

LAPORAN TUGAS BESAR 1

Diajukan sebagai Pemenuhan Tugas Besar IF4071 Pemrosesan Ucapan



Oleh:

Denise Felicia Tiowanni	13522013
Erdianti Wiga Putri Andini	13522053
Shazya Audrea Taufik	13522063
M. Naufal Aulia	13522074
Zahira Dina Amalia	13522085

SEKOLAH TEKNIK ELEKTRO DAN INFORMATIKA

INSTITUT TEKNOLOGI BANDUNG

2025

DAFTAR ISI

DAFTAR ISI.....	1
BAB I	
DESKRIPSI PERSOALAN.....	3
I.1 Latar Belakang.....	3
I.2 Rumusan Masalah.....	4
I.3 Tujuan.....	4
BAB II	
LANDASAN TEORI.....	6
II.1 Pengenalan Ucapan.....	6
II.2 Ekstraksi Fitur Menggunakan MFCC.....	6
II.2.1 Konsep Dasar MFCC.....	6
II.2.2 Tahapan Perhitungan MFCC.....	6
II.3 Dynamic Time Warping (DTW).....	7
II.4 Generalized Template dengan Model Rata-rata dan Kovarians.....	8
II.5 Perhitungan Jarak.....	9
BAB III	
PEMBAHASAN.....	10
III.1 Preprocessing dan Feature Extraction.....	10
III.2 Implementasi Dynamic Time Warping (DTW).....	12
III.2.1 Pembuatan Generalized Template (Training).....	12
III.2.2 Proses Pengenalan (Testing).....	13
III.3 Implementasi Skenario Pengujian.....	13
III.3.1 Pengaturan Data Latih dan Uji.....	14
III.3.2 Closed Scenario.....	14
III.3.3 Open Scenario.....	15
III.4 Implementasi Visualizer Grafik.....	17
III.4.1 Visualisasi Waveform dan VAD.....	17
III.4.2 Visualisasi Fitur MFCC 39 Dimensi (per Segmen).....	18
III.4.3 Visualisasi Penyelarasan DTW (Ilustrasi).....	19
III.4.4 Visualisasi Jarak Vokal (Bar Chart).....	19
III.4.5 Visualisasi Confusion Matrix (Heatmap).....	20
BAB IV	
HASIL DAN ANALISIS.....	22
IV.1 Hasil Eksperimen.....	22
IV.1.1 Hasil Tes dengan Template ‘Wiga’.....	22
IV.1.2 Hasil Tes dengan Template ‘Zya’.....	37
IV.1.3 Hasil Tes dengan Template ‘Hira’.....	53

IV.1.4 Hasil Tes dengan Template ‘Densu’	69
IV.1.5 Hasil Tes dengan Template ‘Naufal’	84
IV.1.6 Hasil Tes Sampel 1 (Template Kelompok).....	101
IV.1.7 Hasil Tes Sampel 2 (Template Kelompok).....	102
IV.1.8 Hasil Tes Sampel 3 (Template Kelompok).....	103
IV.1.9 Hasil Tes Sampel 1 (Template Kelompok Lain).....	104
IV.1.10 Hasil Tes Sampel 2 (Template Kelompok Lain).....	106
IV.1.11 Hasil Tes Sampel 3 (Template Kelompok Lain).....	107
IV.1.12 Hasil Confusion Matrix.....	109
IV.2 Analisis.....	109
IV.2.1 Pola Fitur MFCC (Vektor Rata-rata).....	109
IV.2.2 Analisis Matriks Kovarians.....	110
IV.2.3 Analisis Confusion Matrix.....	110
IV.2.4 Analisis Detail dari Visualisasi Per-Sampel.....	111
IV.2.5 Masalah yang Diidentifikasi.....	114
IV.2.6 Analisis Keterbatasan Dataset dan Bias Gender.....	114
BAB V	
KESIMPULAN DAN SARAN.....	117
V.1 Kesimpulan.....	117
V.2 Saran.....	118
DAFTAR REFERENSI.....	120
LAMPIRAN.....	121

BAB I

DESKRIPSI PERSOALAN

I.1 Latar Belakang

Pengenalan ucapan (*speech recognition*) merupakan salah satu cabang penting dalam bidang pengolahan sinyal digital dan kecerdasan buatan. Sistem pengenalan ucapan memungkinkan komputer untuk memahami dan menafsirkan sinyal suara manusia sehingga dapat digunakan dalam berbagai aplikasi seperti asisten virtual, sistem navigasi suara, hingga perangkat bantu bagi penyandang disabilitas.

Salah satu komponen dasar dalam pengenalan ucapan adalah kemampuan sistem untuk mengenali fonem atau vokal dasar dari suatu bahasa. Pada Bahasa Indonesia, terdapat lima vokal utama, yaitu a, i, u, e, o, yang menjadi fondasi bagi pembentukan kata dan kalimat. Oleh karena itu, sistem pengenalan vokal yang andal menjadi langkah awal menuju sistem pengenalan ucapan yang lebih kompleks.

Dalam proyek ini, dibangun sebuah sistem pengenalan suara, khususnya vokal Bahasa Indonesia menggunakan metode *Dynamic Time Warping* (DTW) dengan ekstraksi fitur *Mel-Frequency Cepstral Coefficients* (MFCC) berdimensi 39. DTW digunakan untuk mencocokkan pola waktu antar sinyal ucapan yang memiliki panjang atau tempo berbeda, sedangkan MFCC berfungsi sebagai representasi ciri akustik dari sinyal suara.

Sistem ini dikembangkan dalam bahasa Python dan diuji menggunakan dataset yang direkam oleh anggota kelompok dan kelompok lain. Evaluasi dilakukan pada dua skenario, yaitu *closed scenario* (pengujian dengan pembicara yang sama dengan *template*) dan *open scenario* (pengujian dengan pembicara berbeda). Tujuan akhirnya adalah memperoleh sistem yang mampu mengenali vokal secara akurat meskipun terdapat variasi antar pembicara.

I.2 Rumusan Masalah

Berdasarkan latar belakang tersebut, rumusan masalah dalam proyek ini adalah sebagai berikut:

1. Bagaimana cara membangun sistem pengenalan vokal Bahasa Indonesia menggunakan metode *Dynamic Time Warping* (DTW) dengan pendekatan *Generalized Template* (rata-rata dan *covariance*)?
2. Bagaimana cara mengekstraksi fitur MFCC 39 dimensi yang mencakup MFCC, Delta, dan Delta-Delta dari sinyal suara?
3. Bagaimana kinerja sistem dalam mengenali lima vokal Bahasa Indonesia pada dua skenario pengujian, yaitu *closed* dan *open scenario*?
4. Seberapa tinggi tingkat akurasi yang dapat dicapai oleh sistem dalam kedua skenario tersebut?

I.3 Tujuan

Adapun tujuan dari pembuatan proyek ini adalah sebagai berikut:

1. Mengimplementasikan sistem pengenalan vokal menggunakan *Dynamic Time Warping* (DTW) pada data ucapan berformat audio.
2. Melakukan ekstraksi fitur suara menggunakan MFCC 39 dimensi yang terdiri atas MFCC, Delta, dan Delta-Delta.
3. Melakukan evaluasi performa sistem pada dua skenario pengujian (*closed* dan *open*) serta menghitung akurasi pengenalan.
4. Menyediakan visualisasi hasil pengolahan sinyal, seperti waveform, spektrum MFCC, DTW *alignment path*, dan *confusion matrix*.

I.4 Batasan Masalah

Proyek yang dikembangkan memiliki batasan sebagai berikut:

1. Sistem hanya mengenali lima vokal Bahasa Indonesia: a, i, u, e, o.
2. Ekstraksi fitur dilakukan menggunakan MFCC 39 dimensi (13 MFCC, 13 Delta, dan 13 Delta-Delta) dari *library* `python_speech_features`.
3. Metode pencocokan pola menggunakan *Dynamic Time Warping* (DTW) dengan model *Generalized Template* berbasis rata-rata dan kovariansi.

4. Implementasi dilakukan menggunakan Python tanpa menggunakan model pembelajaran mendalam (*deep learning*).
5. Evaluasi hanya mencakup perhitungan akurasi *closed*, *open*, dan rata-rata keduanya, tanpa memperhitungkan aspek waktu eksekusi atau efisiensi komputasi.

BAB II

LANDASAN TEORI

II.1 Pengenalan Ucapan

Pengenalan ucapan (*Automatic Speech Recognition* atau ASR) merupakan bidang dalam pemrosesan sinyal suara yang bertujuan untuk mengenali kata atau fonem yang diucapkan manusia dan mengubahnya menjadi representasi teks. ASR membandingkan sinyal ucapan masukan terhadap sekumpulan *template* atau model yang telah dilatih sebelumnya, sehingga mampu mengenali kata yang paling sesuai.

Salah satu tantangan utama dalam ASR adalah variasi pengucapan antar pembicara, kecepatan bicara yang berbeda, serta adanya kebisingan lingkungan. Oleh karena itu, sistem ASR membutuhkan proses ekstraksi ciri (*feature extraction*) yang baik serta algoritma pencocokan yang mampu menyesuaikan perbedaan waktu dan tempo ucapan.

II.2 Ekstraksi Fitur Menggunakan MFCC

II.2.1 Konsep Dasar MFCC

Mel Frequency Cepstral Coefficients (MFCC) merupakan metode ekstraksi fitur yang paling populer dalam sistem pengenalan ucapan. Metode ini merepresentasikan amplop spektrum daya (*power spectrum envelope*) dari sinyal suara secara efisien dan sesuai dengan persepsi pendengaran manusia. MFCC meniru cara kerja koklea di telinga manusia, yakni sensitivitas terhadap perbedaan frekuensi yang menurun seiring meningkatnya frekuensi.

II.2.2 Tahapan Perhitungan MFCC

Secara umum, proses pembentukan fitur MFCC terdiri dari beberapa tahap utama:

1. *Pre-emphasis*

Sinyal suara dilewatkan melalui filter untuk memperkuat komponen frekuensi tinggi yang lemah (akibat sifat alami pita suara manusia).

2. *Framing*

Sinyal dibagi menjadi potongan pendek (*frame*) berdurasi 20–25 ms dengan *overlap* sekitar 10 ms.

3. *Windowing*

Setiap *frame* dikalikan dengan *window function* (biasanya *Hamming window*) untuk mengurangi *spectral leakage* pada saat melakukan transformasi *Fourier*.

4. *Discrete Fourier Transform (DFT)*

Digunakan untuk mengubah sinyal dari domain waktu menjadi domain frekuensi, sehingga dapat dianalisis spektrum dayanya.

5. *Mel Filterbank*

Spektrum hasil FFT dilewatkan pada sekumpulan filter (berbentuk seperti segitiga) yang tersusun sesuai skala mel. Filter di frekuensi rendah lebih rapat dibanding frekuensi tinggi yang meniru sensitivitas pendengaran manusia.

6. Log Energi

Energi dari setiap filter diubah ke skala logaritmik agar sesuai dengan persepsi telinga manusia terhadap intensitas suara.

7. *Cepstrum*

Dilakukan pada log-energi hasil *filterbank* untuk mendekorrelasi fitur dan menghasilkan koefisien cepstral.

MFCC umumnya terdiri atas 39 dimensi, yaitu 12 koefisien cepstral, 12 delta (perubahan antar frame), 12 delta-delta (percepatan), serta energi, delta energi, dan delta-delta energi. Fitur-fitur ini kemudian digunakan sebagai representasi numerik yang menggambarkan karakteristik fonem yang diucapkan.

II.3 *Dynamic Time Warping (DTW)*

Dynamic Time Warping (DTW) adalah algoritma pencocokan pola dengan *dynamic programming* untuk mengukur kesamaan antara dua urutan data yang mungkin berbeda dalam panjang atau kecepatan waktu. Dalam konteks pengenalan ucapan, DTW berfungsi untuk mencari penyelarasan antara sinyal masukan dengan template yang tersimpan dalam *speech corpus*.

DTW menghitung jarak antara vektor fitur pada setiap frame masukan dan frame template menggunakan pengukuran jarak, biasanya *Euclidean Distance*. Proses ini dilakukan pada *trellis* atau matriks dua dimensi yang merepresentasikan semua kemungkinan penyelarasan. Algoritma

kemudian mencari jalur dengan total biaya (*cost*) minimum dari awal hingga akhir sehingga menggambarkan kesesuaian terbaik antara dua sinyal.

Pada pendekatan *single template*, setiap kata direpresentasikan oleh satu contoh rekaman. Pendekatan tersebut sederhana namun tidak mampu menangani variasi pengucapan dengan baik. Untuk mengatasinya, digunakan pendekatan *multiple templates*, yaitu menyimpan beberapa template untuk setiap kata. Namun, pendekatan ini meningkatkan kompleksitas komputasi karena setiap masukan harus dibandingkan dengan banyak template.

II.4 Generalized Template dengan Model Rata-rata dan Kovarians

Untuk memperoleh sistem yang efisien dan tahan terhadap variasi antar pembicara, digunakan pendekatan *Generalized Template*. Ide utamanya adalah menggabungkan beberapa *template* dari pembicara berbeda menjadi satu model representatif dengan menghitung nilai rata-rata dan kovarians pada setiap segmen fitur. Oleh karena itu, prosesnya mencakup hal-hal berikut.

1. *Segmental Vector*

Fitur *template* dibagi ke dalam beberapa segmen sehingga tiap segmen berisi vektor-vektor fitur yang memiliki karakteristik mirip. Segmentasi dapat dilakukan secara uniform atau menggunakan metode *Segmental K-Means* agar variasi dalam setiap segmen diminimalkan.

2. *Averaging Segment*

Untuk setiap segmen, dihitung rata-rata vektor (*mean vector*) dari semua *frame* di dalamnya. Rata-rata tersebut menjadi representasi tunggal dari segmen dan mengurangi ukuran *template*.

3. *Covariance of Each Segment*

Selain rata-rata, setiap segmen juga menyimpan informasi tentang seberapa besar variasi antar vektor dalam segmen tersebut melalui matriks kovarians. Kovarians menggambarkan seberapa lebar distribusi data pada segmen tersebut. Dengan memasukkan kovarians ke dalam perhitungan jarak, sistem dapat menjadi lebih toleran terhadap variasi pengucapan.

II.5 Perhitungan Jarak

Jarak antar *frame* umumnya dihitung menggunakan *Euclidean Distance*, namun untuk *generalized template*, dapat digunakan pendekatan probabilistik dengan fungsi *likelihood* berdasarkan *multivariate Gaussian distribution*. Dalam hal ini, setiap segmen *template* dianggap sebagai distribusi *Gaussian* dengan rerata (m_j) dan kovarians (C_j) tertentu, lalu perbandingan dilakukan berdasarkan probabilitas bahwa vektor input berasal dari distribusi tersebut. Fungsi *Gaussian likelihood* didefinisikan sebagai:

$$Gaussian(x; m_j, C_j) = \frac{1}{\sqrt{(2\pi)^D |C_j|}} e^{-0.5(x-m_j)^T C_j^{-1} (x-m_j)}$$

Untuk keperluan perhitungan jarak, biasanya digunakan negative log-likelihood:

$$d(x, m_j) = -\log(Gaussian(x; m_j, C_j)) = 0.5 \log((2\pi)^D |C_j|) + 0.5(x - m_j)^T C_j^{-1} (x - m_j)$$

Jarak juga menyertakan kovarians untuk menormalkan kontribusi tiap dimensi fitur. Dengan demikian, segmen yang memiliki variansi besar akan memberikan jarak yang lebih kecil untuk perbedaan kecil sehingga lebih toleran terhadap variasi pembicara.

