

**BAB III****ANALISIS DAN PERANCANGAN****3.1 Analisis**

Pada bagian analisis, terdapat penjelasan terkait identifikasi masalah dan pemecahan masalah dari penelitian berikut:

**3.1.1 Identifikasi Masalah**

Identifikasi lagu berdasarkan ingatan melodi sebagian tanpa informasi tambahan merupakan kendala bagi para penggemar musik untuk menemukan lagu yang sesuai. Akibatnya, pencarian lagu tersebut menjadi sebuah permasalahan, terutama ketika hanya terdapat sebagian melodi.

**3.1.2 Pemecahan Masalah**

Sistem pendeteksi lagu melalui voice recognition menggunakan CNN berbasis website dengan inputan humming manusia dapat diterapkan untuk mengatasi masalah identifikasi lagu berdasarkan ingatan melodis. Berikut adalah beberapa cara menerapkan sistem tersebut untuk mengatasi masalah dalam pengenalan lagu :

- A. Mengumpulkan dataset yang terdiri dari rekaman humming dari berbagai individu. Dataset ini akan digunakan untuk melatih dan menguji model.
- B. Menerapkan teknik ekstraksi fitur pada sinyal audio, baik dari lagu asli maupun humming. Teknik yang umum digunakan termasuk mel-spektrogram atau MFCC (Mel-Frequency Cepstral Coefficients). Proses ini mengubah sinyal audio mentah menjadi representasi yang lebih bermakna untuk pembelajaran mesin.
- C. Merancang struktur Convolutional Neural Network (CNN) yang sesuai untuk tugas pengenalan lagu. Arsitektur ini biasanya terdiri dari lapisan konvolusi, pooling, dan fully connected, dengan jumlah dan konfigurasi yang disesuaikan dengan kompleksitas tugas.
- D. Pelatihan dilakukan pada model CNN menggunakan data lagu yang telah dirancang sebagai dataset. Proses ini melibatkan pembagian

- data menjadi set pelatihan dan pengujian, normalisasi data, dan optimisasi parameter model untuk meningkatkan akurasi.
- E. Menguji performa model menggunakan set data pengujian dan melakukan penyesuaian jika diperlukan. Ini mungkin termasuk fine-tuning hyperparameter atau memodifikasi arsitektur model.
  - F. Pengembangan antarmuka website deteksi lagu, merancang dan mengimplementasikan antarmuka web yang memungkinkan pengguna untuk merekam humming mereka. Antarmuka ini harus user-friendly dan responsif.
  - G. Mengintegrasikan model CNN yang telah dilatih dengan backend sistem web. Ini termasuk pemrosesan input pengguna, ekstraksi fitur dari humming, dan penggunaan model untuk prediksi.
  - H. Implementasi sistem pencocokan Mengembangkan algoritma untuk mencocokkan hasil prediksi model dengan database lagu. Ini mungkin melibatkan teknik pencarian efisien atau penggunaan struktur data khusus.
  - I. Pengujian dan deployment sistem dengan melakukan pengujian menyeluruh pada sistem keseluruhan untuk memastikan akurasi, kecepatan respons, dan keandalan. Setelah itu, mendeploy sistem ke lingkungan produksi.
  - J. Pemeliharaan dan peningkatan berkelanjutan Secara berkala memperbarui dataset dan model untuk meningkatkan akurasi dan memperluas cakupan lagu yang dapat dikenali. Juga, mengoptimalkan performa sistem berdasarkan umpan balik pengguna dan analisis penggunaan.

### 3.2 Perancangan

Di tahap perancangan terdapat perancangan sistem yang akan menggambarkan alur kerja sistem agar lebih mudah dipakai.

