

IF2123 Aljabar Linier dan Geometri
Laporan Tugas Besar 2



Disusun oleh:

Kelompok 28 (Hatsune Mix[ue])

Rhio Bimo Prakoso S (13523123)
Rafael Marchel Darma Wijaya (13523146)
I Made Wiweka (13523160)

Dosen Pengampu:

Dr. Judhi Santoso, M.Sc.
Arrival Dwi Sentosa, S. Kom., M.T

PROGRAM STUDI TEKNIK INFORMATIKA
SEKOLAH TEKNIK ELEKTRO DAN INFORMATIKA
INSTITUT TEKNOLOGI BANDUNG
JL. GANESA 10, BANDUNG 40132
2024

Daftar Isi

Daftar Isi	2
BAB I	
DESKRIPSI MASALAH	3
BAB II	
TEORI SINGKAT	5
I. PCA	5
II. MIR	8
BAB III	
ARSITEKTUR WEBSITE DAN PROGRAM INFORMATION RETRIEVAL	11
BAB IV	
EKSPERIMEN	13
4.1. Album Picture Finder	13
4.2. Music Information Retrieval (MIR)	13
BAB V	
KESIMPULAN	21
5.1. Kesimpulan	21
5.2. Saran	21
5.3. Refleksi	22
LAMPIRAN	23
Referensi:	23
Tautan:	23

BAB I

DESKRIPSI MASALAH

Tugas ini bertujuan untuk mengembangkan sebuah website yang dapat mendeteksi dan mencari suara serta gambar yang paling cocok dalam dataset. Adapun spesifikasi dan ketentuan tugas adalah sebagai berikut:

1. Bahasa Pemrograman

Pemilihan bahasa pemrograman untuk frontend dan backend sepenuhnya bebas. Pilihlah bahasa yang sesuai dan memberikan keuntungan bagi tim.

2. Dataset

Dataset audio bebas, dapat diperoleh dari sumber seperti *Kaggle* atau lainnya.

Dataset yang akan diuji akan diumumkan kemudian.

3. Fungsi Utama Program

Program harus mampu menampilkan daftar audio yang telah diunggah dari dataset.

Program harus dapat mencari kecocokan audio berdasarkan metode *query and/or by humming*, menggunakan format file MIDI atau WAV.

Program harus menunjukkan hasil audio atau gambar dengan tingkat kemiripan antara 55% hingga 100%.

Pagination harus diterapkan untuk menghindari *infinite scrolling* saat jumlah data terlalu banyak.

Program harus menunjukkan jumlah audio yang ditemukan, total audio dalam dataset, dan waktu eksekusi pencarian.

BAB II

TEORI SINGKAT

I. Principal Component Analysis (PCA)

Principal Component Analysis (PCA) dapat digunakan untuk mencari kumpulan audio melalui deteksi wajah berbagai orang, misalkan wajah seorang penyanyi. Setiap audio pastinya memiliki gambar albumnya sendiri masing-masing. Untuk memudahkan pengguna yang hanya memiliki gambar dari suatu album, maka dibutuhkan teknik Principal Component Analysis dengan proses berikut:

a. Pemrosesan Gambar

- Konversi gambar ke skala abu-abu (grayscale) untuk mengurangi kompleksitas dengan rumus berikut:

$$I(x,y) = 0.2989 \cdot R(x,y) + 0.5870 \cdot G(x,y) + 0.1140 \cdot B(x,y)$$

- Gambar diubah ke ukuran seragam dan direpresentasikan dalam bentuk vektor 1D. Jika gambar memiliki dimensi $M \times N$, maka hasilnya adalah vektor dengan panjang $M \cdot N$

$$I = [I_1, I_2, \dots, I_{M \cdot N}]$$

b. Standarisasi Data

- Data distandarisasi dengan cara mengurangi rata-rata setiap piksel, sehingga nilai data berada di sekitar nol;

$$\mu_j = \frac{1}{N} \sum_{i=1}^N x_{ij}$$

dimana:

- x_{ij} : nilai piksel ke- j pada gambar ke- i ,
- N : jumlah total gambar dalam dataset.