BAB III

PEMBAHASAN

III.1 *Preprocessing dan Feature Extraction*

Tahap preprocessing dan ekstraksi fitur merupakan fondasi dari sistem ini, yang bertujuan untuk mengubah sinyal audio mentah menjadi representasi numerik (vektor fitur) yang informatif. Proses ini diimplementasikan dalam *class* `VowelRecognitionDTW`, terutama pada metode `extract_mfcc_39`.

Alur implementasinya adalah sebagai berikut:

1. **Pemuatan Audio:** Sinyal audio dimuat menggunakan `librosa.load` dengan *sample rate* yang telah ditentukan (misalnya, 16000 Hz) untuk memastikan konsistensi.
2. **Pre-emphasis:** Filter pre-emphasis (`librosa.effects.preemphasis`) diterapkan pada sinyal untuk memperkuat komponen frekuensi tinggi. Ini berguna karena frekuensi tinggi cenderung memiliki amplitudo yang lebih rendah dan seringkali membawa informasi penting (forman) dalam pengenalan vokal.
3. **Voice Activity Detection (VAD):** Jika diaktifkan (`use_vad=True`), sinyal audio akan diproses oleh metode `voice_activity_detection`. Metode ini menggunakan `librosa.effects.split` untuk mengidentifikasi segmen-segmen audio yang mengandung suara (bukan hening) berdasarkan level desibel (`top_db=20`). Implementasi ini hanya mengambil bagian audio dari awal segmen suara pertama hingga akhir segmen suara terakhir, secara efektif membuang keheningan di awal dan akhir rekaman.
4. **Normalisasi Amplitudo:** Sinyal audio dinormalisasi dengan membaginya dengan nilai amplitudo maksimum absolutnya. Ini memastikan bahwa semua klip audio memiliki rentang volume yang serupa dan mencegah fitur didominasi oleh perbedaan kenyaringan.

5. Ekstraksi Fitur 39 Dimensi

Hal ini adalah inti dari ekstraksi fitur yang dibagi menjadi 3 bagian berikut.

- a. **13 MFCC (Fitur Statis):** 13 koefisien MFCC dasar diekstraksi menggunakan *library* `python_speech_features.mfcc`. Parameter kunci yang digunakan adalah:
 - `numcep=13`: Menghasilkan 13 koefisien MFCC.
 - `nfilt=26`: Menggunakan 26 *filterbank* Mel.
 - `nfft=512`: Ukuran *window* FFT.
 - `winstep=0.01` (10ms) dan `winlen=0.025` (25ms): Menunjukkan *frame shift* dan *frame length*.
 - `appendEnergy=True`: Koefisien pertama (C0) adalah energi dari *frame*.

```
mfcc_features = mfcc(audio, samplerate=sr, numcep=13, ...)
```

- b. **13 Delta (Fitur Dinamis):** Koefisien Delta (turunan pertama) dihitung dari 13 MFCC di atas menggunakan `python_speech_features.delta`. Fitur ini menangkap laju perubahan fitur MFCC dari waktu ke waktu.

```
delta_features = delta(mfcc_features, 2)
```

- c. **13 Delta-Delta (Fitur Akselerasi):** Koefisien Delta-Delta (turunan kedua) dihitung dari fitur Delta. Fitur ini menangkap akselerasi atau laju perubahan dari fitur Delta.

```
delta2_features = delta(delta_features, 2)
```

6. Ekstraksi MFCC (13 Dimensi):

Fitur MFCC diekstraksi menggunakan *library* `python_speech_features.mfcc`. Parameter kunci yang digunakan adalah:

- a. `numcep=13`: Menghasilkan 13 koefisien MFCC.
- b. `nfilt=26`: Menggunakan 26 *filterbank* Mel.
- c. `nfft=512`: Ukuran *window* FFT.
- d. `winstep=0.01` (10ms) dan `winlen=0.025` (25ms): Menunjukkan *frame shift* dan *frame length*.
- e. `appendEnergy=True`: Koefisien pertama (C0) adalah energi dari *frame*.

7. **Median Filtering:** Setiap *stream* koefisien MFCC (misalnya, semua C1, semua C2, dst.) dihaluskan menggunakan `scipy.signal.medfilt` dengan `kernel_size=3`. Ini membantu mengurangi variasi kecil atau *noise* antar-*frame*.
8. **Konkatenasi Fitur:** Ketiga set fitur (MFCC, Delta, Delta-Delta) digabungkan secara horizontal (`np.hstack`) untuk menghasilkan satu matriks fitur dengan total 39 dimensi ($13 + 13 + 13$) untuk setiap *frame*.

```
features_39 = np.hstack([mfcc_features, delta_features, delta2_features])
```

9. **Normalisasi Fitur (Z-Score):** Jika diaktifkan (`normalize=True`), seluruh matriks fitur 39 dimensi dinormalisasi menggunakan *Z-score* (*mean-variance normalization*). Rata-rata dari setiap kolom fitur dikurangkan, dan hasilnya dibagi dengan standar deviasi kolom tersebut. Ini memastikan setiap dimensi fitur memiliki rata-rata 0 dan standar deviasi 1.

III.2 Implementasi *Dynamic Time Warping* (DTW)

Sistem ini menggunakan implementasi DTW dengan pendekatan *Generalized Template* yang dimodelkan dengan statistik Gaussian dan diimplementasikan melalui metode `build_generalized_templates` serta `dtw_distance_generalized`. Pendekatan ini terdiri atas dua fase sebagai berikut.

III.2.1 Pembuatan *Generalized Template* (Training)

Proses ini (dijalankan oleh `build_generalized_templates`) bertujuan untuk membuat satu model statistik untuk setiap vokal (a, i, u, e, o) dari beberapa contoh audio.

1. Pengumpulan Fitur: Untuk satu vokal (misal, 'a'), sistem mengumpulkan semua fitur MFCC-39 dari semua file template yang terkait dengan vokal tersebut (dari semua pembicara dalam set data template).
2. Segmentasi Seragam: Setiap sekuens fitur (dari satu file audio) dibagi menjadi `n_segments` (misalnya, 3) segmen yang seragam ukurannya. Ini diimplementasikan dalam `segment_features`.
3. Kalkulasi Statistik Segmen: Sistem kemudian mengagregasi segmen-segmen ini.
 - a. Semua frame dari segmen 1 (dari semua file 'a') dikumpulkan.
 - b. Semua frame dari segmen 2 (dari semua file 'a') dikumpulkan, dan seterusnya.

- c. Untuk setiap kumpulan frame per segmen, metode `compute_segment_statistics` menghitung dua parameter statistik:
 - i. Vektor Rata-rata (Mean): Vektor 39-dimensi yang mewakili rata-rata fitur di segmen tersebut.
 - ii. Matriks Kovarians (*Covariance*): Matriks 39x39 yang mewakili hubungan dan varians antar dimensi fitur di segmen tersebut.
4. Model Vokal: Hasil akhirnya adalah *template* tergeneralisasi untuk vokal 'a' yang terdiri dari `n_segments` pasang (vektor rata-rata, matriks kovarians). Ini secara efektif menciptakan model *Gaussian Mixture Model* (GMM) sederhana yang dipaksakan mengalir dalam segmen-segmen yang kaku.

III.2.2 Proses Pengenalan (*Testing*)

Saat *file* audio uji masuk, proses pengenalan (dalam fungsi `recognize` dan `dtw_distance_generalized`) adalah sebagai berikut:

1. **Ekstraksi & Segmentasi:** Audio uji diproses melalui `extract_mfcc_39` dan `segment_features` untuk menghasilkan `n_segments` segmen fitur uji.
2. **Perhitungan Jarak:** Jarak dihitung menggunakan *Negative Gaussian Log-Likelihood* (`gaussian_log_likelihood`).
3. **Jarak Segmental:** Sistem membandingkan segmen secara kaku:
 - a. Jarak untuk segmen 1 adalah rata-rata *log-likelihood* dari semua *frame* di segmen 1 uji terhadap model Gaussian (*mean, cov*) dari segmen 1 *template*.
 - b. Proses ini diulang untuk semua segmen.
4. **Jarak Total:** Jarak total untuk satu vokal adalah rata-rata dari semua jarak segmental.
5. **Keputusan:** Jarak total dihitung untuk semua model vokal (a, i, u, e, o). Vokal yang menghasilkan *negative log-likelihood* total terendah (jarak terkecil) dipilih sebagai vokal yang dikenali.

III.3 Implementasi Skenario Pengujian

Sistem ini dirancang untuk dievaluasi menggunakan dua skenario berbeda, yang implementasinya diatur dalam file `main.py` dan dieksekusi oleh metode `evaluate_all_scenarios` di `VowelRecognitionDTW.py`.

III.3.1 Pengaturan Data Latih dan Uji

Dalam main.py, data dibagi berdasarkan foldernya:

- **templates_us:** Berisi audio dari speaker yang dikenali yaitu dari kelompok sendiri (misal: 'densu', 'hira', 'naufal', 'wiga', 'zya').
 - **Data Latih:** main.py memuat sebagian besar file dari folder ini (misal, file 1-2) sebagai data latih menggunakan recognizer.add_template.
 - **Data Uji (Closed):** File yang tersisa (misal, file 3) dari setiap speaker disimpan sebagai test_data_us.
- **templates_other:** Berisi audio dari speaker yang tidak dikenali yaitu dari kelompok lain.
 - **Data Uji (Open):** File dari folder ini tidak pernah digunakan untuk melatih model. File-file ini hanya digunakan sebagai data uji dan disimpan sebagai test_data_other.

III.3.2 Closed Scenario

Skenario ini menguji kemampuan model untuk mengenali vokal dari **speaker yang sama** dengan yang ada di data latih.

- **Implementasi:** Dijalankan oleh metode test_closed_scenario.
- **Proses:** Metode ini menggunakan model penuh yang telah dilatih menggunakan semua data latih dari templates_us. Model ini kemudian diujikan terhadap test_data_us.
- **Logika:** Karena speaker dalam test_data_us juga ada dalam data latih, ini disebut skenario "tertutup" (*speaker-dependent*).

```
def test_closed_scenario(self, test_data):
    """
    Skenario Closed
    """
    correct = 0
    total = 0
    results = []

    print(f"\n--- DETAILED CLOSED SCENARIO RESULTS ---")

    for audio_path, true_vowel, test_person_id in test_data:
        # Memanggil recognizer utama (yang sudah dilatih penuh)
        recognized_vowel, distance, all_distances, final_vowel_distances =
self.recognize(audio_path)

        is_correct = (recognized_vowel == true_vowel)
```

```

        if is_correct:
            correct += 1
        total += 1

    self.print_detailed_prediction(
        audio_path, true_vowel, test_person_id,
        recognized_vowel, distance, final_vowel_distances,
        is_correct, "CLOSED"
    )

    results.append({
        'audio': audio_path,
        'true': true_vowel,
        'predicted': recognized_vowel,
        'person': test_person_id,
        'correct': is_correct,
        'distance': distance,
        'final_vowel_distances': final_vowel_distances
    })

    accuracy = (correct / total) * 100 if total > 0 else 0
    print(f"\n  CLOSED SCENARIO SUMMARY: {correct}/{total} correct =
{accuracy:.2f}%")
    return accuracy, results

```

III.3.3 *Open Scenario*

Skenario ini menguji kemampuan model untuk mengenali vokal dari speaker baru yang tidak ada dalam data latih.

- **Implementasi:** Dijalankan oleh metode `test_open_scenario` di dalam loop `evaluate_all_scenarios`.
- **Proses:** Implementasinya sedikit berbeda. Kode ini mensimulasikan "Open Scenario" beberapa kali.
 1. Model dibangun ulang (`self.build_generalized_templates()`) hanya menggunakan template dari satuspeaker (misal, hanya 'densu').
 2. Model 'densu-only' ini kemudian diujikan terhadap data `test_data_other` (speaker yang sama sekali tidak dikenal).
 3. Proses ini diulang untuk speaker latih lainnya (misal, model 'hira-only' vs `test_data_other`, dst.).
- **Logika:** Karena `test_data_other` berisi speaker yang tidak ada di data latih (baik 'densu', 'hira', dsb.), ini adalah skenario "terbuka" (*speaker-independent*). Akurasi Open Scenario secara keseluruhan adalah rata-rata dari semua pengujian ini.


```

def test_open_scenario(self, test_data, template_person_ids):
    """
    Skenario Open
    """
    correct = 0
    total = 0
    results = []

    # 1. Simpan model penuh yang asli
    original_templates = self.templates.copy()
    original_generalized = self.generalized_templates.copy()

    # 2. Filter template agar hanya berisi speaker yang diinginkan
    filtered_templates = {}
    for vowel in self.templates:
        filtered_templates[vowel] = {
            pid: feat for pid, feat in self.templates[vowel].items()
            if pid in template_person_ids
        }

    # 3. Bangun ulang model HANYA dengan template yang sudah difilter
    self.templates = filtered_templates
    self.generalized_templates = {}
    self.build_generalized_templates()

    print(f"\n--- DETAILED OPEN SCENARIO RESULTS ---")
    print(f"Using templates from: {template_person_ids}")

    # 4. Jalankan pengenalan (sekarang menggunakan model terbatas)
    for audio_path, true_vowel, test_person_id in test_data:
        recognized_vowel, distance, all_distances, final_vowel_distances =
self.recognize(audio_path)

        is_correct = (recognized_vowel == true_vowel)
        if is_correct:
            correct += 1
            total += 1

        self.print_detailed_prediction(
            audio_path, true_vowel, test_person_id,
            recognized_vowel, distance, final_vowel_distances,
            is_correct, "OPEN"
        )

        results.append({
            'audio': audio_path,
            'true': true_vowel,
            'predicted': recognized_vowel,
            'person': test_person_id,
            'correct': is_correct,
            'distance': distance,
            'final_vowel_distances': final_vowel_distances
        })

    # 5. Kembalikan model ke kondisi semula (model penuh)

```

```

self.templates = original_templates
self.generalized_templates = original_generalized

accuracy = (correct / total) * 100 if total > 0 else 0
print(f"\n OPEN SCENARIO SUMMARY: {correct}/{total} correct =
{accuracy:.2f}%")
return accuracy, results

```

III.4 Implementasi *Visualizer* Grafik

Untuk membantu analisis, *class* VowelDTWVisualizer diimplementasikan menggunakan matplotlib dan seaborn. *Script* main.py memanggil beberapa fungsi visualisasi.