#### 3.2.1 Perancangan Sistem

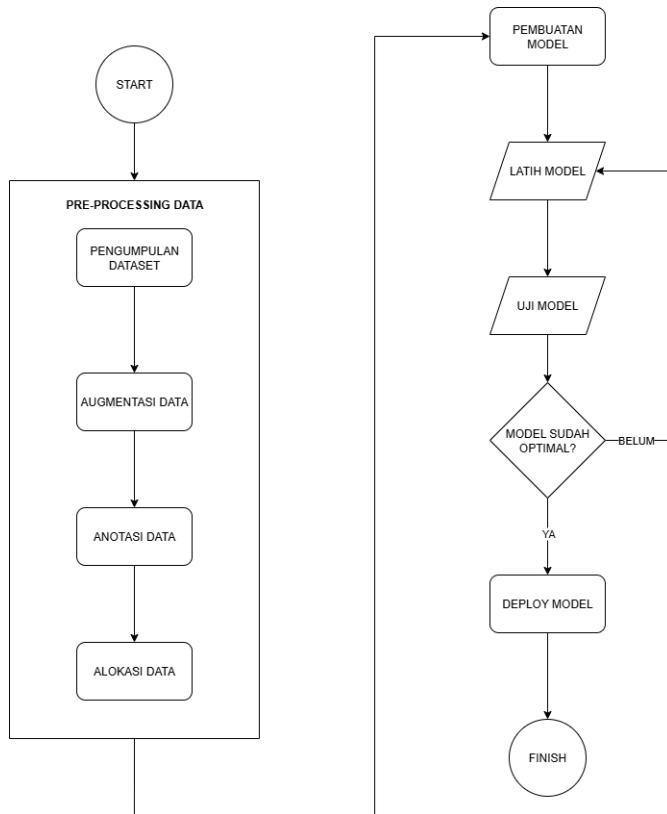
Di tahap ini akan digambarkan alur kerja dari sistem deteksi lagu secara umum dan cara kerja dari CNN.



Gambar 3.1 Alur Deteksi Lagu

### 3.2.2 Perancangan Sistem CNN

Proses dalam sistem deteksi lagu menggunakan CNN adalah sebagai berikut.



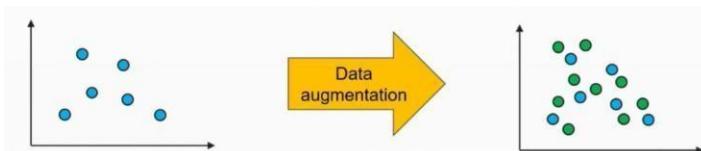
Gambar 3.2 Alur Sistem CNN

### 3.2.2.1 Pengumpulan Dataset

Pada perancangan sistem CNN pada penelitian ini penulis mengumpulkan dataset melalui rekaman humming langsung dari sukarelawan.

### 3.2.2.2 Augmentasi Data

Pre-processing data dibagi menjadi 3 tahap. Yang pertama yaitu augmentasi data. Di tahap ini data pelatihan akan diperbanyak dengan cara memanipulasi data yang ada. Teknik ini digunakan untuk meningkatkan performa model dengan menyediakan variasi data yang lebih banyak tanpa perlu mengumpulkan data baru.



Gambar 3.3 Ilustrasi Augmentasi Data

### 3.2.2.3 Anotasi Data

Anotasi data adalah proses memberikan label atau informasi tambahan pada data mentah untuk memberikan konteks atau identifikasi yang diperlukan untuk pelatihan model machine learning. Dalam konteks sistem pendekripsi lagu melalui humming dengan metode CNN, anotasi data biasanya melibatkan pemberian label pada data audio humming untuk menunjukkan lagu atau kategori apa yang sedang dihumming oleh pengguna.

### 3.2.2.4 Alokasi Data

Pembagian dataset dilakukan menjadi dua jenis, yaitu data training dan untuk melatih model untuk mendapat akurasi, lalu akan diuji oleh data testing.

### **3.2.2.5 Pelatihan Data**

Pelatihan data dalam sistem ini dimulai dengan tahap perancangan, yang melibatkan partisipasi pengguna dalam memberikan suara atau humming melalui perangkat ponsel. Suara yang diperoleh kemudian dikonversi ke dalam format WAV, lalu diproses untuk menghasilkan spektrogram. Langkah ini sangat penting untuk memastikan bahwa data yang digunakan dalam proses pelatihan model memiliki kualitas tinggi dan sesuai untuk dianalisis lebih lanjut.

Setelah proses ekstraksi fitur melalui spektrogram, data dimasukkan ke dalam model Convolutional Neural Network (CNN). Model ini terdiri atas beberapa lapisan yang memiliki fungsi spesifik dalam mengenali pola-pola pada data suara. Adapun lapisan-lapisan yang digunakan dalam arsitektur CNN ini antara lain:

#### **1. Convolutional Layer**

Lapisan ini bertugas untuk mengekstraksi fitur-fitur penting dari data input, seperti tepi, tekstur, atau pola, menggunakan filter atau kernel. Proses konvolusi ini menghasilkan feature maps yang menangkap karakteristik lokal dari data.