- Lalu kurangi piksel tersebut dengan rata-rata yang sudah dihitung untuk melakukan standarisasi.

$$x_{ij}' = x_{ij} - \mu_j$$

c. PCA dengan SVD

- Matriks kovarian dihitung untuk data terstandarisasi, kemudian dilakukan dekomposisi nilai singular (SVD) untuk mendapatkan komponen utama.

$$C = \frac{1}{N} X'^T X'$$

- X' : matriks data yang sudah distandarisasi.
- Lalu lakukan dekomposisi nilai singular untuk mendapatkan komponen utama.

$$\mathbf{C} = \mathbf{U}\Sigma\mathbf{U}^T$$

- \mathbf{U} : matriks eigenvector (komponen utama),
- Σ : matriks eigenvalue (menunjukkan varian data di sepanjang komponen utama).
- Lalu ambil n jumlah component utama teratas dari hasil SVD dan lakukan proyeksikan gambar ke komponen utama. Pilih k-komponen utama teratas ($k \ll M \cdot N$) dan proyeksikan data:

$$\mathbf{Z} = \mathbf{X}'\mathbf{U}_k$$

- \mathbf{U}_k : matriks eigenvector dengan n-dimensi.

d. Penghitungan Similaritas

- Similaritas antara gambar query dan dataset dihitung menggunakan jarak Euclidean.

- Lalu lakukan pengurutan kecocokan dari yang paling tinggi direpresentasikan dengan gambar query dalam ruang komponen utama dengan proyeksi yang sama;

$$\mathbf{q} = (\mathbf{q}' - \mu)\mathbf{U}_k$$

Dimana:

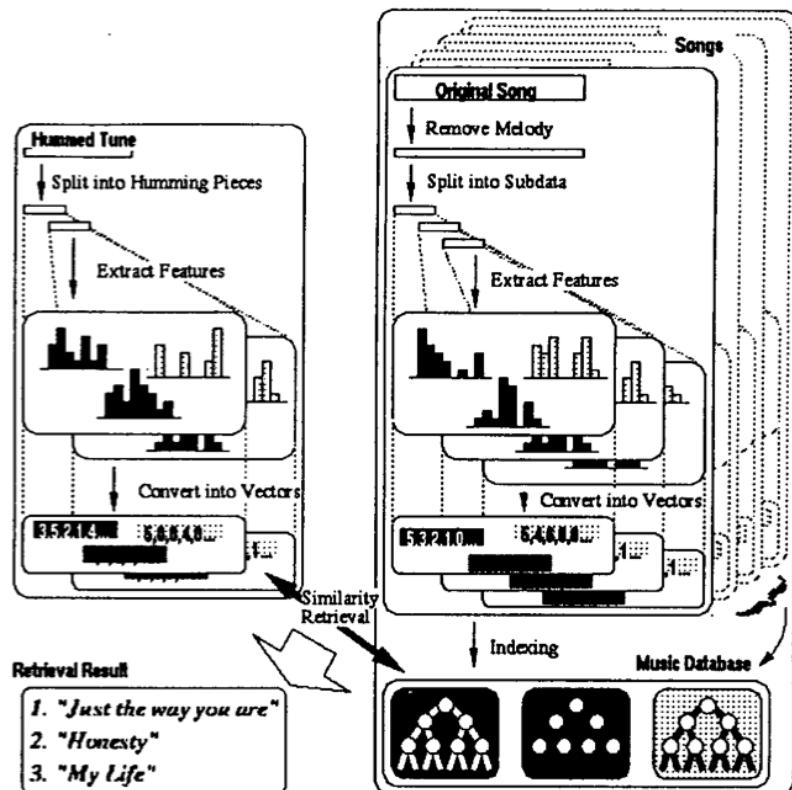
- \mathbf{q} = Vektor proyeksi dari gambar query ke ruang komponen utama (PCA).
- \mathbf{q}' : Gambar query dalam format vektor (setelah grayscale, resize, dan flattening).
- μ : Rata-rata piksel dari dataset (per piksel).
- \mathbf{U}_k : Matriks eigenvector dengan k dimensi utama dari PCA.
- Kemudian, hitung jarak Euclidean antara gambar query dengan semua gambar dalam dataset:

$$d(\mathbf{q}, \mathbf{z}_i) = \sqrt{\sum_{j=1}^k (q_j - z_{ij})^2}$$

- $d(\mathbf{q}, \mathbf{z}_i)$ = Jarak antara gambar query \mathbf{q} dan gambar ke- i dalam ruang komponen utama.
- \mathbf{z}_i = Vektor proyeksi dari gambar ke- i dalam dataset ke ruang komponen utama.
- q_j : Elemen ke- j dari vektor proyeksi query \mathbf{q} .
- z_{ij} : Elemen ke- j dari vektor proyeksi gambar ke- i , yaitu \mathbf{z}_i .
- k : Jumlah dimensi ruang komponen utama yang dipilih.
- Gambar dengan jarak terkecil dianggap paling mirip.