III.4.1 Visualisasi *Waveform* dan VAD

- **Tujuan:** Menampilkan bentuk gelombang audio dan menyoroti area yang dideteksi sebagai aktivitas suara (VAD) untuk memverifikasi proses *preprocessing*.
- Kode Pemanggil (main.py):

```

try:
    viz.plot_waveform_with_vad(
        sample_path,
        top_db=20,
        save_path=os.path.join(IMAGE_DIR,
f"{sample_base}_waveform_vad.png")
    )
    logger.log(" Waveform + VAD visualization created")
except Exception as e:
    logger.log(f" Error creating waveform visualization: {e}", "ERROR")

```

- Kode Implementasi (VowelDTWVisualizer.py):

```

def plot_waveform_with_vad(self, audio_path, top_db=20, save_path=None):
    y, sr = librosa.load(audio_path, sr=self.recognizer.sample_rate)
    y = librosa.effects.preemphasis(y)
    intervals = librosa.effects.split(y, top_db=top_db)

    plt.figure(figsize=(12, 3))
    times = np.arange(len(y))/sr
    plt.plot(times, y, linewidth=0.8, label='Sinyal audio')
    for i, (s, e) in enumerate(intervals):
        label = 'Aktivitas suara' if i == 0 else ''
        plt.axvspan(s/sr, e/sr, alpha=0.2, color='green', label=label)
    # ... (kode untuk judul, label, dan penyimpanan plot) ...
    plt.show()

```

III.4.2 Visualisasi Fitur MFCC 39 Dimensi (per Segmen)

- **Tujuan:** Memvisualisasikan 39 fitur sebagai tiga *heatmap* terpisah (MFCC, Delta, Delta-Delta) dan menunjukkan batas segmen pembagian rata.
- Kode Pemanggil (main.py):

```
try:
    viz.plot_mfcc39(
        sample_path,
        save_prefix=os.path.join(IMAGES_DIR, f"{sample_base}_mfcc39")
    )
    logger.log(" MFCC-39 visualization created")
except Exception as e:
    logger.log(f" Error creating MFCC visualization: {e}", "ERROR")
```

- Kode Implementasi (VowelDTWVisualizer.py):

```
def plot_mfcc39_with_segments(self, audio_path, save_prefix=None):
    feats = self.recognizer.extract_mfcc_39(audio_path)
    segments = self.recognizer.segment_features(feats)

    mfcc_13 = feats[:, :13].T
    delta_13 = feats[:, 13:26].T
    delta2_13 = feats[:, 26:39].T

    def _show_with_segments(mat, title, segments_data):
        plt.figure(figsize=(12, 4))
        librosa.display.specshow(mat, x_axis='time', cmap='viridis')
        # ... (kode untuk judul dan colorbar) ...

        n_frames = mat.shape[1]
        segment_size = n_frames // self.recognizer.n_segments
        for i in range(1, self.recognizer.n_segments):
            boundary = i * segment_size
            plt.axvline(x=boundary, color='red', linestyle='--', alpha=0.7,
linewidth=2)

        # ... (kode untuk teks segmen dan penyimpanan plot) ...
        plt.show()

    _show_with_segments(mfcc_13, "MFCC (13)", segments)
    _show_with_segments(delta_13, "Delta (13)", segments)
    _show_with_segments(delta2_13, "Delta-Delta (13)", segments)

def plot_mfcc39(self, audio_path, save_prefix=None):
    self.plot_mfcc39_with_segments(audio_path, save_prefix)
```

III.4.3 Visualisasi Penyelarasan DTW (Ilustrasi)

- **Tujuan:** Memberikan ilustrasi visual tentang cara kerja DTW *klasik* (bukan metode yang digunakan dalam pengenalan) dengan menunjukkan matriks biaya dan jalur *warping*.
- Kode Pemanggil (main.py):

```
# ... (logika 'if' untuk menemukan template) ...
try:
    # ... (mengambil template_feats dan test_feats) ...
    viz.plot_dtw_alignment(
        template_feats, test_feats,
        title=f"DTW {sample_true.upper()} | template:{first_pid} vs
test:{sample_person}",
        save_path=os.path.join(IMAGES_DIR,
f"{sample_base}_dtw_alignment.png")
    )
    logger.log(" DTW alignment visualization created")
except Exception as e:
    logger.log(f" Error creating DTW visualization: {e}", "ERROR")
```

- Kode Implementasi (VowelDTWVisualizer.py):

```
def plot_dtw_alignment(self, template_features, test_features,
title="Penyelarasan DTW", save_path=None):
    A = template_features[:, :13] # Hanya menggunakan 13 MFCC untuk
visualisasi
    B = test_features[:, :13]
    D = cdist(A, B, metric='euclidean') # Matriks biaya Euclidean
    C, wp = librosa.sequence.dtw(C=D) # Jalur warping

    plt.figure(figsize=(8, 6))
    plt.imshow(D.T, origin='lower', aspect='auto', cmap='hot')
    path_i = [p[0] for p in wp]
    path_j = [p[1] for p in wp]
    plt.plot(path_i, path_j, color='cyan', linewidth=2, label='Jalur
warping')
    # ... (kode untuk judul, label, dan penyimpanan plot) ...
    plt.show()
```

III.4.4 Visualisasi Jarak Vokal (*Bar Chart*)

- **Tujuan:** Menampilkan hasil akhir pengenalan untuk satu audio uji dalam bentuk diagram batang, menunjukkan jarak (skor *log-likelihood*) ke setiap model vokal.
- Kode Pemanggil (main.py):

```
try:
    pred, dist, _, final_d = recognizer.recognize(sample_path)
    viz.plot_vowel_distances_bar(
        final_d,
```

```

        title=f"Distances for {os.path.basename(sample_path)}
(true={sample_true}, pred={pred})",
        save_path=os.path.join(IMAGES_DIR,
f"{sample_base}_vowel_distances.png")
    )
    logger.log(f" Distance bar chart created (true={sample_true},
pred={pred})")
except Exception as e:
    logger.log(f" Error creating distance visualization: {e}", "ERROR")

```

- Kode Implementasi (VowelDTWVisualizer.py):

```

def plot_vowel_distances_bar(self, final_vowel_distances: dict,
true_vowel=None,
                        title="Jarak final per vokal", save_path=None):
    vowels = list(final_vowel_distances.keys())
    vals = [final_vowel_distances[v] for v in vowels]
    colors = ['green' if v == true_vowel else 'lightblue' for v in vowels]

    idx = np.arange(len(vowels))
    plt.figure(figsize=(8, 4))
    bars = plt.bar(idx, vals, color=colors, edgecolor='black',
linewidth=0.5)
    plt.xticks(idx, vowels)
    # ... (kode untuk judul, label, teks, dan penyimpanan plot) ...
    plt.show()

```

III.4.5 Visualisasi *Confusion Matrix (Heatmap)*

- **Tujuan:** Merangkum kinerja model di seluruh set data uji (misalnya, semua hasil *Closed Scenario*) dalam bentuk *heatmap* matriks konfusi.
- Kode Pemanggil (main.py):

```

try:
    # ... (logika untuk mengumpulkan all_closed_results) ...
    if len(all_closed_results) > 0:
        viz.plot_confusion_heatmap(
            all_closed_results,
            vowels=recognizer.vowels,
            title="Confusion Matrix (Closed Scenario)",
            save_path=os.path.join(IMAGES_DIR,
"confusion_matrix_closed.png")
        )
        logger.log("Confusion matrix heatmap created")
    # ... (penanganan error) ...

```

- Kode Implementasi (VowelDTWVisualizer.py):

```

def plot_confusion_heatmap(self, results, vowels=('a','i','u','e','o'),
                        title="Matriks Konfusi", save_path=None):

```

```

idx = {v:i for i,v in enumerate(vowels)}
cm = np.zeros((len(vowels), len(vowels)), dtype=int)
for r in results:
    if r['true'] in idx and r['predicted'] in idx:
        cm[idx[r['true']], idx[r['predicted']]] += 1

# ... (kode untuk menghitung akurasi per kelas) ...

plt.figure(figsize=(7, 5.5))
sns.heatmap(cm, annot=True, fmt='d', cmap='Blues',
            xticklabels=vowels, yticklabels=vowels,
            square=True, cbar_kws={'label': 'Jumlah'})
# ... (kode untuk judul, label, dan penyimpanan plot) ...
plt.show()

```

BAB IV

HASIL DAN ANALISIS

IV.1 Hasil Eksperimen

IV.1.1 Hasil Tes dengan *Template* 'Wiga'

```
=====
=== EVALUATION WITH TEMPLATES FROM WIGA ===
=====

Total test files (templates_us): 25
Total test files (templates_other): 50

[CLOSED SCENARIO] (all templates vs test from templates_us)

--- DETAILED CLOSED SCENARIO RESULTS ---

Test: densu - a 3.wav
Person: densu
Actual: a | Predicted: a | CORRECT
Best Distance: 70.2069
All distances:
  a: 70.2069 <- CHOSEN (CORRECT)
  o: 74.3216
  u: 76.2315
  e: 96.8769
  i: 101.2702

Test: densu - i 3.wav
Person: densu
Actual: i | Predicted: i | CORRECT
Best Distance: 53.9492
All distances:
  i: 53.9492 <- CHOSEN (CORRECT)
  u: 75.7062
  e: 78.5462
  a: 95.7307
  o: 100.7484

Test: densu - u 3.wav
Person: densu
Actual: u | Predicted: u | CORRECT
Best Distance: 55.6178
All distances:
  u: 55.6178 <- CHOSEN (CORRECT)
  o: 72.8369
  a: 75.8265
  i: 92.3293
  e: 105.2425

Test: densu - e 3.wav
Person: densu
Actual: e | Predicted: e | CORRECT
```

Best Distance: 30.6416
All distances:
e: 30.6416 <- CHOSEN (CORRECT)
i: 64.1177
u: 65.0695
a: 73.9451
o: 74.4279

Test: densu - o 3.wav
Person: densu
Actual: o | Predicted: a | WRONG
Best Distance: 54.9172
All distances:
a: 54.9172 <- CHOSEN
o: 61.5603 (CORRECT)
u: 65.1129
e: 101.6156
i: 108.8738

Test: hira - a 3.wav
Person: hira
Actual: a | Predicted: a | CORRECT
Best Distance: 63.2745
All distances:
a: 63.2745 <- CHOSEN (CORRECT)
u: 72.9438
o: 77.6447
i: 89.6612
e: 108.5596

Test: hira - i 3.wav
Person: hira
Actual: i | Predicted: u | WRONG
Best Distance: 80.0211
All distances:
u: 80.0211 <- CHOSEN
e: 81.3058
a: 89.3321
i: 90.1186 (CORRECT)
o: 94.0986

Test: hira - u 3.wav
Person: hira
Actual: u | Predicted: u | CORRECT
Best Distance: 52.2315
All distances:
u: 52.2315 <- CHOSEN (CORRECT)
o: 56.9199
a: 82.7891
i: 87.1686
e: 94.8692

Test: hira - e 3.wav
Person: hira
Actual: e | Predicted: e | CORRECT

Best Distance: 31.6649
All distances:
e: 31.6649 <- CHOSEN (CORRECT)
u: 67.2318
i: 67.2943
o: 82.4711
a: 93.1578

Test: hira - o 3.wav
Person: hira
Actual: o | Predicted: a | WRONG
Best Distance: 52.5049
All distances:
a: 52.5049 <- CHOSEN
o: 57.9618 (CORRECT)
u: 64.8192
e: 81.7117
i: 104.5562

Test: naufal - a 4.wav
Person: naufal
Actual: a | Predicted: a | CORRECT
Best Distance: 57.6231
All distances:
a: 57.6231 <- CHOSEN (CORRECT)
u: 72.1626
e: 78.4054
o: 83.8841
i: 89.6910

Test: naufal - i 4.wav
Person: naufal
Actual: i | Predicted: i | CORRECT
Best Distance: 59.4456
All distances:
i: 59.4456 <- CHOSEN (CORRECT)
e: 75.9879
u: 78.6571
a: 93.3244
o: 142.4249

Test: naufal - u 4.wav
Person: naufal
Actual: u | Predicted: u | CORRECT
Best Distance: 52.9969
All distances:
u: 52.9969 <- CHOSEN (CORRECT)
o: 68.3638
e: 83.6060
a: 83.9304
i: 103.4871

Test: naufal - e 4.wav
Person: naufal
Actual: e | Predicted: e | CORRECT

Best Distance: 71.9031
All distances:
e: 71.9031 <- CHOSEN (CORRECT)
u: 76.3323
a: 79.4679
i: 83.3870
o: 100.6227

Test: naufal - o 4.wav
Person: naufal
Actual: o | Predicted: o | CORRECT
Best Distance: 64.2112
All distances:
o: 64.2112 <- CHOSEN (CORRECT)
u: 65.4305
a: 67.4120
e: 71.1539
i: 94.5248

Test: wiga - a 3.wav
Person: wiga
Actual: a | Predicted: a | CORRECT
Best Distance: 53.1276
All distances:
a: 53.1276 <- CHOSEN (CORRECT)
o: 70.0291
u: 71.0580
e: 84.9119
i: 104.1089

Test: wiga - i 3.wav
Person: wiga
Actual: i | Predicted: i | CORRECT
Best Distance: 80.4008
All distances:
i: 80.4008 <- CHOSEN (CORRECT)
u: 83.4983
a: 88.8513
e: 114.4101
o: 126.2325

Test: wiga - u 3.wav
Person: wiga
Actual: u | Predicted: u | CORRECT
Best Distance: 52.9289
All distances:
u: 52.9289 <- CHOSEN (CORRECT)
o: 71.6992
a: 89.0054
e: 105.4107
i: 108.4208

Test: wiga - e 3.wav
Person: wiga
Actual: e | Predicted: u | WRONG

Best Distance: 86.2586

All distances:

u: 86.2586 <- CHOSEN

i: 90.7028

a: 109.7578

e: 143.0313 (CORRECT)

o: 162.3077

Test: wiga - o 3.wav

Person: wiga

Actual: o | Predicted: u | WRONG

Best Distance: 63.5333

All distances:

u: 63.5333 <- CHOSEN

o: 94.7962 (CORRECT)

a: 95.8924

e: 123.4073

i: 126.2884

Test: zya - a 3.wav

Person: zya

Actual: a | Predicted: a | CORRECT

Best Distance: 52.8507

All distances:

a: 52.8507 <- CHOSEN (CORRECT)

u: 63.6856

o: 68.4934

e: 70.3881

i: 100.6245

Test: zya - i 3.wav

Person: zya

Actual: i | Predicted: i | CORRECT

Best Distance: 47.4606

All distances:

i: 47.4606 <- CHOSEN (CORRECT)

u: 66.4740

e: 66.7781

o: 94.1389

a: 100.8038

Test: zya - u 3.wav

Person: zya

Actual: u | Predicted: u | CORRECT

Best Distance: 49.1872

All distances:

u: 49.1872 <- CHOSEN (CORRECT)

o: 59.0017

a: 76.4671

e: 92.9694

i: 109.2422

Test: zya - e 3.wav

Person: zya

Actual: e | Predicted: e | CORRECT

Best Distance: 51.3478
All distances:
e: 51.3478 <- CHOSEN (CORRECT)
u: 64.8213
a: 73.6815
o: 75.4071
i: 90.7024

Test: zya - o 3.wav
Person: zya
Actual: o | Predicted: o | CORRECT
Best Distance: 61.4085
All distances:
o: 61.4085 <- CHOSEN (CORRECT)
a: 68.8325
u: 70.5674
e: 83.1020
i: 90.7356

CLOSED SCENARIO SUMMARY: 20/25 correct = 80.00%

[OPEN SCENARIO] (templates from wiga vs test from templates_other)
Test files: 50

--- DETAILED OPEN SCENARIO RESULTS ---
Using templates from: ['wiga']

Test: dab - a 2.wav
Person: Dabbir
Actual: a | Predicted: u | WRONG
Best Distance: 5845.0311
All distances:
u: 5845.0311 <- CHOSEN
e: 14395.1300
i: 304927.8081
a: 2801848.7628 (CORRECT)
o: 5545492.6510

Test: dab - i 2.wav
Person: Dabbir
Actual: i | Predicted: u | WRONG
Best Distance: 6474.7020
All distances:
u: 6474.7020 <- CHOSEN
e: 7402.9704
i: 458418.6542 (CORRECT)
a: 5498143.6685
o: 6478586.9961

Test: dab - u 2.wav
Person: Dabbir
Actual: u | Predicted: u | CORRECT
Best Distance: 4400.9275
All distances:
u: 4400.9275 <- CHOSEN (CORRECT)

e: 16005.0705
i: 597894.1006
o: 4627688.6353
a: 4754654.4901

Test: dab - e 2.wav

Person: Dabbir

Actual: e | Predicted: u | WRONG

Best Distance: 8660.1778

All distances:

u: 8660.1778 <- CHOSEN
e: 21641.2557 (CORRECT)
i: 673039.4756
a: 4173094.0320
o: 4934256.2167

