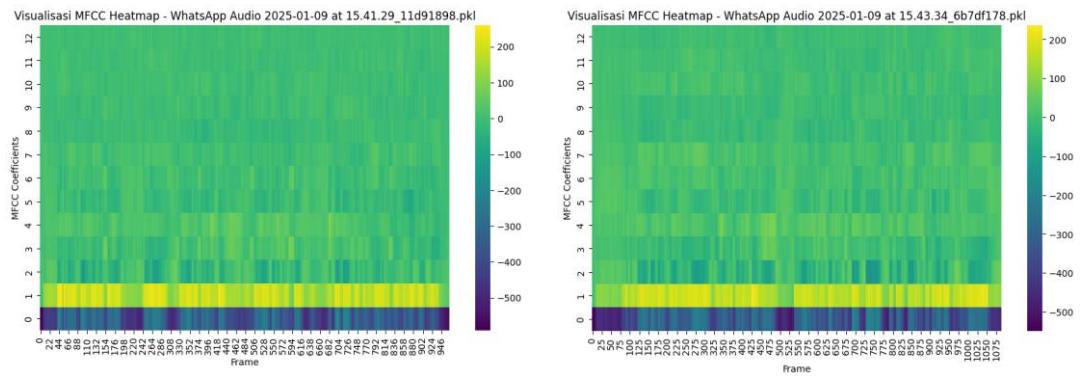
- Visualisasi MFCC logat Minang

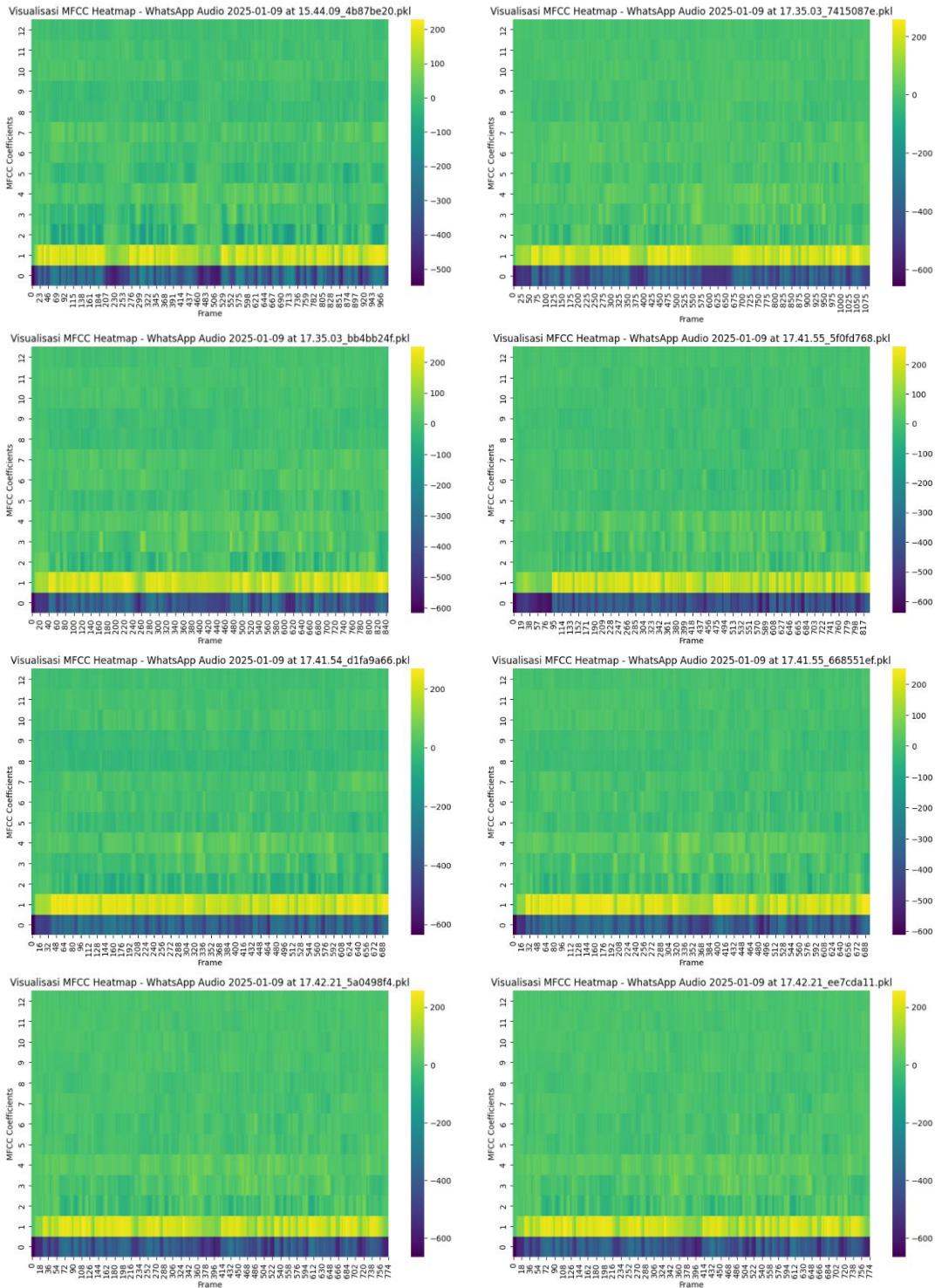






- Visualisasi MFCC logat Bali





## B. Pembahasan

### 1. Variasi Dimesi Matriks MFCC

Dimensi rata-rata matriks MFCC memberikan gambaran kuantitatif mengenai karakteristik durasi dan pola bicara dari setiap logat. Setiap logat memiliki dimensi yang berbeda, yang mencerminkan keunikan dalam pengucapan, tempo, dan intonasi berbicara. Berikut penjelasan lebih rinci:

a. Logat Timur (Dimesi Rata-rata  $13 \times 1171.2$ )

Logat Timur menunjukkan dimensi MFCC yang tinggi. Hal ini dapat disebabkan oleh:

- **Durasi Ucapan Lebih Panjang:** Gaya bicara logat Timur cenderung lebih lambat dan jelas, menghasilkan lebih banyak frame dalam proses ekstraksi MFCC.
- **Dominasi Vokal dan Intonasi:** Pengucapan vokal yang panjang dengan jeda antar kata yang lebih lebar menghasilkan pola suara yang lebih kompleks.
- **Konteks Budaya:** Gaya berbicara yang lebih santai dan formal, seperti pada percakapan adat atau budaya lokal, dapat berkontribusi pada dimensi MFCC yang lebih besar.

b. Logat Jawa (Dimesi Rata-rata  $13 \times 1200$ )

Dimensi rata-rata logat Jawa merupakan salah satu yang terbesar. Ini menunjukkan:

- **Pengucapan Formal dan Terstruktur:** Logat Jawa sering diasosiasikan dengan nilai kehalusan dan kehati-hatian dalam berbicara. Pola ini mencerminkan durasi ucapan yang konsisten dan panjang.