II. MIR

Music Information Retrieval (MIR) bidang dalam ilmu komputer dan musik yang fokus pada teknik dan metode untuk menganalisis, memahami, mencari, dan mengorganisasi informasi dari data musik. Tujuan utama MIR adalah mengembangkan teknologi yang memungkinkan komputer memahami musik dengan cara yang mirip dengan manusia, baik dari segi struktur, melodi, ritme, maupun konteks.



Gambar 1. Tahapan pada MIR dengan metode Query By Humming

1. Pemrosesan Audio

- File audio (format MIDI/WAV) diproses untuk fokus pada melodi utama.

- Teknik *windowing* membagi melodi menjadi segmen 20-40 beat dengan sliding window 4-8 beat.
- Proses normalisasi tempo dan pitch dilakukan untuk mengurangi variasi suara humming.

2. Ekstraksi Fitur

Tiga metode ekstraksi fitur digunakan:

- Absolute Tone Based (ATB): Menghitung distribusi nada absolut dalam rentang MIDI (0-127).
- Relative Tone Based (RTB): Menghitung perubahan antar nada untuk menangkap pola melodi (-127 hingga +127).
- First Tone Based (FTB): Menganalisis nada relatif terhadap nada pertama sebagai referensi (-127 hingga +127).
- Normalisasi: memastikan bahwa semua nilai dalam histogram berada dalam skala probabilitas;

$$H_{norm} = \frac{H[d]}{\sum_d H[d]}$$

*Dimana H adalah Histogram dan d adalah bin dari histogram tersebut.

3. Penghitungan Similaritas

Kemiripan dihitung menggunakan Cosine Similarity, yang mengukur seberapa mirip dua vektor berdasarkan sudut cosinusnya.

- Cosine Similarity mendekati 1 menunjukkan tingkat kemiripan yang tinggi. Berikut rumusnya:

$$\cos(\theta) = \frac{\mathbf{A} \cdot \mathbf{B}}{\|\mathbf{A}\| \|\mathbf{B}\|} = \frac{\sum_{i=1}^n A_i B_i}{\sqrt{\sum_{i=1}^n A_i^2} \sqrt{\sum_{i=1}^n B_i^2}}$$

Gambar 2. Cosine Similarity Formula

BAB III

ARSITEKTUR WEBSITE DAN PROGRAM INFORMATION RETRIEVAL

Website kami dirancang untuk mendukung sistem information retrieval, dengan fokus utama pada pencarian kemiripan album dan audio. Pada bagian Front-End, website ini menggunakan Vite, React, dan JavaScript untuk menciptakan antarmuka pengguna yang responsif, cepat, dan efisien. Kami juga mengintegrasikan Three.js untuk mendukung server-side rendering serta rendering 3D, sehingga performa website tetap optimal di berbagai perangkat.

Pada bagian Back-End, proses retrieval dikelola menggunakan Python, yang terbagi menjadi dua bagian utama: image retrieval dan audio retrieval. Untuk image retrieval, library seperti Pillow digunakan untuk mengolah gambar, seperti mengonversi matriks RGB menjadi vektor, sementara NumPy dimanfaatkan untuk perhitungan matematis dan manipulasi data numerik. Pada audio retrieval, kami memanfaatkan library basic-pitch untuk mengonversi file audio (seperti MP3 atau WAV) ke format MIDI. Selanjutnya, NumPy dan math digunakan untuk menghitung metrik seperti rata-rata, standar deviasi, dan histogram nada, yang meliputi Absolute Tone-Based (ATB), Relative Tone-Based (RTB), dan First Tone-Based (FTB).

Manajemen file dan integrasi database juga menjadi aspek penting dari pengembangan website ini. Library seperti zipfile digunakan untuk menangani file terkompresi. Library seperti shutil dan os digunakan untuk membuat folder sementara sebagai tempat penyimpanan file yang diunduh dari database. Durasi

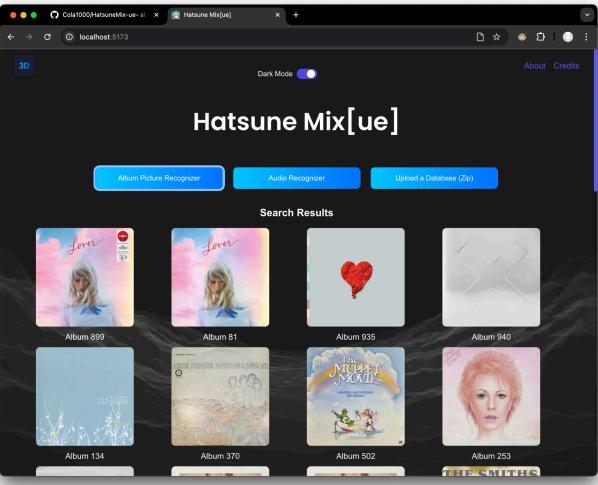
eksekusi program dicatat menggunakan modul time untuk memastikan performa yang optimal.