Test: dab - o 2.wav

Person: Dabbir

Actual: o | Predicted: u | WRONG

Best Distance: 5439.8528

All distances:

u: 5439.8528 <- CHOSEN
e: 19704.0098
i: 432380.1523
a: 3183473.6397
o: 5659227.9751 (CORRECT)

Test: Evelyn - a 2.wav

Person: Evelyn

Actual: a | Predicted: u | WRONG

Best Distance: 5348.9733

All distances:

u: 5348.9733 <- CHOSEN
e: 44025.2812
i: 358979.4822
a: 5436245.3273 (CORRECT)
o: 6144957.1857

Test: Evelyn - i 2.wav

Person: Evelyn

Actual: i | Predicted: u | WRONG

Best Distance: 5815.9674

All distances:

u: 5815.9674 <- CHOSEN
e: 50602.1162
i: 343634.8584 (CORRECT)
a: 5043416.5619
o: 7125835.0880

Test: Evelyn - u 2.wav

Person: Evelyn

Actual: u | Predicted: u | CORRECT

Best Distance: 4626.3162

All distances:

u: 4626.3162 <- CHOSEN (CORRECT)

e: 19771.4625
i: 660575.2274
a: 3770985.2731
o: 4493818.9227

Test: Evelyn - e 2.wav

Person: Evelyn

Actual: e | Predicted: u | WRONG

Best Distance: 4940.9311

All distances:

u: 4940.9311 <- CHOSEN
e: 51264.8269 (CORRECT)
i: 417021.9751
a: 5477300.9000
o: 6804221.0626

Test: Evelyn - o 2.wav

Person: Evelyn

Actual: o | Predicted: u | WRONG

Best Distance: 5783.9857

All distances:

u: 5783.9857 <- CHOSEN
e: 15863.6181
i: 935049.5959
a: 4945213.5782
o: 6743522.7514 (CORRECT)

Test: Fed - a 2.wav

Person: Fed

Actual: a | Predicted: u | WRONG

Best Distance: 10273.7804

All distances:

u: 10273.7804 <- CHOSEN
e: 19208.8957
i: 875181.9054
a: 4980374.7420 (CORRECT)
o: 8844601.8774

Test: Fed - i 2.wav

Person: Fed

Actual: i | Predicted: u | WRONG

Best Distance: 4860.2246

All distances:

u: 4860.2246 <- CHOSEN
e: 15307.2325
i: 454323.0337 (CORRECT)
a: 3938165.7598
o: 6643294.3790

Test: Fed - u 2.wav

Person: Fed

Actual: u | Predicted: u | CORRECT

Best Distance: 6108.3934

All distances:

u: 6108.3934 <- CHOSEN (CORRECT)

e: 17609.3618
i: 643600.6364
a: 4063576.9472
o: 6312662.4751

Test: Fed - e 2.wav

Person: Fed

Actual: e | Predicted: u | WRONG

Best Distance: 9594.0093

All distances:

u: 9594.0093 <- CHOSEN
e: 38679.7086 (CORRECT)
i: 683840.5004
a: 5796943.7106
o: 6605666.2976

Test: Fed - o 2.wav

Person: Fed

Actual: o | Predicted: u | WRONG

Best Distance: 10867.3197

All distances:

u: 10867.3197 <- CHOSEN
e: 13002.0912
i: 558675.7087
a: 4698967.4672
o: 6273797.5144 (CORRECT)

Test: Akbar - a 2.wav

Person: Akbar

Actual: a | Predicted: u | WRONG

Best Distance: 12136.9322

All distances:

u: 12136.9322 <- CHOSEN
e: 20107.9889
i: 689541.4784
a: 4538306.0969 (CORRECT)
o: 6625660.7761

Test: Akbar - i 2.wav

Person: Akbar

Actual: i | Predicted: u | WRONG

Best Distance: 4654.1445

All distances:

u: 4654.1445 <- CHOSEN
e: 15565.3156
i: 520522.2103 (CORRECT)
a: 5507129.7758
o: 6632042.4607

Test: Akbar - u 2.wav

Person: Akbar

Actual: u | Predicted: u | CORRECT

Best Distance: 5168.0735

All distances:

u: 5168.0735 <- CHOSEN (CORRECT)

e: 16181.0204
i: 714617.1428
o: 3974242.5772
a: 4924327.3582

Test: Akbar - e 2.wav

Person: Akbar

Actual: e | Predicted: u | WRONG

Best Distance: 6786.3018

All distances:

u: 6786.3018 <- CHOSEN
e: 15356.6058 (CORRECT)
i: 373082.4308
a: 5252586.8514
o: 7459220.5242

Test: Akbar - o 2.wav

Person: Akbar

Actual: o | Predicted: u | WRONG

Best Distance: 6888.0479

All distances:

u: 6888.0479 <- CHOSEN
e: 19773.1103
i: 416611.2345
a: 4130394.6305
o: 6488409.2401 (CORRECT)

Test: Justin - a 2.wav

Person: Justin

Actual: a | Predicted: u | WRONG

Best Distance: 7638.9264

All distances:

u: 7638.9264 <- CHOSEN
e: 36802.1145
i: 737007.7995
a: 4914529.7648 (CORRECT)
o: 6771080.8901

Test: Justin - i 2.wav

Person: Justin

Actual: i | Predicted: u | WRONG

Best Distance: 6023.1488

All distances:

u: 6023.1488 <- CHOSEN
e: 30374.7741
i: 928977.7136 (CORRECT)
a: 5207354.9346
o: 7307917.1377

Test: Justin - u 2.wav

Person: Justin

Actual: u | Predicted: u | CORRECT

Best Distance: 5161.7084

All distances:

u: 5161.7084 <- CHOSEN (CORRECT)

e: 37361.6867
i: 558606.6912
a: 5750984.9179
o: 6801830.8135

Test: Justin - e 2.wav

Person: Justin

Actual: e | Predicted: u | WRONG

Best Distance: 7046.5056

All distances:

u: 7046.5056 <- CHOSEN
e: 33339.1982 (CORRECT)
i: 968823.7364
a: 5820290.8832
o: 7252282.9643

Test: Justin - o 2.wav

Person: Justin

Actual: o | Predicted: u | WRONG

Best Distance: 4845.9072

All distances:

u: 4845.9072 <- CHOSEN
e: 36239.6460
i: 320636.7093
a: 4707467.8544
o: 5213901.3813 (CORRECT)

Test: Ucup - a 2.wav

Person: Ucup

Actual: a | Predicted: u | WRONG

Best Distance: 5171.2070

All distances:

u: 5171.2070 <- CHOSEN
e: 44872.2923
i: 391850.4315
a: 4983690.9379 (CORRECT)
o: 6509626.1721

Test: Ucup - i 2.wav

Person: Ucup

Actual: i | Predicted: u | WRONG

Best Distance: 5661.8723

All distances:

u: 5661.8723 <- CHOSEN
e: 39050.7980
i: 338062.9820 (CORRECT)
a: 5021424.4247
o: 7700651.5996

Test: Ucup - u 2.wav

Person: Ucup

Actual: u | Predicted: u | CORRECT

Best Distance: 4007.2765

All distances:

u: 4007.2765 <- CHOSEN (CORRECT)

e: 37773.2405
i: 391460.1820
a: 5435118.6392
o: 6690477.7429

Test: Ucup - e 2.wav

Person: Ucup

Actual: e | Predicted: u | WRONG

Best Distance: 4970.4434

All distances:

u: 4970.4434 <- CHOSEN
e: 17320.3009 (CORRECT)
i: 168628.6197
a: 4224881.4475
o: 6356843.3688

Test: Ucup - o 2.wav

Person: Ucup

Actual: o | Predicted: u | WRONG

Best Distance: 5855.1585

All distances:

u: 5855.1585 <- CHOSEN
e: 21385.2735
i: 355667.1774
a: 4052004.4160
o: 5032576.2125 (CORRECT)

Test: gon - a 2.wav

Person: Gonza

Actual: a | Predicted: u | WRONG

Best Distance: 6595.4334

All distances:

u: 6595.4334 <- CHOSEN
e: 17018.3626
i: 300837.8929
a: 4626790.0045 (CORRECT)
o: 6484271.6271

Test: gon - i 2.wav

Person: Gonza

Actual: i | Predicted: u | WRONG

Best Distance: 6764.8704

All distances:

u: 6764.8704 <- CHOSEN
e: 17652.1490
i: 304785.4438 (CORRECT)
a: 4207657.8661
o: 6899451.1713

Test: gon - u 2.wav

Person: Gonza

Actual: u | Predicted: u | CORRECT

Best Distance: 3677.6003

All distances:

u: 3677.6003 <- CHOSEN (CORRECT)

e: 26426.7614
i: 365843.6766
a: 3769968.0292
o: 5451845.1964

Test: gon - e 2.wav

Person: Gonza

Actual: e | Predicted: u | WRONG

Best Distance: 5978.8107

All distances:

u: 5978.8107 <- CHOSEN
e: 24147.1694 (CORRECT)
i: 641386.8287
a: 5331015.8188
o: 8099157.5474

Test: gon - o 2.wav

Person: Gonza

Actual: o | Predicted: u | WRONG

Best Distance: 6009.5797

All distances:

u: 6009.5797 <- CHOSEN
e: 25346.7127
i: 517389.6526
a: 4187628.5246
o: 5106581.3133 (CORRECT)

Test: ji - a 2.wav

Person: Angie

Actual: a | Predicted: u | WRONG

Best Distance: 5477.6109

All distances:

u: 5477.6109 <- CHOSEN
e: 23068.5953
i: 954029.9740
a: 3824393.1876 (CORRECT)
o: 7730503.8946

Test: ji - i 2.wav

Person: Angie

Actual: i | Predicted: u | WRONG

Best Distance: 8194.5803

All distances:

u: 8194.5803 <- CHOSEN
e: 11543.1076
i: 417713.0494 (CORRECT)
a: 5441840.1785
o: 7251119.2523

Test: ji - u 2.wav

Person: Angie

Actual: u | Predicted: u | CORRECT

Best Distance: 6904.7193

All distances:

u: 6904.7193 <- CHOSEN (CORRECT)

e: 9905.5928
i: 504169.1477
a: 4162183.5396
o: 5303241.5994

Test: ji - e 2.wav

Person: Angie

Actual: e | Predicted: u | WRONG

Best Distance: 6623.9095

All distances:

u: 6623.9095 <- CHOSEN
e: 11122.5397 (CORRECT)
i: 450455.1225
a: 4182862.3457
o: 5995657.1711

Test: ji - o 2.wav

Person: Angie

Actual: o | Predicted: u | WRONG

Best Distance: 3924.6877

All distances:

u: 3924.6877 <- CHOSEN
e: 12939.4644
i: 537454.9965
a: 3317049.2716
o: 3543809.1189 (CORRECT)

Test: nan - a 2.wav

Person: Adnan

Actual: a | Predicted: u | WRONG

Best Distance: 17704.9383

All distances:

u: 17704.9383 <- CHOSEN
e: 20627.9570
i: 425404.7395
a: 4237653.8291 (CORRECT)
o: 6327598.5127

Test: nan - i 2.wav

Person: Adnan

Actual: i | Predicted: u | WRONG

Best Distance: 5843.0238

All distances:

u: 5843.0238 <- CHOSEN
e: 11619.8757
i: 428953.6668 (CORRECT)
o: 5842220.1557
a: 6223385.2177

Test: nan - u 2.wav

Person: Adnan

Actual: u | Predicted: u | CORRECT

Best Distance: 4849.0506

All distances:

u: 4849.0506 <- CHOSEN (CORRECT)

e: 17943.4078
i: 276180.8863
a: 3537520.0471
o: 5681301.6091

Test: nan - e 2.wav

Person: Adnan

Actual: e | Predicted: u | WRONG

Best Distance: 13239.3398

All distances:

u: 13239.3398 <- CHOSEN
e: 15985.4903 (CORRECT)
i: 307801.9380
a: 4711006.4625
o: 5433769.8694

Test: nan - o 2.wav

Person: Adnan

Actual: o | Predicted: u | WRONG

Best Distance: 4860.2607

All distances:

u: 4860.2607 <- CHOSEN
e: 17172.2145
i: 412796.4247
a: 4064791.3693
o: 6258176.8183 (CORRECT)

Test: zul - a 2.wav

Person: Azul

Actual: a | Predicted: u | WRONG

Best Distance: 5277.9655

All distances:

u: 5277.9655 <- CHOSEN
e: 22453.0075
i: 421347.9317
a: 2307748.2292 (CORRECT)
o: 4546535.6632

Test: zul - i 2.wav

Person: Azul

Actual: i | Predicted: u | WRONG

Best Distance: 6010.0295

All distances:

u: 6010.0295 <- CHOSEN
e: 11682.4959
i: 200526.7491 (CORRECT)
a: 4398458.2133
o: 5867192.4389

Test: zul - u 2.wav

Person: Azul

Actual: u | Predicted: u | CORRECT

Best Distance: 4542.8702

All distances:

u: 4542.8702 <- CHOSEN (CORRECT)

```
e: 14167.6131
i: 536620.7767
a: 3803337.6399
o: 4096433.5575
```

Test: zul - e 2.wav

Person: Azul

Actual: e | Predicted: u | WRONG

Best Distance: 4504.4768

All distances:

```
u: 4504.4768 <- CHOSEN
e: 15174.0335 (CORRECT)
i: 512730.2817
a: 4731455.1961
o: 5191737.7040
```

Test: zul - o 2.wav

Person: Azul

Actual: o | Predicted: u | WRONG

Best Distance: 4821.2591

All distances:

```
u: 4821.2591 <- CHOSEN
e: 14583.5018
i: 585037.9612
a: 3561830.1352
o: 4812228.8886 (CORRECT)
```

OPEN SCENARIO SUMMARY: 10/50 correct = 20.00%

--- SUMMARY FOR WIGA ---

Closed Accuracy: 80.00%

Open Accuracy: 20.00%

Average Accuracy: 50.00%

IV.1.2 Hasil Tes dengan *Template 'Zya'*

```
=====
=== EVALUATION WITH TEMPLATES FROM ZYA ===
=====
```

Total test files (templates_us): 25

Total test files (templates_other): 50

[CLOSED SCENARIO] (all templates vs test from templates_us)