- **Tempo Sedang:** Tidak secepat logat Sunda, tetapi tetap menjaga irama bicara yang berkesinambungan.
- **Penggunaan Intonasi yang Halus:** Intonasi yang mendalam tetapi terkontrol membuat logat Jawa menghasilkan pola spektral yang lebih kaya dan kompleks.

c. Logat Medan (Dimesi Rata-rata  $13 \times 1227.5$ )

Dimensi logat Medan adalah yang terbesar di antara semua logat. Hal ini mencerminkan:

- **Energi Bicara yang Tinggi:** Gaya berbicara yang lugas, tegas, dan penuh energi menghasilkan frame suara yang lebih banyak.
- **Durasi dan Pola Intonasi yang Kaya:** Pengucapan cenderung lebih panjang dan kuat, sehingga menghasilkan representasi spektral yang lebih besar.
- **Karakteristik Regional:** Keunikan logat Medan dengan ritme bicara yang penuh tekanan turut memengaruhi hasil dimensi MFCC yang tinggi.

d. Logat Sunda (Dimesi Rata-rata  $13 \times 865.3$ )

Logat Sunda memiliki dimensi terkecil, yang mengindikasikan:

- **Gaya Bicara Cepat dan Ringan:** Partisipan berbicara dengan tempo yang cepat, menghasilkan durasi rekaman yang lebih pendek.
- **Pengucapan Singkat:** Gaya berbicara yang cenderung efisien dengan sedikit jeda menghasilkan pola suara yang sederhana.
- **Irama Bicara yang Khas:** Intonasi ringan khas Sunda turut berkontribusi pada hasil MFCC dengan dimensi yang lebih kecil.

e. Logat Minang (Dimesi Rata-rata  $13 \times 1005.6$ )

Dimensi logat Minang berada di antara logat Jawa dan Sunda, menunjukkan:

- **Kecepatan Bicara yang Moderat:** Logat ini tidak terlalu cepat seperti Sunda, tetapi juga tidak selama logat Timur atau Medan.
- **Penggunaan Ritme yang Berirama:** Logat Minang memiliki ritme yang khas, sering digunakan dalam pantun atau pidato adat, yang memberikan pola suara yang konsisten.
- **Intonasi Khas:** Pola suara cenderung dinamis dengan tekanan tertentu yang membuat dimensi MFCC cukup signifikan.

f. Logat Bali (Dimesi Rata-rata  $13 \times 876.1$ )

Dimensi logat Bali sedikit lebih besar dari logat Sunda, tetapi tetap kecil dibandingkan logat lain. Hal ini menunjukkan:

- **Gaya Bicara yang Tenang:** Walaupun tenang, tempo bicara tetap cenderung cepat, menghasilkan durasi rekaman yang relatif pendek.
- **Penggunaan Vokal yang Stabil:** Logat Bali memiliki intonasi yang lebih datar dibandingkan logat lainnya, menghasilkan dimensi yang lebih kecil tetapi tetap konsisten.
- **Budaya Lokal:** Keterpengaruhannya budaya Bali yang kaya dengan nyanyian dan intonasi khas mungkin memengaruhi pola bicara dalam sampel.

## 2. Visualisasi MFCC sebagai Representasi Pola Logat

Visualisasi MFCC memberikan wawasan penting tentang bagaimana pola suara dari berbagai logat dapat dianalisis secara visual. Setiap logat memiliki pola spektral yang berbeda, seperti distribusi energi pada frekuensi tertentu, yang direpresentasikan oleh warna dalam spektrogram MFCC. Dengan menganalisis spektrogram ini, pola unik dari setiap logat dapat diidentifikasi secara lebih intuitif.

Beberapa poin penting terkait visualisasi MFCC dalam penelitian ini adalah sebagai berikut:

- **Perbedaan Pola Visual Antarlogat:** Pada logat Timur, visualisasi MFCC menunjukkan pola dengan intensitas tinggi pada rentang frekuensi tertentu, yang mencerminkan pengucapan vokal yang panjang. Sebaliknya, logat Sunda menunjukkan pola yang lebih rapat, yang menggambarkan kecepatan berbicara yang lebih cepat.
- **Distribusi Frekuensi yang Unik:** Setiap logat memiliki distribusi energi pada frekuensi yang berbeda. Sebagai contoh, logat Minang memiliki energi dominan pada frekuensi rendah hingga menengah, sedangkan logat Jawa memiliki distribusi energi yang lebih merata di seluruh rentang frekuensi.
- **Konsistensi Pola Antarpartisipan:** Visualisasi MFCC juga menunjukkan bahwa pola spektral cenderung konsisten antarpartisipan dari daerah yang sama, meskipun terdapat sedikit variasi yang disebabkan oleh karakteristik individu, seperti intonasi atau nada bicara.

## 3. Hubungan Dimensi MFCC dan Keunikan Logat

MFCC adalah representasi numerik dari pola frekuensi suara, sehingga dapat menangkap ciri khas dari masing-masing logat. Variasi dimensi matriks MFCC tidak hanya dipengaruhi oleh logat, tetapi juga oleh kecepatan berbicara dan jumlah jeda yang terjadi selama perekaman. Hal ini menunjukkan bahwa logat-logat tertentu memiliki pola spektral yang konsisten dan dapat dijadikan indikator untuk membedakan logat.

## KESIMPULAN

Analisis matriks MFCC dari berbagai logat menunjukkan bahwa setiap logat memiliki karakteristik unik yang tercermin dari dimensi rata-rata dan pola spektral yang dihasilkan. Logat Medan memiliki dimensi MFCC terbesar, mencerminkan gaya bicara yang lugas dan penuh energi, sedangkan logat Sunda memiliki dimensi terkecil, mengindikasikan tempo bicara yang cepat dan ringan. Logat lain, seperti Jawa, Timur, Minang, dan Bali, menampilkan variasi dimensi MFCC yang dipengaruhi oleh keunikan pengucapan, intonasi, dan budaya lokal masing-masing. Hal ini menunjukkan bahwa MFCC mampu menangkap perbedaan subtil dalam pola bicara setiap logat, menjadikannya fitur yang relevan untuk pengklasifikasian logat.

Hasil analisis ini juga memberikan dasar yang kuat untuk persiapan data dalam pelatihan model machine learning. Matriks MFCC yang telah dihasilkan dan disimpan dalam format terstruktur dapat digunakan sebagai *input feature* untuk algoritma pembelajaran seperti classification atau clustering. Dengan pola spektral yang konsisten antarpartisipan dari logat yang sama, MFCC dapat membantu meningkatkan akurasi model dengan menyediakan representasi data yang kaya akan informasi akustik. Ini menjadikan MFCC sebagai langkah awal yang krusial dalam membangun sistem pengenalan logat berbasis machine learning.