Website ini tidak hanya dirancang untuk memberikan solusi yang andal dalam sistem information retrieval, tetapi juga menghadirkan elemen hiburan dan orisinalgan kombinasi ini, sistem mampu melakukan pencocokan yang akurat untuk lagu potongan maupun versi lengkapnya.itas, memberikan pengalaman pengguna yang menarik dan menyenangkan.

BAB IV

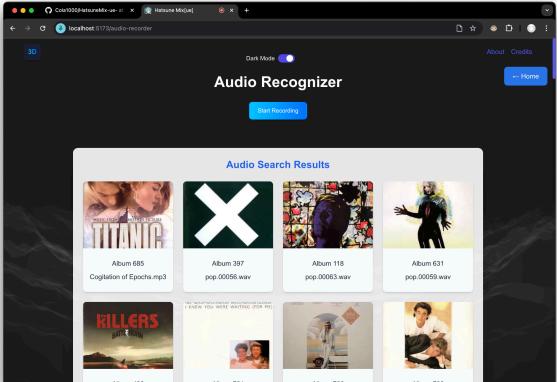
EKSPERIMEN

4.1. Album Picture Finder

Query	Keluaran
	

```
1  {
2      "similarity_rank": 1,
3      "similarity_percentage": 100.0,
4      "id": 899,
5      "title": "Album 899",
6      "imageSrc": "src/backend/database/picture/R-14401235-1573773330-8576-jpeg.jpg",
7      "songs": [
8          {
9              "id": "899_1",
10             "file": "pop.00087.wav"
11         }
12     ],
13 },
14 {
15     "similarity_rank": 2,
16     "similarity_percentage": 92.93,
17     "id": 81,
18     "title": "Album 81",
19     "imageSrc": "src/backend/database/picture/R-14035434-1570872261-6195-jpeg.jpg",
20     "songs": [
21         {
22             "id": "81_1",
23             "file": "pop.00002.wav"
24         }
25     ],
26 },
27 {
28     "similarity_rank": 3,
29     "similarity_percentage": 84.69,
30     "id": 935,
31     "title": "Album 935",
32     "imageSrc": "src/backend/database/picture/R-1617049-1624331284-3666-jpeg.jpg",
33     "songs": [
34         {
35             "id": "935_1",
36             "file": "pop.00038.wav"
37         }
38     ],
39 },
40 {
41     "similarity_rank": 4,
42     "similarity_percentage": 83.91,
43     "id": 940,
44     "title": "Album 940",
45     "imageSrc": "src/backend/database/picture/R-9595898-1484385049-3470-jpeg.jpg",
46     "songs": [
47         {
48             "id": "940_1",
49             "file": "pop.00074.wav"
50         }
51     ],
52 },
53 {
54     "similarity_rank": 5,
55     "similarity_percentage": 83.54,
56     "id": 134,
57     "title": "Album 134",
58     "imageSrc": "src/backend/database/picture/R-386751-1327155923-jpeg.jpg",
59     "songs": [
60         {
61             "id": "134_1",
62             "file": "pop.00012.wav"
63         }
64     ],
65 }
```

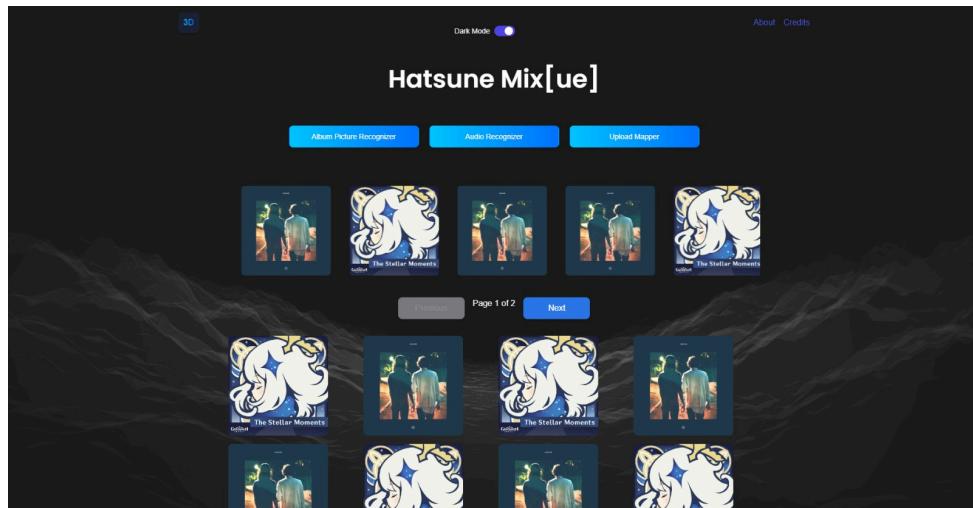
4.2. Music Information Retrieval (MIR)