--- DETAILED CLOSED SCENARIO RESULTS ---

Test: densu - a 3.wav

Person: densu

Actual: a | Predicted: a | CORRECT

Best Distance: 70.2069

All distances:

```
a: 70.2069 <- CHOSEN (CORRECT)
```

o: 74.3216
u: 76.2315
e: 96.8769
i: 101.2702

Test: densu - i 3.wav

Person: densu

Actual: i | Predicted: i | CORRECT

Best Distance: 53.9492

All distances:

i: 53.9492 <- CHOSEN (CORRECT)
u: 75.7062
e: 78.5462
a: 95.7307
o: 100.7484

Test: densu - u 3.wav

Person: densu

Actual: u | Predicted: u | CORRECT

Best Distance: 55.6178

All distances:

u: 55.6178 <- CHOSEN (CORRECT)
o: 72.8369
a: 75.8265
i: 92.3293
e: 105.2425

Test: densu - e 3.wav

Person: densu

Actual: e | Predicted: e | CORRECT

Best Distance: 30.6416

All distances:

e: 30.6416 <- CHOSEN (CORRECT)
i: 64.1177
u: 65.0695
a: 73.9451
o: 74.4279

Test: densu - o 3.wav

Person: densu

Actual: o | Predicted: a | WRONG

Best Distance: 54.9172

All distances:

a: 54.9172 <- CHOSEN
o: 61.5603 (CORRECT)
u: 65.1129
e: 101.6156
i: 108.8738

Test: hira - a 3.wav

Person: hira

Actual: a | Predicted: a | CORRECT

Best Distance: 63.2745

All distances:

a: 63.2745 <- CHOSEN (CORRECT)

u: 72.9438
o: 77.6447
i: 89.6612
e: 108.5596

Test: hira - i 3.wav

Person: hira

Actual: i | Predicted: u | WRONG

Best Distance: 80.0211

All distances:

u: 80.0211 <- CHOSEN
e: 81.3058
a: 89.3321
i: 90.1186 (CORRECT)
o: 94.0986

Test: hira - u 3.wav

Person: hira

Actual: u | Predicted: u | CORRECT

Best Distance: 52.2315

All distances:

u: 52.2315 <- CHOSEN (CORRECT)
o: 56.9199
a: 82.7891
i: 87.1686
e: 94.8692

Test: hira - e 3.wav

Person: hira

Actual: e | Predicted: e | CORRECT

Best Distance: 31.6649

All distances:

e: 31.6649 <- CHOSEN (CORRECT)
u: 67.2318
i: 67.2943
o: 82.4711
a: 93.1578

Test: hira - o 3.wav

Person: hira

Actual: o | Predicted: a | WRONG

Best Distance: 52.5049

All distances:

a: 52.5049 <- CHOSEN
o: 57.9618 (CORRECT)
u: 64.8192
e: 81.7117
i: 104.5562

Test: naufal - a 4.wav

Person: naufal

Actual: a | Predicted: a | CORRECT

Best Distance: 57.6231

All distances:

a: 57.6231 <- CHOSEN (CORRECT)

u: 72.1626
e: 78.4054
o: 83.8841
i: 89.6910

Test: naufal - i 4.wav

Person: naufal

Actual: i | Predicted: i | CORRECT

Best Distance: 59.4456

All distances:

i: 59.4456 <- CHOSEN (CORRECT)
e: 75.9879
u: 78.6571
a: 93.3244
o: 142.4249

Test: naufal - u 4.wav

Person: naufal

Actual: u | Predicted: u | CORRECT

Best Distance: 52.9969

All distances:

u: 52.9969 <- CHOSEN (CORRECT)
o: 68.3638
e: 83.6060
a: 83.9304
i: 103.4871

Test: naufal - e 4.wav

Person: naufal

Actual: e | Predicted: e | CORRECT

Best Distance: 71.9031

All distances:

e: 71.9031 <- CHOSEN (CORRECT)
u: 76.3323
a: 79.4679
i: 83.3870
o: 100.6227

Test: naufal - o 4.wav

Person: naufal

Actual: o | Predicted: o | CORRECT

Best Distance: 64.2112

All distances:

o: 64.2112 <- CHOSEN (CORRECT)
u: 65.4305
a: 67.4120
e: 71.1539
i: 94.5248

Test: wiga - a 3.wav

Person: wiga

Actual: a | Predicted: a | CORRECT

Best Distance: 53.1276

All distances:

a: 53.1276 <- CHOSEN (CORRECT)

o: 70.0291
u: 71.0580
e: 84.9119
i: 104.1089

Test: wiga - i 3.wav

Person: wiga

Actual: i | Predicted: i | CORRECT

Best Distance: 80.4008

All distances:

i: 80.4008 <- CHOSEN (CORRECT)
u: 83.4983
a: 88.8513
e: 114.4101
o: 126.2325

Test: wiga - u 3.wav

Person: wiga

Actual: u | Predicted: u | CORRECT

Best Distance: 52.9289

All distances:

u: 52.9289 <- CHOSEN (CORRECT)
o: 71.6992
a: 89.0054
e: 105.4107
i: 108.4208

Test: wiga - e 3.wav

Person: wiga

Actual: e | Predicted: u | WRONG

Best Distance: 86.2586

All distances:

u: 86.2586 <- CHOSEN
i: 90.7028
a: 109.7578
e: 143.0313 (CORRECT)
o: 162.3077

Test: wiga - o 3.wav

Person: wiga

Actual: o | Predicted: u | WRONG

Best Distance: 63.5333

All distances:

u: 63.5333 <- CHOSEN
o: 94.7962 (CORRECT)
a: 95.8924
e: 123.4073
i: 126.2884

Test: zya - a 3.wav

Person: zya

Actual: a | Predicted: a | CORRECT

Best Distance: 52.8507

All distances:

a: 52.8507 <- CHOSEN (CORRECT)

u: 63.6856
o: 68.4934
e: 70.3881
i: 100.6245

Test: zya - i 3.wav

Person: zya

Actual: i | Predicted: i | CORRECT

Best Distance: 47.4606

All distances:

i: 47.4606 <- CHOSEN (CORRECT)

u: 66.4740

e: 66.7781

o: 94.1389

a: 100.8038

Test: zya - u 3.wav

Person: zya

Actual: u | Predicted: u | CORRECT

Best Distance: 49.1872

All distances:

u: 49.1872 <- CHOSEN (CORRECT)

o: 59.0017

a: 76.4671

e: 92.9694

i: 109.2422

Test: zya - e 3.wav

Person: zya

Actual: e | Predicted: e | CORRECT

Best Distance: 51.3478

All distances:

e: 51.3478 <- CHOSEN (CORRECT)

u: 64.8213

a: 73.6815

o: 75.4071

i: 90.7024

Test: zya - o 3.wav

Person: zya

Actual: o | Predicted: o | CORRECT

Best Distance: 61.4085

All distances:

o: 61.4085 <- CHOSEN (CORRECT)

a: 68.8325

u: 70.5674

e: 83.1020

i: 90.7356

CLOSED SCENARIO SUMMARY: 20/25 correct = 80.00%

[OPEN SCENARIO] (templates from zya vs test from templates_other)

Test files: 50

=== BUILDING GENERALIZED TEMPLATES ===

```
Building generalized template for vowel: a
- Created 3 segments with mean and covariance parameters
Building generalized template for vowel: i
- Created 3 segments with mean and covariance parameters
Building generalized template for vowel: u
- Created 3 segments with mean and covariance parameters
Building generalized template for vowel: e
- Created 3 segments with mean and covariance parameters
Building generalized template for vowel: o
- Created 3 segments with mean and covariance parameters
Generalized templates built successfully!
```

--- DETAILED OPEN SCENARIO RESULTS ---

Using templates from: ['zya']

```
Test: dab - a 2.wav
Person: Dabbir
Actual: a | Predicted: u | WRONG
Best Distance: 6386.0748
All distances:
  u: 6386.0748 <- CHOSEN
  e: 2811935.7275
  o: 5123224.1840
  a: 5170680.7771 (CORRECT)
  i: 5276273.4929
```

```
Test: dab - i 2.wav
Person: Dabbir
Actual: i | Predicted: u | WRONG
Best Distance: 6897.0206
All distances:
  u: 6897.0206 <- CHOSEN
  e: 3165669.4996
  i: 3235178.3217 (CORRECT)
  o: 4343854.6153
  a: 7980968.4360
```

```
Test: dab - u 2.wav
Person: Dabbir
Actual: u | Predicted: u | CORRECT
Best Distance: 4354.8990
All distances:
  u: 4354.8990 <- CHOSEN (CORRECT)
  e: 3334724.3736
  o: 3754669.7834
  i: 4868166.5149
  a: 6700984.7466
```

```
Test: dab - e 2.wav
Person: Dabbir
Actual: e | Predicted: u | WRONG
Best Distance: 5345.9550
All distances:
  u: 5345.9550 <- CHOSEN
  e: 3047411.4381 (CORRECT)
```

i: 3700305.9100
o: 4482259.9157
a: 7057844.7926

Test: dab - o 2.wav
Person: Dabbir
Actual: o | Predicted: u | WRONG
Best Distance: 4091.3763
All distances:
u: 4091.3763 <- CHOSEN
e: 2751312.8848
o: 3858223.1989 (CORRECT)
i: 4458382.2617
a: 6613537.6291

Test: Evelyn - a 2.wav
Person: Evelyn
Actual: a | Predicted: u | WRONG
Best Distance: 6940.5461
All distances:
u: 6940.5461 <- CHOSEN
e: 4546660.4794
i: 5012556.9075
o: 5319725.3652
a: 8178283.7390 (CORRECT)

Test: Evelyn - i 2.wav
Person: Evelyn
Actual: i | Predicted: u | WRONG
Best Distance: 9889.3004
All distances:
u: 9889.3004 <- CHOSEN
i: 4494704.6267 (CORRECT)
e: 4575343.2480
o: 6479115.1990
a: 7491367.8169

Test: Evelyn - u 2.wav
Person: Evelyn
Actual: u | Predicted: u | CORRECT
Best Distance: 6070.1617
All distances:
u: 6070.1617 <- CHOSEN (CORRECT)
e: 4598588.7617
o: 5181526.3601
i: 5712483.4198
a: 6273824.7689

Test: Evelyn - e 2.wav
Person: Evelyn
Actual: e | Predicted: u | WRONG
Best Distance: 8711.8767
All distances:
u: 8711.8767 <- CHOSEN
e: 4388118.7332 (CORRECT)

i: 4909993.1285
o: 5970301.1626
a: 7012168.4665

Test: Evelyn - o 2.wav

Person: Evelyn

Actual: o | Predicted: u | WRONG

Best Distance: 4581.8637

All distances:

u: 4581.8637 <- CHOSEN
e: 3516019.8440
o: 4644920.6570 (CORRECT)
a: 6265399.5936
i: 6942469.3458

Test: Fed - a 2.wav

Person: Fed

Actual: a | Predicted: u | WRONG

Best Distance: 8178.3233

All distances:

u: 8178.3233 <- CHOSEN
e: 3530952.6240
i: 6298748.4335
o: 6905044.5122
a: 7411492.1340 (CORRECT)

Test: Fed - i 2.wav

Person: Fed

Actual: i | Predicted: u | WRONG

Best Distance: 10198.6031

All distances:

u: 10198.6031 <- CHOSEN
e: 3536006.9593
i: 3813518.6494 (CORRECT)
o: 6563361.2343
a: 9460762.1992

Test: Fed - u 2.wav

Person: Fed

Actual: u | Predicted: u | CORRECT

Best Distance: 4601.5425

All distances:

u: 4601.5425 <- CHOSEN (CORRECT)
e: 3422808.3688
o: 5359706.4197
i: 5704308.4815
a: 7294633.6857

Test: Fed - e 2.wav

Person: Fed

Actual: e | Predicted: u | WRONG

Best Distance: 7009.5128

All distances:

u: 7009.5128 <- CHOSEN
e: 4388599.3727 (CORRECT)

i: 5544403.9633
o: 7897157.8281
a: 8953073.5841

Test: Fed - o 2.wav

Person: Fed

Actual: o | Predicted: u | WRONG

Best Distance: 5920.7581

All distances:

u: 5920.7581 <- CHOSEN
e: 2780802.7661
i: 4932578.3310
a: 5385231.4564
o: 5738784.6793 (CORRECT)

Test: Akbar - a 2.wav

Person: Akbar

Actual: a | Predicted: u | WRONG

Best Distance: 7419.5939

All distances:

u: 7419.5939 <- CHOSEN
e: 4215159.4317
i: 5582803.4479
o: 5719033.8986
a: 6897892.7585 (CORRECT)

Test: Akbar - i 2.wav

Person: Akbar

Actual: i | Predicted: u | WRONG

Best Distance: 7450.5508

All distances:

u: 7450.5508 <- CHOSEN
i: 2668640.4694 (CORRECT)
e: 4025483.4518
o: 6662827.9205
a: 8290226.1270

Test: Akbar - u 2.wav

Person: Akbar

Actual: u | Predicted: u | CORRECT

Best Distance: 4026.9461

All distances:

u: 4026.9461 <- CHOSEN (CORRECT)
e: 3383738.5318
o: 3597027.8940
i: 4292733.2101
a: 7843673.6313

Test: Akbar - e 2.wav

Person: Akbar

Actual: e | Predicted: u | WRONG

Best Distance: 8527.3903

All distances:

u: 8527.3903 <- CHOSEN
e: 3870473.5735 (CORRECT)

i: 4388811.3741
o: 6650322.1890
a: 9533235.1377

Test: Akbar - o 2.wav

Person: Akbar

Actual: o | Predicted: u | WRONG

Best Distance: 9016.4163

All distances:

u: 9016.4163 <- CHOSEN
e: 3621061.9962
i: 4498470.7895
o: 5341086.6094 (CORRECT)
a: 6881421.7740

Test: Justin - a 2.wav

Person: Justin

Actual: a | Predicted: u | WRONG

Best Distance: 8358.8571

All distances:

u: 8358.8571 <- CHOSEN
e: 4548270.4911
i: 5789144.9857
o: 5800536.3107
a: 7343600.1675 (CORRECT)

Test: Justin - i 2.wav

Person: Justin

Actual: i | Predicted: u | WRONG

Best Distance: 7423.2138

All distances:

u: 7423.2138 <- CHOSEN
e: 5043306.8309
i: 5355597.5941 (CORRECT)
o: 6202153.6655
a: 7535551.9631

Test: Justin - u 2.wav

Person: Justin

Actual: u | Predicted: u | CORRECT

Best Distance: 8333.4284

All distances:

u: 8333.4284 <- CHOSEN (CORRECT)
i: 5433845.0001
e: 5563738.5945
o: 6697112.6597
a: 8451431.7834

Test: Justin - e 2.wav

Person: Justin

Actual: e | Predicted: u | WRONG

Best Distance: 8494.1870

All distances:

u: 8494.1870 <- CHOSEN
e: 5241535.2070 (CORRECT)

i: 5306671.2152
o: 7240300.3559
a: 8233423.5924

Test: Justin - o 2.wav

Person: Justin

Actual: o | Predicted: u | WRONG

Best Distance: 8032.2903

All distances:

u: 8032.2903 <- CHOSEN
o: 4924707.6849 (CORRECT)
e: 4994684.0485
i: 5660642.0642
a: 8010825.2385

Test: Ucup - a 2.wav

Person: Ucup

Actual: a | Predicted: u | WRONG

Best Distance: 11099.5919

All distances:

u: 11099.5919 <- CHOSEN
e: 4453001.3277
i: 5242351.0169
o: 5559220.5476
a: 8074955.0352 (CORRECT)

Test: Ucup - i 2.wav

Person: Ucup

Actual: i | Predicted: u | WRONG

Best Distance: 9776.0822

All distances:

u: 9776.0822 <- CHOSEN
e: 4603294.9977
i: 4952134.9225 (CORRECT)
o: 5832777.4878
a: 7951486.2606

Test: Ucup - u 2.wav

Person: Ucup

Actual: u | Predicted: u | CORRECT

Best Distance: 8055.3521

All distances:

u: 8055.3521 <- CHOSEN (CORRECT)
e: 4247195.0001
o: 4994635.0653
i: 5542766.8810
a: 7703472.2966

Test: Ucup - e 2.wav

Person: Ucup

Actual: e | Predicted: u | WRONG

Best Distance: 4363.6721

All distances:

u: 4363.6721 <- CHOSEN
e: 3934408.9486 (CORRECT)

i: 4197842.9347
o: 6410773.4283
a: 8624490.7169

Test: Ucup - o 2.wav

Person: Ucup

Actual: o | Predicted: u | WRONG

Best Distance: 4831.4468

All distances:

u: 4831.4468 <- CHOSEN
e: 3304576.1635
o: 4717172.9434 (CORRECT)
i: 4811175.8353
a: 6327146.6271

Test: gon - a 2.wav

Person: Gonza

Actual: a | Predicted: u | WRONG

Best Distance: 9068.0404

All distances:

u: 9068.0404 <- CHOSEN
e: 2592307.3295
i: 4771707.7610
o: 5482608.4265
a: 7757977.8840 (CORRECT)