#### **2. Maxpooling 2D**

Lapisan ini digunakan untuk mengurangi dimensi spasial (lebar dan tinggi) dari feature maps, sehingga mengurangi jumlah parameter dan komputasi, sekaligus mencegah overfitting. Maxpooling mengambil nilai maksimum dari wilayah tertentu, mempertahankan fitur dominan.

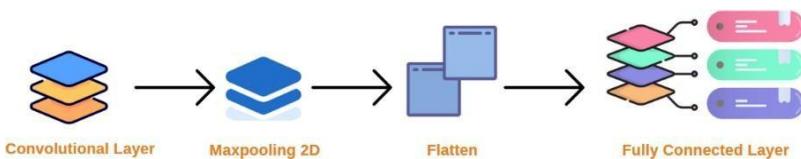
#### **3. Flatten**

Proses ini mengubah data multidimensi dari lapisan sebelumnya menjadi vektor satu dimensi. Langkah ini diperlukan untuk

menghubungkan output dari lapisan konvolusional ke lapisan fully connected.

#### 4. Fully Connected Layer

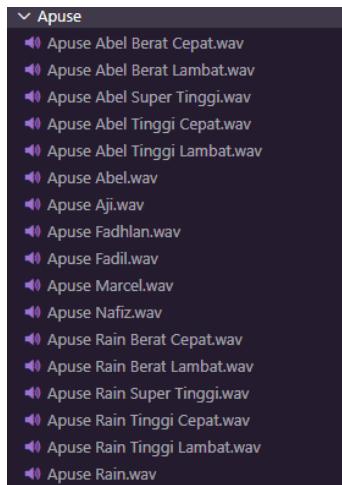
Lapisan ini mengintegrasikan fitur-fitur yang diekstraksi sebelumnya untuk melakukan klasifikasi atau regresi. Setiap neuron pada lapisan ini terhubung dengan semua neuron dari lapisan sebelumnya, memungkinkan pengambilan keputusan berdasarkan fitur yang telah diproses.



Gambar 3.4 Layer di Dalam CNN

#### 3.2.3 Perancangan Data

Dataset yang digunakan adalah sebagai berikut :



Gambar 3.5 Data Humming dari Lagu

Dataset yang digunakan adalah rekaman langsung dari sukarelawan yang di mana dataset tersebut berisi humming dengan berbagai variasi dari setiap lagu.

### 3.2.4 Perhitungan Sistem Deteksi

Proses deteksi lagu dari humming dimulai dengan merekam suara pengguna selama 10 detik melalui website, menghasilkan file WAV dengan 220500 titik data (sample rate 22050 Hz). Audio diproses untuk mengekstrak mel spectrogram, tempo, dan onset rate. Tempo dihitung dengan rumus:

$$\text{Tempo (BPM)} = \frac{\text{Jumlah beat}}{10} \times 60$$

Misalnya, deteksi 10 beat menghasilkan 60 BPM. Lalu dilanjutkan dengan Laju onset dihitung dengan:

$$\text{Laju Onset} = \frac{\text{Jumlah onset}}{10}$$

Contohnya, 20 onset menghasilkan 2 onset/detik. Setelah itu *Mel spectrogram* dibuat dengan mengubah suara ke skala desibel:

$$\text{mel\_db} = 10 \times \log_{10} \left( \frac{\text{Kekuatan suara}}{\text{Kekuatan maksimum}} \right)$$

Misalnya kekuatan 0.5 dan maksimum 1.0 menghasilkan -3.01 dB, lalu dinormalisasi:

$$\text{Nilai baru} = \frac{\text{Nilai lama} - \text{Rata-rata}}{\text{Standar deviasi}}$$

Contoh (misal dengan standar deviasi sebesar 20) :

$$\frac{-3.01 - (-50)}{20} \approx 2.35$$

Hasil fitur ini dimasukkan ke model CNN yang menghasilkan probabilitas lagu menggunakan layer CNN:

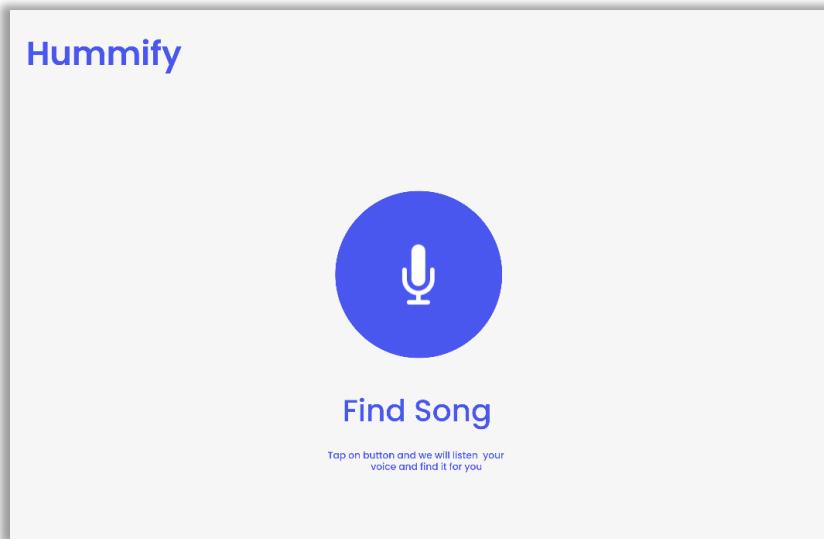
$$\text{Probabilitas}_i = \frac{e^{\text{nilai}_i}}{\sum e^{\text{nilai}_j}}$$

Misalnya nilai mentah (raw scores) yang dihasilkan oleh lapisan terakhir model Convolutional Neural Network (CNN) sebelum diterapkan fungsi softmax untuk menghasilkan probabilitas kelas lagu adalah berikut:

[-2.0, 1.5, -1.0, -2.5, -3.0], menghasilkan [0.03, 0.88, 0.07, 0.02, 0.01], mendekati [0.05, 0.80, 0.10, 0.03, 0.02] dengan "Apuse" (0.80) sebagai prediksi. Hasil ditampilkan di website dengan nama lagu dan tingkat tingkat keyakinan.

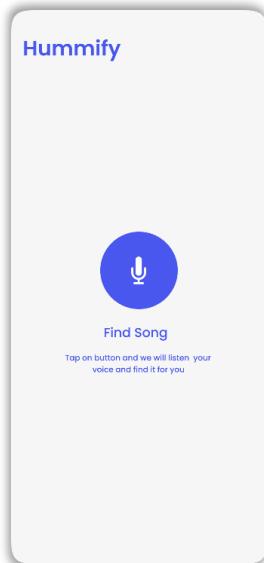
### 3.2.5 Perancangan User Interface / *Mock-up* aplikasi

Berikut adalah tampilan Mock-up dari website pendekripsi lagu menggunakan CNN



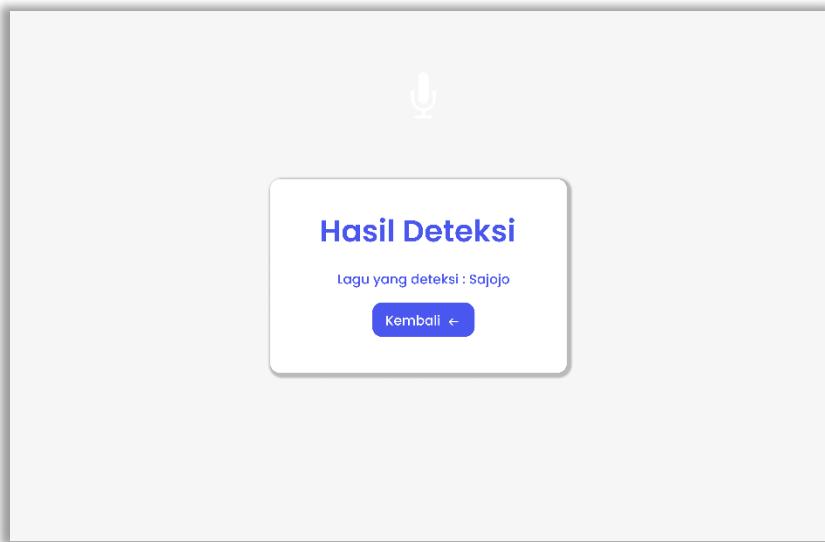
*Gambar 3.6 Halaman Home Website*

Tampilan utama aplikasi pendekripsi lagu, seperti terlihat pada gambar di atas, memiliki satu tombol yang berfungsi untuk menangkap suara humming sebagai input.