Query	Keluaran
Cogitation of Epochs_trimm ed.mp3	 A screenshot of a web-based music search application titled "Audio Recognizer". The interface includes a "Dark Mode" toggle, a "Start Recording" button, and a "Home" link. Below these are sections for "Audio Search Results" and "Recent Searches". The "Audio Search Results" section displays eight album covers with their respective titles and IDs: Album 685 (Cogitation of Epochs.mp3), Album 397 (pop.00056.wav), Album 118 (pop.00063.wav), Album 631 (pop.00059.wav), Album 496 (KILLERS - Killers), Album 721 (I KNOW YOU WERE WATCHING, DICKIE BOY), Album 708 (I KNOW YOU WERE WATCHING, DICKIE BOY), and Album 793 (I KNOW YOU WERE WATCHING, DICKIE BOY). The "Recent Searches" section shows a list of terms: Cogitation of Epochs_trimm ed.mp3, pop.00056.wav, pop.00063.wav, pop.00059.wav, KILLERS - Killers, I KNOW YOU WERE WATCHING, DICKIE BOY, and I KNOW YOU WERE WATCHING, DICKIE BOY.

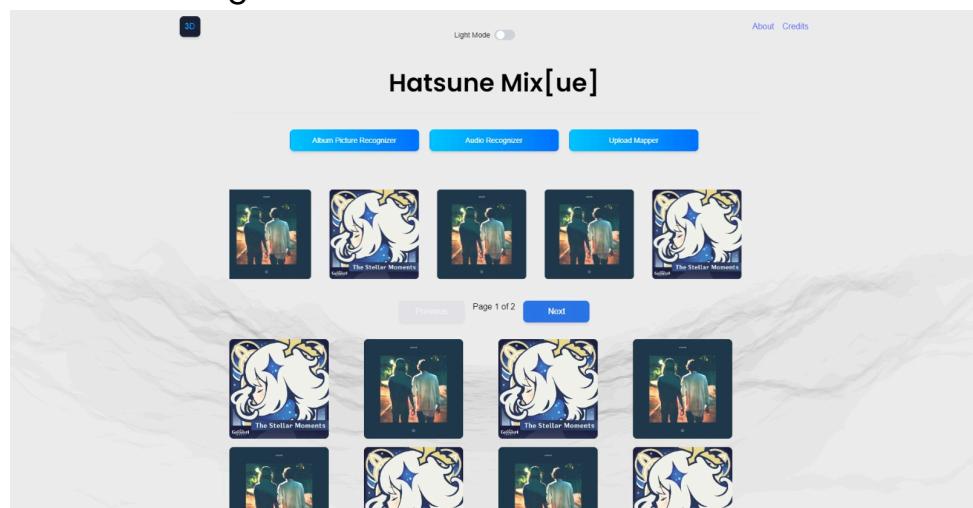
```
1  {
2      "similarity_rank": 1,
3      "similarity_percentage": 0.9842,
4      "id": 685,
5      "title": "Album 685",
6      "imageSrc": "src/backend/database/picture/R-467990-1185556320-jpeg.jpg",
7      "song": {
8          "id": "685_1",
9          "file": "Cogitation of Epochs.mp3"
10     }
11 },
12 {
13     "similarity_rank": 2,
14     "similarity_percentage": 0.9809,
15     "id": 397,
16     "title": "Album 397",
17     "imageSrc": "src/backend/database/picture/R-2546873-1512920036-1732-jpeg.jpg",
18     "song": {
19         "id": "397_1",
20         "file": "pop.00056.wav"
21     }
22 },
23 {
24     "similarity_rank": 3,
25     "similarity_percentage": 0.9791,
26     "id": 118,
27     "title": "Album 118",
28     "imageSrc": "src/backend/database/picture/R-470173-1345458522-1643-jpeg.jpg",
29     "song": {
30         "id": "118_1",
31         "file": "pop.00063.wav"
32     }
33 },
34 {
35     "similarity_rank": 4,
36     "similarity_percentage": 0.9735,
37     "id": 631,
38     "title": "Album 631",
39     "imageSrc": "src/backend/database/picture/R-6793557-1430654646-5820-jpeg.jpg",
40     "song": {
41         "id": "631_1",
42         "file": "pop.00059.wav"
43     }
44 },
45 {
46     "similarity_rank": 5,
47     "similarity_percentage": 0.9725,
48     "id": 496,
49     "title": "Album 496",
50     "imageSrc": "src/backend/database/picture/R-3884382-1348025929-4300-jpeg.jpg",
51     "song": {
52         "id": "496_1",
53         "file": "pop.00061.wav"
54     }
55 }
```