Test: gon - i 2.wav

Person: Gonza

Actual: i | Predicted: u | WRONG

Best Distance: 6957.8628

All distances:

u: 6957.8628 <- CHOSEN
e: 3285606.1459
i: 3579254.3727 (CORRECT)
o: 5926768.7482
a: 8887140.4616

Test: gon - u 2.wav

Person: Gonza

Actual: u | Predicted: u | CORRECT

Best Distance: 6769.2620

All distances:

u: 6769.2620 <- CHOSEN (CORRECT)
e: 3377272.6240
i: 5318825.5415
o: 5799868.0273
a: 7158947.4646

Test: gon - e 2.wav

Person: Gonza

Actual: e | Predicted: u | WRONG

Best Distance: 3445.2161

All distances:

u: 3445.2161 <- CHOSEN
e: 2796460.6870 (CORRECT)

i: 4109809.4251
o: 6511997.7604
a: 8234717.5202

Test: gon - o 2.wav
Person: Gonza
Actual: o | Predicted: u | WRONG
Best Distance: 5754.0207
All distances:
u: 5754.0207 <- CHOSEN
i: 3086427.3417
e: 3095837.6769
o: 3728817.0223 (CORRECT)
a: 6442360.7575

Test: ji - a 2.wav
Person: Angie
Actual: a | Predicted: u | WRONG
Best Distance: 6618.9971
All distances:
u: 6618.9971 <- CHOSEN
e: 3871058.0416
i: 5086836.2962
o: 5513525.5924
a: 6905838.1703 (CORRECT)

Test: ji - i 2.wav
Person: Angie
Actual: i | Predicted: u | WRONG
Best Distance: 6264.8886
All distances:
u: 6264.8886 <- CHOSEN
e: 3265268.4116
i: 3728647.1807 (CORRECT)
o: 4773582.9771
a: 8369670.8453

Test: ji - u 2.wav
Person: Angie
Actual: u | Predicted: u | CORRECT
Best Distance: 4274.3843
All distances:
u: 4274.3843 <- CHOSEN (CORRECT)
e: 3824818.7068
o: 4934090.4229
i: 5035425.8085
a: 6445864.3168

Test: ji - e 2.wav
Person: Angie
Actual: e | Predicted: u | WRONG
Best Distance: 8577.2847
All distances:
u: 8577.2847 <- CHOSEN
e: 3413651.1733 (CORRECT)

i: 4968431.2602
o: 5631107.4473
a: 7626540.1068

Test: ji - o 2.wav
Person: Angie
Actual: o | Predicted: u | WRONG
Best Distance: 4522.7207
All distances:
u: 4522.7207 <- CHOSEN
i: 4016519.4446
e: 4070230.1403
o: 4771996.6299 (CORRECT)
a: 6390105.1500

Test: nan - a 2.wav
Person: Adnan
Actual: a | Predicted: u | WRONG
Best Distance: 5880.4178
All distances:
u: 5880.4178 <- CHOSEN
e: 3300447.7511
a: 4647266.7623 (CORRECT)
i: 4958935.5970
o: 6419183.1784

Test: nan - i 2.wav
Person: Adnan
Actual: i | Predicted: u | WRONG
Best Distance: 4290.2728
All distances:
u: 4290.2728 <- CHOSEN
i: 3545713.1173 (CORRECT)
e: 3739042.7089
o: 5015811.1263
a: 8391422.2541

Test: nan - u 2.wav
Person: Adnan
Actual: u | Predicted: u | CORRECT
Best Distance: 6140.4450
All distances:
u: 6140.4450 <- CHOSEN (CORRECT)
i: 3215603.5322
e: 3261874.9215
o: 5037807.4255
a: 7432175.6766

Test: nan - e 2.wav
Person: Adnan
Actual: e | Predicted: u | WRONG
Best Distance: 6789.5013
All distances:
u: 6789.5013 <- CHOSEN
e: 1826676.2471 (CORRECT)

i: 3831666.9774
a: 6198963.2942
o: 6998482.2095

Test: nan - o 2.wav

Person: Adnan

Actual: o | Predicted: u | WRONG

Best Distance: 4645.9725

All distances:

u: 4645.9725 <- CHOSEN
e: 3699631.7510
i: 4286210.1420
o: 5277022.6596 (CORRECT)
a: 6936995.8215

Test: zul - a 2.wav

Person: Azul

Actual: a | Predicted: u | WRONG

Best Distance: 5572.5358

All distances:

u: 5572.5358 <- CHOSEN
e: 2477196.4334
o: 4221152.3261
i: 4879002.8822
a: 5159003.0715 (CORRECT)

Test: zul - i 2.wav

Person: Azul

Actual: i | Predicted: u | WRONG

Best Distance: 7552.6595

All distances:

u: 7552.6595 <- CHOSEN
e: 4844712.6646
i: 4909051.0654 (CORRECT)
o: 5002022.2106
a: 7719548.1182

Test: zul - u 2.wav

Person: Azul

Actual: u | Predicted: u | CORRECT

Best Distance: 3897.0143

All distances:

u: 3897.0143 <- CHOSEN (CORRECT)
e: 3626945.9484
o: 3872691.7596
i: 5390943.7801
a: 6617539.0658

Test: zul - e 2.wav

Person: Azul

Actual: e | Predicted: u | WRONG

Best Distance: 7354.8331

All distances:

u: 7354.8331 <- CHOSEN
e: 2663103.5327 (CORRECT)

```
o: 4538008.0532
i: 4669181.3025
a: 7161026.1139
```

Test: zul - o 2.wav

Person: Azul

Actual: o | Predicted: u | WRONG

Best Distance: 3611.3851

All distances:

u: 3611.3851 <- CHOSEN

e: 3152557.8625

o: 4157594.8123 (CORRECT)

i: 6361357.4032

a: 6407010.5759

OPEN SCENARIO SUMMARY: 10/50 correct = 20.00%

--- SUMMARY FOR ZYA ---

Closed Accuracy: 80.00%

Open Accuracy: 20.00%

Average Accuracy: 50.00%

IV.1.3 Hasil Tes dengan *Template* 'Hira'

```
=====
=== EVALUATION WITH TEMPLATES FROM HIRA ===
=====
```

Total test files (templates_us): 25

Total test files (templates_other): 50

[CLOSED SCENARIO] (all templates vs test from templates_us)

--- DETAILED CLOSED SCENARIO RESULTS ---

Test: densu - a 3.wav

Person: densu

Actual: a | Predicted: a | CORRECT

Best Distance: 70.2069

All distances:

a: 70.2069 <- CHOSEN (CORRECT)

o: 74.3216

u: 76.2315

e: 96.8769

i: 101.2702

Test: densu - i 3.wav

Person: densu

Actual: i | Predicted: i | CORRECT

Best Distance: 53.9492

All distances:

i: 53.9492 <- CHOSEN (CORRECT)

u: 75.7062

e: 78.5462
a: 95.7307
o: 100.7484

Test: densu - u 3.wav

Person: densu

Actual: u | Predicted: u | CORRECT

Best Distance: 55.6178

All distances:

u: 55.6178 <- CHOSEN (CORRECT)
o: 72.8369
a: 75.8265
i: 92.3293
e: 105.2425

Test: densu - e 3.wav

Person: densu

Actual: e | Predicted: e | CORRECT

Best Distance: 30.6416

All distances:

e: 30.6416 <- CHOSEN (CORRECT)
i: 64.1177
u: 65.0695
a: 73.9451
o: 74.4279

Test: densu - o 3.wav

Person: densu

Actual: o | Predicted: a | WRONG

Best Distance: 54.9172

All distances:

a: 54.9172 <- CHOSEN
o: 61.5603 (CORRECT)
u: 65.1129
e: 101.6156
i: 108.8738

Test: hira - a 3.wav

Person: hira

Actual: a | Predicted: a | CORRECT

Best Distance: 63.2745

All distances:

a: 63.2745 <- CHOSEN (CORRECT)
u: 72.9438
o: 77.6447
i: 89.6612
e: 108.5596

Test: hira - i 3.wav

Person: hira

Actual: i | Predicted: u | WRONG

Best Distance: 80.0211

All distances:

u: 80.0211 <- CHOSEN
e: 81.3058

a: 89.3321
i: 90.1186 (CORRECT)
o: 94.0986

Test: hira - u 3.wav

Person: hira

Actual: u | Predicted: u | CORRECT

Best Distance: 52.2315

All distances:

u: 52.2315 <- CHOSEN (CORRECT)

o: 56.9199

a: 82.7891

i: 87.1686

e: 94.8692

Test: hira - e 3.wav

Person: hira

Actual: e | Predicted: e | CORRECT

Best Distance: 31.6649

All distances:

e: 31.6649 <- CHOSEN (CORRECT)

u: 67.2318

i: 67.2943

o: 82.4711

a: 93.1578

Test: hira - o 3.wav

Person: hira

Actual: o | Predicted: a | WRONG

Best Distance: 52.5049

All distances:

a: 52.5049 <- CHOSEN

o: 57.9618 (CORRECT)

u: 64.8192

e: 81.7117

i: 104.5562

Test: naufal - a 4.wav

Person: naufal

Actual: a | Predicted: a | CORRECT

Best Distance: 57.6231

All distances:

a: 57.6231 <- CHOSEN (CORRECT)

u: 72.1626

e: 78.4054

o: 83.8841

i: 89.6910

Test: naufal - i 4.wav

Person: naufal

Actual: i | Predicted: i | CORRECT

Best Distance: 59.4456

All distances:

i: 59.4456 <- CHOSEN (CORRECT)

e: 75.9879

u: 78.6571
a: 93.3244
o: 142.4249

Test: naufal - u 4.wav
Person: naufal
Actual: u | Predicted: u | CORRECT
Best Distance: 52.9969
All distances:
u: 52.9969 <- CHOSEN (CORRECT)
o: 68.3638
e: 83.6060
a: 83.9304
i: 103.4871

Test: naufal - e 4.wav
Person: naufal
Actual: e | Predicted: e | CORRECT
Best Distance: 71.9031
All distances:
e: 71.9031 <- CHOSEN (CORRECT)
u: 76.3323
a: 79.4679
i: 83.3870
o: 100.6227

Test: naufal - o 4.wav
Person: naufal
Actual: o | Predicted: o | CORRECT
Best Distance: 64.2112
All distances:
o: 64.2112 <- CHOSEN (CORRECT)
u: 65.4305
a: 67.4120
e: 71.1539
i: 94.5248

Test: wiga - a 3.wav
Person: wiga
Actual: a | Predicted: a | CORRECT
Best Distance: 53.1276
All distances:
a: 53.1276 <- CHOSEN (CORRECT)
o: 70.0291
u: 71.0580
e: 84.9119
i: 104.1089

Test: wiga - i 3.wav
Person: wiga
Actual: i | Predicted: i | CORRECT
Best Distance: 80.4008
All distances:
i: 80.4008 <- CHOSEN (CORRECT)
u: 83.4983

a: 88.8513
e: 114.4101
o: 126.2325

Test: wiga - u 3.wav

Person: wiga

Actual: u | Predicted: u | CORRECT

Best Distance: 52.9289

All distances:

u: 52.9289 <- CHOSEN (CORRECT)
o: 71.6992
a: 89.0054
e: 105.4107
i: 108.4208

Test: wiga - e 3.wav

Person: wiga

Actual: e | Predicted: u | WRONG

Best Distance: 86.2586

All distances:

u: 86.2586 <- CHOSEN
i: 90.7028
a: 109.7578
e: 143.0313 (CORRECT)
o: 162.3077

Test: wiga - o 3.wav

Person: wiga

Actual: o | Predicted: u | WRONG

Best Distance: 63.5333

All distances:

u: 63.5333 <- CHOSEN
o: 94.7962 (CORRECT)
a: 95.8924
e: 123.4073
i: 126.2884

Test: zya - a 3.wav

Person: zya

Actual: a | Predicted: a | CORRECT

Best Distance: 52.8507

All distances:

a: 52.8507 <- CHOSEN (CORRECT)
u: 63.6856
o: 68.4934
e: 70.3881
i: 100.6245

Test: zya - i 3.wav

Person: zya

Actual: i | Predicted: i | CORRECT

Best Distance: 47.4606

All distances:

i: 47.4606 <- CHOSEN (CORRECT)
u: 66.4740

e: 66.7781
o: 94.1389
a: 100.8038

Test: zya - u 3.wav

Person: zya

Actual: u | Predicted: u | CORRECT

Best Distance: 49.1872

All distances:

u: 49.1872 <- CHOSEN (CORRECT)
o: 59.0017
a: 76.4671
e: 92.9694
i: 109.2422

Test: zya - e 3.wav

Person: zya

Actual: e | Predicted: e | CORRECT

Best Distance: 51.3478

All distances:

e: 51.3478 <- CHOSEN (CORRECT)
u: 64.8213
a: 73.6815
o: 75.4071
i: 90.7024

Test: zya - o 3.wav

Person: zya

Actual: o | Predicted: o | CORRECT

Best Distance: 61.4085

All distances:

o: 61.4085 <- CHOSEN (CORRECT)
a: 68.8325
u: 70.5674
e: 83.1020
i: 90.7356

CLOSED SCENARIO SUMMARY: 20/25 correct = 80.00%

[OPEN SCENARIO] (templates from hira vs test from templates_other)

Test files: 50

--- DETAILED OPEN SCENARIO RESULTS ---

Using templates from: ['hira']

Test: dab - a 2.wav

Person: Dabbir

Actual: a | Predicted: o | WRONG

Best Distance: 4381.4580

All distances:

o: 4381.4580 <- CHOSEN
i: 26438.2715
u: 2072871.4466
a: 3532213.8097 (CORRECT)
e: 9975533.1885

Test: dab - i 2.wav
 Person: Dabbir
 Actual: i | Predicted: o | WRONG
 Best Distance: 8262.6151
 All distances:
 o: 8262.6151 <- CHOSEN
 i: 22523.2953 (CORRECT)
 u: 3009667.8738
 a: 3815019.6901
 e: 13083034.1418

Test: dab - u 2.wav
 Person: Dabbir
 Actual: u | Predicted: o | WRONG
 Best Distance: 4541.5047
 All distances:
 o: 4541.5047 <- CHOSEN
 i: 29808.6303
 u: 2094882.6914 (CORRECT)
 a: 4661578.6920
 e: 11646592.0897

Test: dab - e 2.wav
 Person: Dabbir
 Actual: e | Predicted: o | WRONG
 Best Distance: 8296.3395
 All distances:
 o: 8296.3395 <- CHOSEN
 i: 22505.8442
 u: 2711877.8807
 a: 4454471.8150
 e: 9313925.0522 (CORRECT)

Test: dab - o 2.wav
 Person: Dabbir
 Actual: o | Predicted: o | CORRECT
 Best Distance: 4901.1191
 All distances:
 o: 4901.1191 <- CHOSEN (CORRECT)
 i: 29436.2846
 u: 1594684.5834
 a: 3682725.4985
 e: 12076824.6462

Test: Evelyn - a 2.wav
 Person: Evelyn
 Actual: a | Predicted: o | WRONG
 Best Distance: 7954.9740
 All distances:
 o: 7954.9740 <- CHOSEN
 i: 76742.9549
 u: 2911662.2946
 a: 4525311.3714 (CORRECT)
 e: 13395007.0423

Test: Evelyn - i 2.wav
Person: Evelyn
Actual: i | Predicted: o | WRONG
Best Distance: 13988.6177
All distances:
o: 13988.6177 <- CHOSEN
i: 51891.2060 (CORRECT)
u: 2814708.3707
a: 4787126.9840
e: 13882058.7708

Test: Evelyn - u 2.wav
Person: Evelyn
Actual: u | Predicted: o | WRONG
Best Distance: 7708.1568
All distances:
o: 7708.1568 <- CHOSEN
i: 33742.2091
u: 2075988.8314 (CORRECT)
a: 3510264.7515
e: 11066783.3768