## DAFTAR PUSTAKA

- [1] Ajinurseto, G., & Islamuddin, N. "Penerapan Metode Mel Frequency Cepstral Coefficients pada Sistem Pengenalan Suara Berbasis Desktop. INFOMATEK": *Jurnal Informatika, Manajemen dan Teknologi*, 25(1), 11-20, 2023
- [2] Alfian, A. "Analisis Audio Menggunakan Speech Recognition untuk Identifikasi dan Deteksi Suara dalam Bahasa Indonesia". Skripsi, Fakultas Sains dan Teknologi, Universitas Islam Negeri Sunan Kalijaga, 2023
- [3] Andriana, A. "SISTEM SPEAKER RECOGNITION (PENGENAL PENGUCAP) UNTUK MENCARI KARAKTERISTIK UCAPAN SESEORANG DENGAN METODE MEL FREQUENCY CEPTRUM COEFFICIENT (MFCC) MENGGUNAKAN SOFTWARE MATLAB". Jurnal: Industri Elektro dan Penerbangan, 3(1), 2013
- [4] Bilaldi, R. "Melestarikan Bahasa Daerah dengan Mengenal 6 Variasi Dialek Bahasa Indonesia", kompasiana, 2020
- [5] Chou W., Juang B., "Pattern Recognition in Speech and Language Processing", CRC Press LLC, New Jersey, 2003
- [6] Habib M. (2024, July 24). Apa itu FFmpeg? Fungsi dan Penggunaannya
- [7] Islamiyah, D., Herwanto, H., Anggrayni, O. D., & Choiri, M. "Pemahaman Dialek dan Variasi Bahasa di Indonesia melalui Kajian Literatur". Journal of Linguistics and Social Studies, 1(1), 33-43, 2024
- [8] Krisnawati, E. "Ragam Bahasa Indonesia: Macam, Ciri-Ciri, dan Contohnya", tиро. 2024
- [9] Ltd, A. P. (2025, January 8). *Open Source Python Library untuk Membuat, Menyimpan, Memuat & Mengekstrak File Audio*. Aspose Pty. Ltd.
- [10] McFee, B., Raffel, C., Liang, D., Ellis, D. P. W., McVicar, M., Battenberg, E., & Nieto, O. (2015). librosa: Audio and music analysis in Python. In Proceedings of the 14th Python in Science Conference (pp. 18-25). Retrieved from <https://librosa.org/doc/latest/index.html>
- [11] Murniaseh, E. "Contoh Dialek, Pengertian, dan Penggunaannya di Masyarakat", tиро.id, 2021.
- [12] Nasution, T. "Metoda mel frequency cepstrum coefficients (mfcc) untuk mengenali ucapan pada bahasa indonesia". *Jurnal Sains dan Teknologi Informasi*, 1(1), 22-31, 2012

- [13] Novela, M., & Basaruddin, T. "Dataset Suara dan Teks Berbahasa Indonesia Pada Rekaman Podcast dan Talk show". *Jurnal Fasilkom*, 11(2), 61-66, 2021
- [14] Nur Azizah U. "6 Jenis Keragaman Budaya di Indonesia dan Faktor yang Memengaruhinya" *detikJateng*, 2024.
- [15] Prihapsari, I. "Karakteristik Bahasa Indonesia Logat Papua dan Relevansinya sebagai Materi Ajar Mata Kuliah Sosiolinguistik di Program Studi Pendidikan Bahasa Indonesia Universitas Sebelas Maret", 2018.
- [16] Samosir, C. "Definisi Logat Bahasa dan Ciri-cirinya dalam Kehidupan Sehari-hari", parapuan , 2024
- [17] Swasty, R. "Mengenal Lebih Dekat dengan Dialek dan Logat, Ciri-Ciri dan Perbedaannya", medcom, 2024
- [18] Umam "8 Macam Keragaman Budaya Indonesia & Contoh Keragaman Budaya Indonesia" Gramedia Blog, 2024.
- [19] Unmaha. (2024, August 17). Membuat Aplikasi Music Player dengan Python dan PyDub. Blog | Universitas Mahakarya Asia | UNMAHA.
- [20] Yunita Putri, A. "Sistem Identifikasi Suara Berbasis Convolutional Neural Network". Skripsi, Fakultas Teknik, Universitas Sriwijaya, 2019.