4.3. UI

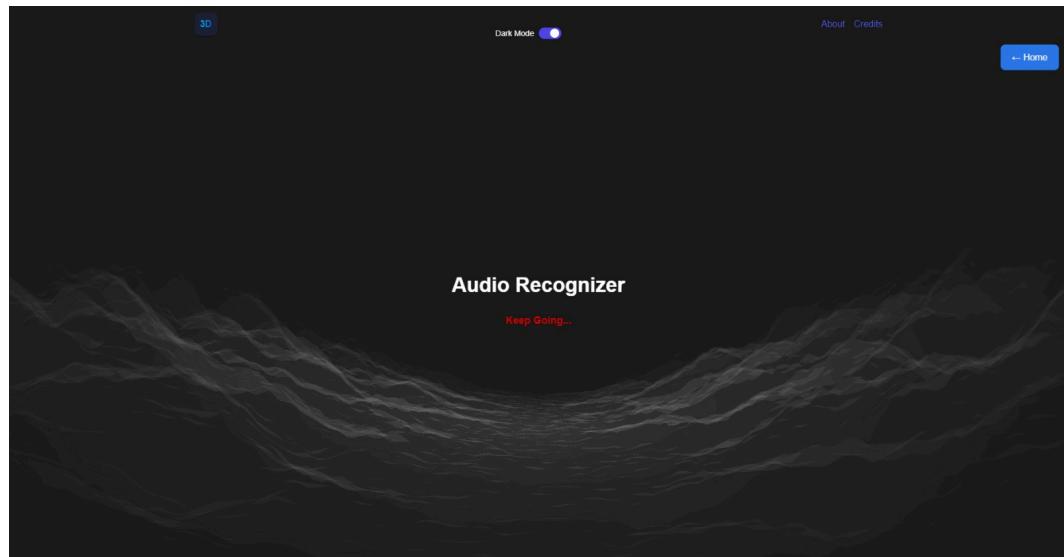
Gambar 4.3.1 | Landing Page



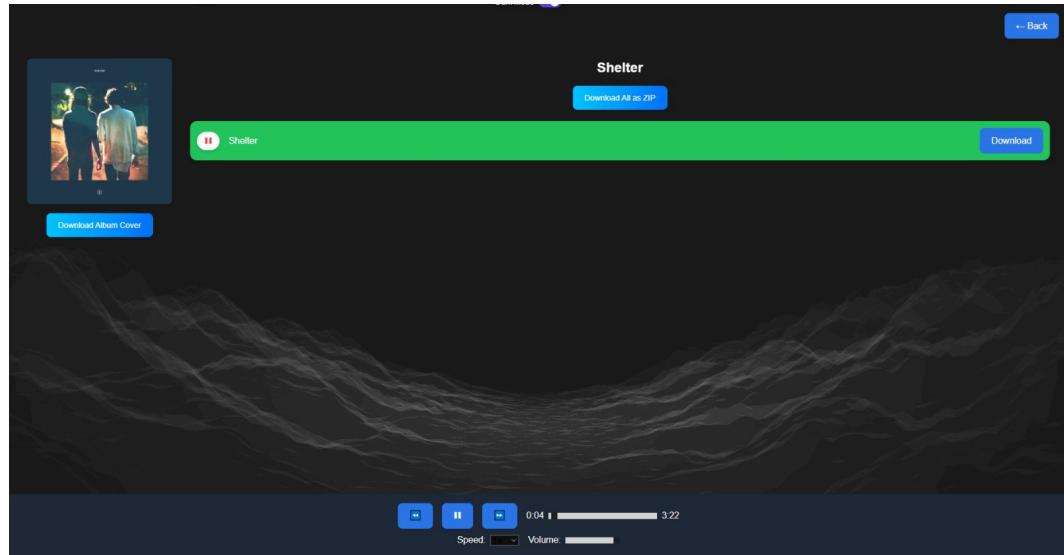
Gambar 4.3.3 | Light Mode



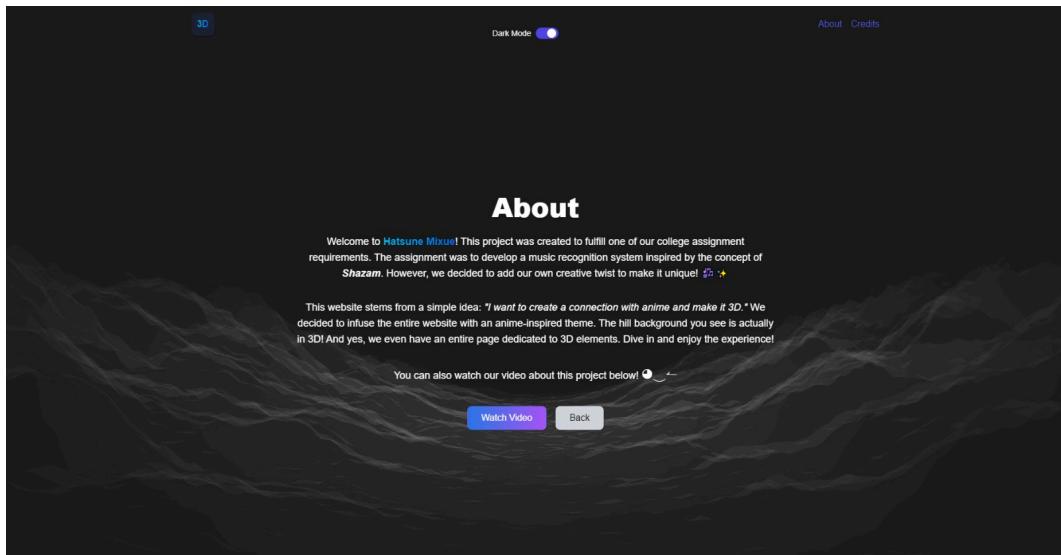
Gambar 4.3.3 | Humming/Singing Page



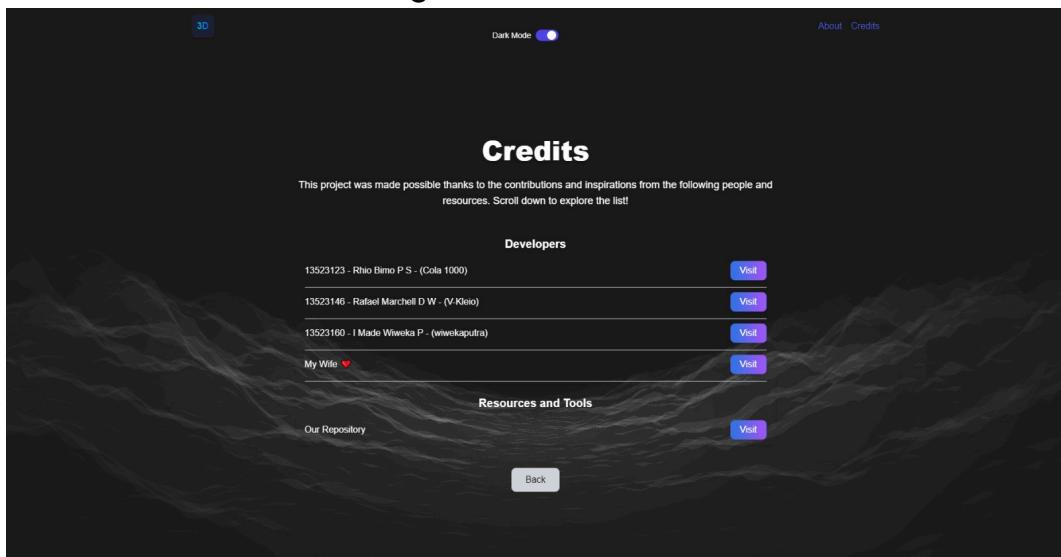
Gambar 4.3.4 | Music & Album Page



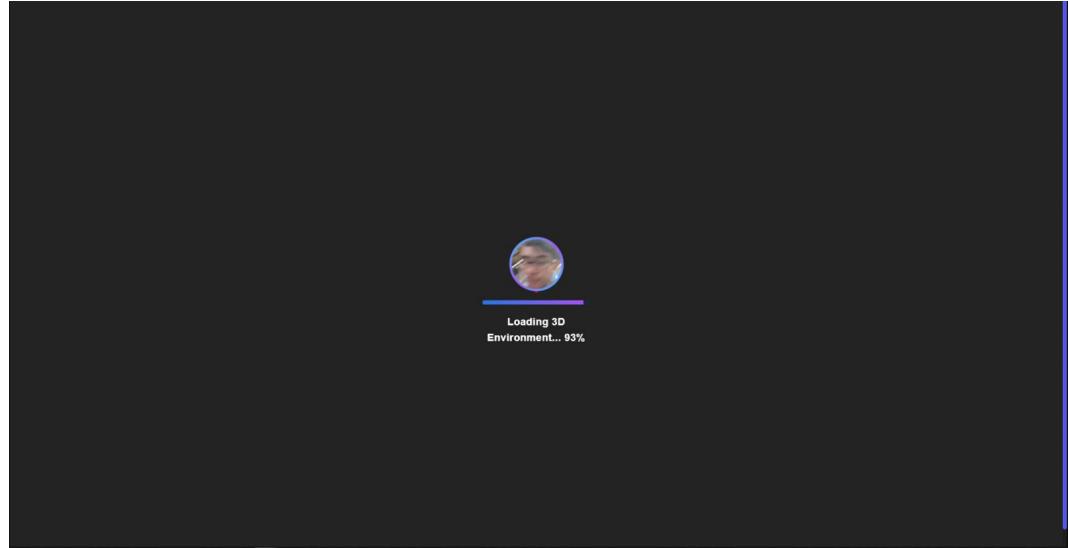
Gambar 4.3.5 | About Page



Gambar 4.3.6 | Credits Page



Gambar 4.3.7 | 3D Environment Loader



Gambar 4.3.8 | 3D Environment View



BAB V

KESIMPULAN

5.1. Kesimpulan

Website yang dikembangkan telah berhasil memenuhi spesifikasi utama yang diberikan. Sistem mampu menjalankan fungsi pencocokan suara menggunakan metode Query by Humming dan pencocokan gambar berbasis PCA. Dengan tingkat akurasi yang mencapai kisaran 55% hingga 100%, sistem ini telah menunjukkan kemampuan yang baik dalam mendeteksi suara dan gambar yang relevan dengan input pengguna. Fitur pagination juga berhasil diimplementasikan untuk memastikan tampilan data tetap efisien dan mudah diakses, terutama ketika dataset berukuran besar.