Test: Evelyn - e 2.wav
Person: Evelyn
Actual: e | Predicted: o | WRONG
Best Distance: 9418.3289
All distances:
o: 9418.3289 <- CHOSEN
i: 67236.9028
u: 2902629.7597
a: 5077370.8382
e: 11155639.4063 (CORRECT)

Test: Evelyn - o 2.wav
Person: Evelyn
Actual: o | Predicted: o | CORRECT
Best Distance: 3464.5578
All distances:
o: 3464.5578 <- CHOSEN (CORRECT)
i: 30322.9702
u: 2088659.3033
a: 4506319.5637
e: 12570257.1828

Test: Fed - a 2.wav
Person: Fed
Actual: a | Predicted: o | WRONG
Best Distance: 5678.4287
All distances:
o: 5678.4287 <- CHOSEN
i: 33861.3356
u: 2107980.2847
a: 2748088.1762 (CORRECT)
e: 12674309.7325

Test: Fed - i 2.wav
Person: Fed
Actual: i | Predicted: o | WRONG
Best Distance: 6775.5674
All distances:
o: 6775.5674 <- CHOSEN
i: 18503.5496 (CORRECT)
u: 2963226.6020
a: 4912221.4368
e: 13075440.4453

Test: Fed - u 2.wav
Person: Fed
Actual: u | Predicted: o | WRONG
Best Distance: 4822.6228
All distances:
o: 4822.6228 <- CHOSEN
i: 30600.1379
u: 2132069.0115 (CORRECT)
a: 6043057.5969
e: 11412074.9365

Test: Fed - e 2.wav
Person: Fed
Actual: e | Predicted: o | WRONG
Best Distance: 7202.4750
All distances:
o: 7202.4750 <- CHOSEN
i: 37131.4039
u: 2977155.0037
a: 4537893.0843
e: 12164991.3342 (CORRECT)

Test: Fed - o 2.wav
Person: Fed
Actual: o | Predicted: o | CORRECT
Best Distance: 6379.0897
All distances:
o: 6379.0897 <- CHOSEN (CORRECT)
i: 31890.3892
u: 2268416.1849
a: 3671862.0368
e: 12632509.5640

Test: Akbar - a 2.wav
Person: Akbar
Actual: a | Predicted: o | WRONG
Best Distance: 7311.8427
All distances:
o: 7311.8427 <- CHOSEN
i: 53238.2485
u: 3351743.6438
a: 5143550.0840 (CORRECT)
e: 12389975.6772

Test: Akbar - i 2.wav
 Person: Akbar
 Actual: i | Predicted: o | WRONG
 Best Distance: 10499.6171
 All distances:
 o: 10499.6171 <- CHOSEN
 i: 20288.5579 (CORRECT)
 u: 2974778.9219
 a: 3472363.6012
 e: 11280587.6678

Test: Akbar - u 2.wav
 Person: Akbar
 Actual: u | Predicted: o | WRONG
 Best Distance: 7357.9547
 All distances:
 o: 7357.9547 <- CHOSEN
 i: 26144.2038
 u: 1877813.6656 (CORRECT)
 a: 5952633.5744
 e: 11954460.6194

Test: Akbar - e 2.wav
 Person: Akbar
 Actual: e | Predicted: o | WRONG
 Best Distance: 9439.7807
 All distances:
 o: 9439.7807 <- CHOSEN
 i: 51245.3558
 u: 3299596.3433
 a: 3570638.6316
 e: 11005068.5873 (CORRECT)

Test: Akbar - o 2.wav
 Person: Akbar
 Actual: o | Predicted: o | CORRECT
 Best Distance: 5236.7158
 All distances:
 o: 5236.7158 <- CHOSEN (CORRECT)
 i: 30266.3369
 u: 3191366.8707
 a: 3321838.8816
 e: 13076053.2936

Test: Justin - a 2.wav
 Person: Justin
 Actual: a | Predicted: o | WRONG
 Best Distance: 6805.1263
 All distances:
 o: 6805.1263 <- CHOSEN
 i: 58942.0653
 u: 2827446.6272
 a: 5099389.7252 (CORRECT)
 e: 12524141.8705

Test: Justin - i 2.wav
 Person: Justin
 Actual: i | Predicted: o | WRONG
 Best Distance: 8329.1539
 All distances:
 o: 8329.1539 <- CHOSEN
 i: 59500.2004 (CORRECT)
 u: 2865021.8019
 a: 4621024.6234
 e: 12362514.6350

Test: Justin - u 2.wav
 Person: Justin
 Actual: u | Predicted: o | WRONG
 Best Distance: 7626.3604
 All distances:
 o: 7626.3604 <- CHOSEN
 i: 81993.5386
 u: 2949017.8309 (CORRECT)
 a: 5276845.8104
 e: 13072183.5699

Test: Justin - e 2.wav
 Person: Justin
 Actual: e | Predicted: o | WRONG
 Best Distance: 7195.7409
 All distances:
 o: 7195.7409 <- CHOSEN
 i: 74925.5429
 u: 3133617.4652
 a: 5034652.8942
 e: 12984644.0049 (CORRECT)

Test: Justin - o 2.wav
 Person: Justin
 Actual: o | Predicted: o | CORRECT
 Best Distance: 7406.8372
 All distances:
 o: 7406.8372 <- CHOSEN (CORRECT)
 i: 95536.0228
 u: 2082577.1844
 a: 4461738.5413
 e: 11627941.5463

Test: Ucup - a 2.wav
 Person: Ucup
 Actual: a | Predicted: o | WRONG
 Best Distance: 6952.1908
 All distances:
 o: 6952.1908 <- CHOSEN
 i: 72085.6323
 u: 2723030.7204
 a: 4161783.8282 (CORRECT)
 e: 12727396.4723

Test: Ucup - i 2.wav
 Person: Ucup
 Actual: i | Predicted: o | WRONG
 Best Distance: 7933.9456
 All distances:
 o: 7933.9456 <- CHOSEN
 i: 71986.3603 (CORRECT)
 u: 3074635.1056
 a: 4678697.9101
 e: 13492906.6904

Test: Ucup - u 2.wav
 Person: Ucup
 Actual: u | Predicted: o | WRONG
 Best Distance: 6660.2352
 All distances:
 o: 6660.2352 <- CHOSEN
 i: 86725.1952
 u: 2197256.1401 (CORRECT)
 a: 4738054.6183
 e: 12401628.3486

Test: Ucup - e 2.wav
 Person: Ucup
 Actual: e | Predicted: o | WRONG
 Best Distance: 7620.5261
 All distances:
 o: 7620.5261 <- CHOSEN
 i: 39741.9344
 u: 2981530.0926
 a: 4353364.7021
 e: 11262151.7079 (CORRECT)

Test: Ucup - o 2.wav
 Person: Ucup
 Actual: o | Predicted: o | CORRECT
 Best Distance: 4799.9672
 All distances:
 o: 4799.9672 <- CHOSEN (CORRECT)
 i: 14988.8842
 u: 3271131.0645
 a: 5251313.5472
 e: 12956570.1484

Test: gon - a 2.wav
 Person: Gonza
 Actual: a | Predicted: o | WRONG
 Best Distance: 9037.8527
 All distances:
 o: 9037.8527 <- CHOSEN
 i: 29187.8259
 u: 2756867.5154
 a: 4080913.3761 (CORRECT)
 e: 11617655.6387

Test: gon - i 2.wav
 Person: Gonza
 Actual: i | Predicted: o | WRONG
 Best Distance: 7084.8648
 All distances:
 o: 7084.8648 <- CHOSEN
 i: 20240.2882 (CORRECT)
 u: 3141603.5902
 a: 4267173.0772
 e: 11646273.9138

Test: gon - u 2.wav
 Person: Gonza
 Actual: u | Predicted: o | WRONG
 Best Distance: 5955.8819
 All distances:
 o: 5955.8819 <- CHOSEN
 i: 28335.2327
 u: 1623564.2742 (CORRECT)
 a: 3345469.5621
 e: 9814022.2970

Test: gon - e 2.wav
 Person: Gonza
 Actual: e | Predicted: o | WRONG
 Best Distance: 6560.8852
 All distances:
 o: 6560.8852 <- CHOSEN
 i: 37657.5655
 u: 3027406.3146
 a: 7409258.1647
 e: 11921798.5102 (CORRECT)

Test: gon - o 2.wav
 Person: Gonza
 Actual: o | Predicted: o | CORRECT
 Best Distance: 7493.5922
 All distances:
 o: 7493.5922 <- CHOSEN (CORRECT)
 i: 39971.3882
 u: 2346412.9923
 a: 3651647.2050
 e: 10899605.2099

Test: ji - a 2.wav
 Person: Angie
 Actual: a | Predicted: o | WRONG
 Best Distance: 6434.8577
 All distances:
 o: 6434.8577 <- CHOSEN
 i: 23294.1125
 u: 2564646.0197
 a: 3036066.8525 (CORRECT)
 e: 9814772.1654

Test: ji - i 2.wav
Person: Angie
Actual: i | Predicted: o | WRONG
Best Distance: 4985.3573
All distances:
o: 4985.3573 <- CHOSEN
i: 31565.1821 (CORRECT)
u: 2709618.9987
a: 3244270.1996
e: 12344010.3747

Test: ji - u 2.wav
Person: Angie
Actual: u | Predicted: o | WRONG
Best Distance: 3353.6828
All distances:
o: 3353.6828 <- CHOSEN
i: 20369.2851
u: 1809292.9675 (CORRECT)
a: 4102810.2546
e: 11920792.5331

Test: ji - e 2.wav
Person: Angie
Actual: e | Predicted: o | WRONG
Best Distance: 6658.5898
All distances:
o: 6658.5898 <- CHOSEN
i: 23572.4897
u: 3098387.7065
a: 4315540.6519
e: 9746357.2615 (CORRECT)

Test: ji - o 2.wav
Person: Angie
Actual: o | Predicted: o | CORRECT
Best Distance: 4296.4020
All distances:
o: 4296.4020 <- CHOSEN (CORRECT)
i: 25797.1297
u: 1274510.8932
a: 3143563.6920
e: 11314867.5163

Test: nan - a 2.wav
Person: Adnan
Actual: a | Predicted: o | WRONG
Best Distance: 3388.3512
All distances:
o: 3388.3512 <- CHOSEN
i: 32510.3908
a: 3228347.2041 (CORRECT)
u: 3680282.0581
e: 11750820.6292

Test: nan - i 2.wav
 Person: Adnan
 Actual: i | Predicted: o | WRONG
 Best Distance: 6677.0777
 All distances:
 o: 6677.0777 <- CHOSEN
 i: 41342.9037 (CORRECT)
 u: 2238812.4688
 a: 5718158.8459
 e: 11135036.6019

Test: nan - u 2.wav
 Person: Adnan
 Actual: u | Predicted: o | WRONG
 Best Distance: 7255.0411
 All distances:
 o: 7255.0411 <- CHOSEN
 i: 21834.4772
 u: 1484223.9520 (CORRECT)
 a: 3586250.0967
 e: 9191599.2478

Test: nan - e 2.wav
 Person: Adnan
 Actual: e | Predicted: o | WRONG
 Best Distance: 6416.3927
 All distances:
 o: 6416.3927 <- CHOSEN
 i: 30413.4396
 u: 2982644.9322
 a: 4300109.7232
 e: 10752767.7236 (CORRECT)

Test: nan - o 2.wav
 Person: Adnan
 Actual: o | Predicted: o | CORRECT
 Best Distance: 10111.7978
 All distances:
 o: 10111.7978 <- CHOSEN (CORRECT)
 i: 38335.2026
 u: 2441711.7014
 a: 4053034.0200
 e: 12123749.3264

Test: zul - a 2.wav
 Person: Azul
 Actual: a | Predicted: o | WRONG
 Best Distance: 4457.5887
 All distances:
 o: 4457.5887 <- CHOSEN
 i: 17927.6208
 u: 1645925.5580
 a: 3413100.9940 (CORRECT)
 e: 9819546.4789

Test: zul - i 2.wav
Person: Azul
Actual: i | Predicted: o | WRONG
Best Distance: 5017.6826
All distances:
o: 5017.6826 <- CHOSEN
i: 36631.1687 (CORRECT)
u: 2496675.7382
a: 5251499.2985
e: 11925774.7302

Test: zul - u 2.wav
Person: Azul
Actual: u | Predicted: o | WRONG
Best Distance: 6141.4741
All distances:
o: 6141.4741 <- CHOSEN
i: 26802.0629
u: 2349108.0520 (CORRECT)
a: 3750112.8676
e: 11274531.8231

Test: zul - e 2.wav
Person: Azul
Actual: e | Predicted: o | WRONG
Best Distance: 5083.9758
All distances:
o: 5083.9758 <- CHOSEN
i: 20754.9044
u: 1573049.6687
a: 4388607.5284
e: 10088507.0597 (CORRECT)

Test: zul - o 2.wav
Person: Azul
Actual: o | Predicted: o | CORRECT
Best Distance: 4325.9460
All distances:
o: 4325.9460 <- CHOSEN (CORRECT)
i: 38557.4099
u: 2135284.8126
a: 5047827.5712
e: 11627434.9635

OPEN SCENARIO SUMMARY: 10/50 correct = 20.00%

--- SUMMARY FOR HIRA ---
Closed Accuracy: 80.00%
Open Accuracy: 20.00%
Average Accuracy: 50.00%