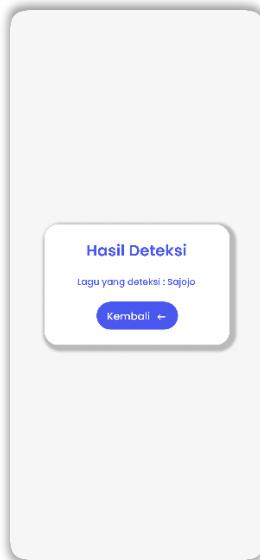
Gambar 3.7 Halaman Home Website Mobile

Tampilan utama aplikasi pendekripsi lagu, seperti terlihat pada gambar di atas, memiliki satu tombol yang berfungsi untuk menangkap suara humming sebagai input. Tampilan berbentuk mobile di website.



Gambar 3.8 Halaman Deteksi Web Desktop

Gambar di atas adalah halaman hasil dari deteksi lagu jika lagu berhasil ditemukan dengan tampilan website. Terdapat judul lagu yang akan ditampilkan.



Gambar 3.9 Halaman Deteksi Web Mobile

Gambar di atas adalah halaman hasil dari deteksi lagu jika lagu berhasil ditemukan dengan tampilan yang disesuaikan dengan tampilan web di mobile. Terdapat judul lagu yang akan ditampilkan.

### 3.3 Rancangan Pengujian

Berikut adalah beberapa rancangan pengujian yang dilakukan.

#### 3.3.1 Black Box Testing

Metode black box digunakan untuk menguji perangkat lunak berdasarkan spesifikasi yang ada, dengan menilai keluaran dari input tertentu tanpa memperhatikan struktur internal sistem. Pengujian ini bertujuan untuk mengetahui apakah sistem telah berjalan sebagaimana mestinya.

Dalam penelitian ini, penentuan kategori hasil pengujian Black Box Testing dilakukan dengan membandingkan antara output yang diharapkan (expected output) dengan output aktual

(actual output) yang dihasilkan oleh sistem. Agar hasil pengujian dapat dievaluasi secara konsisten, ditetapkan tiga kategori interpretasi prediksi, yaitu:

**1. Prediksi Akurat**

Hasil keluaran sistem sesuai dengan output yang diharapkan dan nilai keyakinan (*confidence score*) model berada pada kisaran tinggi ( $\geq 80\%$ ). Kategori ini menunjukkan bahwa sistem berhasil mengenali input humming dengan tepat dan yakin.

**2. Prediksi Lumayan**

Hasil keluaran sistem sesuai dengan output yang diharapkan, namun tingkat keyakinan model relatif moderat (50% – 79%). Hal ini menunjukkan bahwa meskipun sistem mampu mengidentifikasi lagu dengan benar, tingkat keyakinannya belum optimal.

**3. Prediksi Rendah / Tidak Akurat**

Hasil keluaran sistem tidak sesuai dengan output yang diharapkan atau tingkat keyakinan model berada pada kisaran rendah ( $< 50\%$ ). Kategori ini menandakan bahwa sistem gagal memberikan prediksi yang dapat diandalkan.

Penetapan batasan kategori ini ditentukan oleh penulis berdasarkan kebutuhan penelitian serta sebagai acuan evaluasi fungsionalitas sistem. Dengan adanya kategori ini, hasil pengujian Black Box Testing tidak hanya menunjukkan benar atau salah, tetapi juga memberikan gambaran mengenai tingkat keyakinan sistem dalam menghasilkan prediksi.

### 3.3.2 Confusion Matrix

Dalam mengevaluasi model deteksi lagu melalui humming, kita menggunakan confusion matrix untuk menganalisis performa model secara detail. Sebagai contoh, misalkan kita memiliki dataset dengan 30 rekaman humming yang terdiri dari 10 rekaman untuk masing-masing dari 3 lagu berbeda (Lagu A, B, dan C).

Di sini humming Lagu A diprediksi benar, 1 humming Lagu A salah diprediksi sebagai Lagu B dan 1 humming Lagu A salah diprediksi sebagai Lagu C

Evaluasi sistem dilakukan menggunakan confusion matrix untuk memperoleh akurasi, presisi, recall, dan F1-score. Hasil perhitungan menunjukkan akurasi sebesar 77%. Presisi, recall, dan F1-score untuk Lagu A, B, dan C masing-masing bernilai 80%, 70%, dan 80%.

Hasil tersebut menunjukkan bahwa model memiliki performa yang relatif seimbang dalam mengenali ketiga lagu yang diuji. Performa pada Lagu B berada di bawah dua lagu lainnya. Perbedaan ini dapat disebabkan oleh faktor seperti tingkat kompleksitas melodi pada Lagu B atau variasi dalam pola humming yang diberikan oleh pengguna.