Walaupun telah memenuhi spesifikasi, terdapat beberapa keterbatasan yang perlu dicatat, seperti waktu eksekusi yang meningkat pada dataset berukuran besar dan sensitivitas algoritma terhadap variasi kualitas data input. Meskipun demikian, sistem ini telah menjadi landasan yang kuat untuk pengembangan lebih lanjut, terutama dengan potensi integrasi teknologi pembelajaran mesin untuk meningkatkan akurasi dan performa di masa depan.

5.2. Saran

Saran untuk pengembangan program lebih lanjut:

- a. Lebih mengembangkan antarmuka 3D yang lebih menarik namun *practical* juga.
- b. Melakukan optimasi algoritma untuk masalah yang lebih besar dan kompleks.
- c. Menambahkan dokumentasi kode.

5.3. Refleksi

Proyek ini memberikan banyak pelajaran berharga, terutama dalam memahami dan mengimplementasikan algoritma kompleks seperti PCA dan Query by Humming dalam sistem praktis berbasis web. Tantangan yang dihadapi selama pengembangan, seperti optimasi waktu eksekusi dan pengolahan format data yang bervariasi, mendorong kami untuk belajar lebih dalam tentang teknik pemrosesan data dan pengelolaan sistem.

Selain aspek teknis, proyek ini juga menjadi pengalaman penting dalam meningkatkan keterampilan kerja sama tim dan pengelolaan waktu. Kami belajar bahwa komunikasi yang efektif dan pembagian tugas yang jelas sangat krusial dalam menyelesaikan proyek yang melibatkan banyak komponen teknis. Refleksi ini memberikan motivasi untuk terus meningkatkan kemampuan di bidang pengembangan perangkat lunak dan ilmu data.

Walaupun sudah dikerjakan dari minggu pertama, mungkin kami harus mempertimbangkan lagi skill kami beserta ambisi kami ketika diberikan suatu tugas hehe. Rencana sih menurut kami sudah bagus, cuman kami-nya yang *skill issue* wkwkwk.

LAMPIRAN

Referensi:

Rinaldi Munir. "Aljabar dan Geometri untuk Informatika 2024/2025."
<https://informatika.stei.itb.ac.id/~rinaldi.munir/AljabarGeometri/2024-2025/algeo24-25.htm>

Tautan:

- Repository: [Hatsune Mix\[ue\]](#)
- Video:

 I bet you'd see this video then look at the views, it has n...