IV.1.4 Hasil Tes dengan *Template* 'Densu'

```
=====
=== EVALUATION WITH TEMPLATES FROM DENSU ===
=====

Total test files (templates_us): 25
Total test files (templates_other): 50

[CLOSED SCENARIO] (all templates vs test from templates_us)

--- DETAILED CLOSED SCENARIO RESULTS ---

Test: densu - a 3.wav
  Person: densu
  Actual: a | Predicted: a | CORRECT
  Best Distance: 70.2069
  All distances:
    a: 70.2069 <- CHOSEN (CORRECT)
    o: 74.3216
    u: 76.2315
    e: 96.8769
    i: 101.2702

Test: densu - i 3.wav
  Person: densu
  Actual: i | Predicted: i | CORRECT
  Best Distance: 53.9492
  All distances:
    i: 53.9492 <- CHOSEN (CORRECT)
    u: 75.7062
    e: 78.5462
    a: 95.7307
    o: 100.7484

Test: densu - u 3.wav
  Person: densu
  Actual: u | Predicted: u | CORRECT
  Best Distance: 55.6178
  All distances:
    u: 55.6178 <- CHOSEN (CORRECT)
    o: 72.8369
    a: 75.8265
    i: 92.3293
    e: 105.2425

Test: densu - e 3.wav
  Person: densu
  Actual: e | Predicted: e | CORRECT
  Best Distance: 30.6416
  All distances:
    e: 30.6416 <- CHOSEN (CORRECT)
    i: 64.1177
    u: 65.0695
    a: 73.9451
```

o: 74.4279

Test: densu - o 3.wav

Person: densu

Actual: o | Predicted: a | WRONG

Best Distance: 54.9172

All distances:

a: 54.9172 <- CHOSEN

o: 61.5603 (CORRECT)

u: 65.1129

e: 101.6156

i: 108.8738

Test: hira - a 3.wav

Person: hira

Actual: a | Predicted: a | CORRECT

Best Distance: 63.2745

All distances:

a: 63.2745 <- CHOSEN (CORRECT)

u: 72.9438

o: 77.6447

i: 89.6612

e: 108.5596

Test: hira - i 3.wav

Person: hira

Actual: i | Predicted: u | WRONG

Best Distance: 80.0211

All distances:

u: 80.0211 <- CHOSEN

e: 81.3058

a: 89.3321

i: 90.1186 (CORRECT)

o: 94.0986

Test: hira - u 3.wav

Person: hira

Actual: u | Predicted: u | CORRECT

Best Distance: 52.2315

All distances:

u: 52.2315 <- CHOSEN (CORRECT)

o: 56.9199

a: 82.7891

i: 87.1686

e: 94.8692

Test: hira - e 3.wav

Person: hira

Actual: e | Predicted: e | CORRECT

Best Distance: 31.6649

All distances:

e: 31.6649 <- CHOSEN (CORRECT)

u: 67.2318

i: 67.2943

o: 82.4711

a: 93.1578

Test: hira - o 3.wav

Person: hira

Actual: o | Predicted: a | WRONG

Best Distance: 52.5049

All distances:

a: 52.5049 <- CHOSEN

o: 57.9618 (CORRECT)

u: 64.8192

e: 81.7117

i: 104.5562

Test: naufal - a 4.wav

Person: naufal

Actual: a | Predicted: a | CORRECT

Best Distance: 57.6231

All distances:

a: 57.6231 <- CHOSEN (CORRECT)

u: 72.1626

e: 78.4054

o: 83.8841

i: 89.6910

Test: naufal - i 4.wav

Person: naufal

Actual: i | Predicted: i | CORRECT

Best Distance: 59.4456

All distances:

i: 59.4456 <- CHOSEN (CORRECT)

e: 75.9879

u: 78.6571

a: 93.3244

o: 142.4249

Test: naufal - u 4.wav

Person: naufal

Actual: u | Predicted: u | CORRECT

Best Distance: 52.9969

All distances:

u: 52.9969 <- CHOSEN (CORRECT)

o: 68.3638

e: 83.6060

a: 83.9304

i: 103.4871

Test: naufal - e 4.wav

Person: naufal

Actual: e | Predicted: e | CORRECT

Best Distance: 71.9031

All distances:

e: 71.9031 <- CHOSEN (CORRECT)

u: 76.3323

a: 79.4679

i: 83.3870

o: 100.6227

Test: naufal - o 4.wav

Person: naufal

Actual: o | Predicted: o | CORRECT

Best Distance: 64.2112

All distances:

o: 64.2112 <- CHOSEN (CORRECT)

u: 65.4305

a: 67.4120

e: 71.1539

i: 94.5248

Test: wiga - a 3.wav

Person: wiga

Actual: a | Predicted: a | CORRECT

Best Distance: 53.1276

All distances:

a: 53.1276 <- CHOSEN (CORRECT)

o: 70.0291

u: 71.0580

e: 84.9119

i: 104.1089

Test: wiga - i 3.wav

Person: wiga

Actual: i | Predicted: i | CORRECT

Best Distance: 80.4008

All distances:

i: 80.4008 <- CHOSEN (CORRECT)

u: 83.4983

a: 88.8513

e: 114.4101

o: 126.2325

Test: wiga - u 3.wav

Person: wiga

Actual: u | Predicted: u | CORRECT

Best Distance: 52.9289

All distances:

u: 52.9289 <- CHOSEN (CORRECT)

o: 71.6992

a: 89.0054

e: 105.4107

i: 108.4208

Test: wiga - e 3.wav

Person: wiga

Actual: e | Predicted: u | WRONG

Best Distance: 86.2586

All distances:

u: 86.2586 <- CHOSEN

i: 90.7028

a: 109.7578

e: 143.0313 (CORRECT)

o: 162.3077

Test: wiga - o 3.wav

Person: wiga

Actual: o | Predicted: u | WRONG

Best Distance: 63.5333

All distances:

u: 63.5333 <- CHOSEN

o: 94.7962 (CORRECT)

a: 95.8924

e: 123.4073

i: 126.2884

Test: zya - a 3.wav

Person: zya

Actual: a | Predicted: a | CORRECT

Best Distance: 52.8507

All distances:

a: 52.8507 <- CHOSEN (CORRECT)

u: 63.6856

o: 68.4934

e: 70.3881

i: 100.6245

Test: zya - i 3.wav

Person: zya

Actual: i | Predicted: i | CORRECT

Best Distance: 47.4606

All distances:

i: 47.4606 <- CHOSEN (CORRECT)

u: 66.4740

e: 66.7781

o: 94.1389

a: 100.8038

Test: zya - u 3.wav

Person: zya

Actual: u | Predicted: u | CORRECT

Best Distance: 49.1872

All distances:

u: 49.1872 <- CHOSEN (CORRECT)

o: 59.0017

a: 76.4671

e: 92.9694

i: 109.2422

Test: zya - e 3.wav

Person: zya

Actual: e | Predicted: e | CORRECT

Best Distance: 51.3478

All distances:

e: 51.3478 <- CHOSEN (CORRECT)

u: 64.8213

a: 73.6815

o: 75.4071

i: 90.7024

Test: zya - o 3.wav

Person: zya

Actual: o | Predicted: o | CORRECT

Best Distance: 61.4085

All distances:

o: 61.4085 <- CHOSEN (CORRECT)

a: 68.8325

u: 70.5674

e: 83.1020

i: 90.7356

CLOSED SCENARIO SUMMARY: 20/25 correct = 80.00%

[OPEN SCENARIO] (templates from densu vs test from templates_other)

Test files: 50

--- DETAILED OPEN SCENARIO RESULTS ---

Using templates from: ['densu']

Test: dab - a 2.wav

Person: Dabbir

Actual: a | Predicted: a | CORRECT

Best Distance: 3807.7420

All distances:

a: 3807.7420 <- CHOSEN (CORRECT)

u: 847294.8078

i: 2288800.6440

o: 6481228.5485

e: 9725130.9570

Test: dab - i 2.wav

Person: Dabbir

Actual: i | Predicted: a | WRONG

Best Distance: 2691.5105

All distances:

a: 2691.5105 <- CHOSEN

u: 1355975.6307

i: 1744902.2771 (CORRECT)

o: 7211413.1937

e: 11833568.2296

Test: dab - u 2.wav

Person: Dabbir

Actual: u | Predicted: a | WRONG

Best Distance: 3463.5448

All distances:

a: 3463.5448 <- CHOSEN

u: 999973.9590 (CORRECT)

i: 2681350.4614

o: 5601215.2812

e: 10181000.0744

Test: dab - e 2.wav

Person: Dabbir
Actual: e | Predicted: a | WRONG
Best Distance: 3321.2777
All distances:
a: 3321.2777 <- CHOSEN
u: 951153.4067
i: 1742374.3662
o: 5919753.4702
e: 9092350.0444 (CORRECT)

Test: dab - o 2.wav
Person: Dabbir
Actual: o | Predicted: a | WRONG
Best Distance: 3105.7457
All distances:
a: 3105.7457 <- CHOSEN
u: 454118.3929
i: 2178061.7357
o: 7059254.7560 (CORRECT)
e: 10997256.3637

Test: Evelyn - a 2.wav
Person: Evelyn
Actual: a | Predicted: a | CORRECT
Best Distance: 3009.0118
All distances:
a: 3009.0118 <- CHOSEN (CORRECT)
u: 1216292.0820
i: 2685602.4110
o: 8530133.1647
e: 13342057.3645

Test: Evelyn - i 2.wav
Person: Evelyn
Actual: i | Predicted: a | WRONG
Best Distance: 6301.8745
All distances:
a: 6301.8745 <- CHOSEN
u: 1196548.1001
i: 2211677.9292 (CORRECT)
o: 8240201.4437
e: 13399399.6789

Test: Evelyn - u 2.wav
Person: Evelyn
Actual: u | Predicted: a | WRONG
Best Distance: 2948.2039
All distances:
a: 2948.2039 <- CHOSEN
u: 1075355.5227 (CORRECT)
i: 2581973.6277
o: 7639886.7297
e: 13908967.8308

Test: Evelyn - e 2.wav

Person: Evelyn
Actual: e | Predicted: a | WRONG
Best Distance: 4122.3525
All distances:
a: 4122.3525 <- CHOSEN
u: 1335380.8264
i: 2698191.9755
o: 8650183.0786
e: 11255705.7066 (CORRECT)

Test: Evelyn - o 2.wav
Person: Evelyn
Actual: o | Predicted: a | WRONG
Best Distance: 2872.2727
All distances:
a: 2872.2727 <- CHOSEN
u: 1574016.3356
i: 3383406.8663
o: 5621069.2431 (CORRECT)
e: 13341706.1277

Test: Fed - a 2.wav
Person: Fed
Actual: a | Predicted: a | CORRECT
Best Distance: 3136.1342
All distances:
a: 3136.1342 <- CHOSEN (CORRECT)
i: 1800407.7009
u: 1960445.3954
o: 7845231.7793
e: 11814715.0268

Test: Fed - i 2.wav
Person: Fed
Actual: i | Predicted: a | WRONG
Best Distance: 3124.6297
All distances:
a: 3124.6297 <- CHOSEN
u: 1141052.3750
i: 1801842.4523 (CORRECT)
o: 6895221.6703
e: 10140675.7703

Test: Fed - u 2.wav
Person: Fed
Actual: u | Predicted: a | WRONG
Best Distance: 3153.0854
All distances:
a: 3153.0854 <- CHOSEN
u: 916401.8987 (CORRECT)
i: 2182121.2787
o: 6090692.5953
e: 12174188.6771

Test: Fed - e 2.wav

Person: Fed
Actual: e | Predicted: a | WRONG
Best Distance: 2480.8212
All distances:
a: 2480.8212 <- CHOSEN
u: 1440245.2665
i: 3021138.6383
o: 8031649.6558
e: 12316986.1680 (CORRECT)

Test: Fed - o 2.wav
Person: Fed
Actual: o | Predicted: a | WRONG
Best Distance: 2845.4246
All distances:
a: 2845.4246 <- CHOSEN
u: 1713438.1366
i: 2016122.1507
o: 5626565.3089 (CORRECT)
e: 12187419.1669

Test: Akbar - a 2.wav
Person: Akbar
Actual: a | Predicted: a | CORRECT
Best Distance: 2103.5548
All distances:
a: 2103.5548 <- CHOSEN (CORRECT)
u: 2701499.8681
i: 2882920.8548
o: 8991858.2444
e: 13287227.6614

Test: Akbar - i 2.wav
Person: Akbar
Actual: i | Predicted: a | WRONG
Best Distance: 2002.5516
All distances:
a: 2002.5516 <- CHOSEN
u: 1117734.2261
i: 2503059.0932 (CORRECT)
o: 6515207.6729
e: 11494343.0402

Test: Akbar - u 2.wav
Person: Akbar
Actual: u | Predicted: a | WRONG
Best Distance: 3273.6561
All distances:
a: 3273.6561 <- CHOSEN
u: 1245772.1830 (CORRECT)
i: 2475267.2240
o: 6889511.3332
e: 11703506.2606

Test: Akbar - e 2.wav

Person: Akbar
Actual: e | Predicted: a | WRONG
Best Distance: 2448.8800
All distances:
a: 2448.8800 <- CHOSEN
u: 2244541.8550
i: 3568316.5951
o: 7890756.3308
e: 11651747.4389 (CORRECT)

Test: Akbar - o 2.wav
Person: Akbar
Actual: o | Predicted: a | WRONG
Best Distance: 2789.6347
All distances:
a: 2789.6347 <- CHOSEN
u: 1800892.1944
i: 2249383.9462
o: 6285917.4307 (CORRECT)
e: 12887808.3115

Test: Justin - a 2.wav
Person: Justin
Actual: a | Predicted: a | CORRECT
Best Distance: 2742.1139
All distances:
a: 2742.1139 <- CHOSEN (CORRECT)
u: 1484486.7509
i: 2639461.4629
o: 8328509.2458
e: 12203932.1744

Test: Justin - i 2.wav
Person: Justin
Actual: i | Predicted: a | WRONG
Best Distance: 2910.1600
All distances:
a: 2910.1600 <- CHOSEN
u: 1558104.0240
i: 2954377.5532 (CORRECT)
o: 8476373.8406
e: 12726435.1039

Test: Justin - u 2.wav
Person: Justin
Actual: u | Predicted: a | WRONG
Best Distance: 3346.3573
All distances:
a: 3346.3573 <- CHOSEN
u: 1465912.2535 (CORRECT)
i: 2406847.6071
o: 8307064.0354
e: 12741689.8268

Test: Justin - e 2.wav

Person: Justin
Actual: e | Predicted: a | WRONG
Best Distance: 3163.6298
All distances:
a: 3163.6298 <- CHOSEN
u: 1419407.9001
i: 2965856.7076
o: 9330476.4655
e: 13456252.6536 (CORRECT)

Test: Justin - o 2.wav
Person: Justin
Actual: o | Predicted: a | WRONG
Best Distance: 4780.9416
All distances:
a: 4780.9416 <- CHOSEN
u: 1087775.0337
i: 2129745.9758
o: 8232438.8092 (CORRECT)
e: 13439618.1426

Test: Ucup - a 2.wav
Person: Ucup
Actual: a | Predicted: a | CORRECT
Best Distance: 3387.9413
All distances:
a: 3387.9413 <- CHOSEN (CORRECT)
u: 1138562.6960
i: 2556709.2094
o: 8285782.7880
e: 12865287.5732

Test: Ucup - i 2.wav
Person: Ucup
Actual: i | Predicted: a | WRONG
Best Distance: 4510.8802
All distances:
a: 4510.8802 <- CHOSEN
u: 1050886.7949
i: 2887878.4723 (CORRECT)
o: 8557729.1796
e: 13355563.7554

Test: Ucup - u 2.wav
Person: Ucup
Actual: u | Predicted: a | WRONG
Best Distance: 4144.9676
All distances:
a: 4144.9676 <- CHOSEN
u: 1263709.8240 (CORRECT)
i: 2293250.2929
o: 8189372.9015
e: 13018013.7767

Test: Ucup - e 2.wav

Person: Ucup
Actual: e | Predicted: a | WRONG
Best Distance: 2364.0717
All distances:
a: 2364.0717 <- CHOSEN
u: 728893.7371
i: 2465191.3361
o: 8044725.7379
e: 10795610.7375 (CORRECT)

Test: Ucup - o 2.wav
Person: Ucup
Actual: o | Predicted: a | WRONG
Best Distance: 2961.3692
All distances:
a: 2961.3692 <- CHOSEN
u: 943889.7606
i: 1934406.9354
o: 6695272.3685 (CORRECT)
e: 11859298.8165

Test: gon - a 2.wav
Person: Gonza
Actual: a | Predicted: a | CORRECT
Best Distance: 3919.7957
All distances:
a: 3919.7957 <- CHOSEN (CORRECT)
u: 1112585.1389
i: 1823559.4862
o: 7975095.5549
e: 10085487.5149

Test: gon - i 2.wav
Person: Gonza
Actual: i | Predicted: a | WRONG
Best Distance: 2552.7249
All distances:
a: 2552.7249 <- CHOSEN
u: 1330824.6423
i: 1451408.8185 (CORRECT)
o: 6368622.4775
e: 10129241.7455

Test: gon - u 2.wav
Person: Gonza
Actual: u | Predicted: a | WRONG
Best Distance: 2658.7751
All distances:
a: 2658.7751 <- CHOSEN
u: 421458.2684 (CORRECT)
i: 1450416.3192
o: 5862625.4337
e: 9726799.9987

Test: gon - e 2.wav

Person: Gonza
Actual: e | Predicted: a | WRONG
Best Distance: 2788.0977
All distances:
a: 2788.0977 <- CHOSEN
u: 1329704.0632
i: 2303710.8605
o: 8710376.1560
e: 9962862.1815 (CORRECT)

Test: gon - o 2.wav
Person: Gonza
Actual: o | Predicted: a | WRONG
Best Distance: 2862.6573
All distances:
a: 2862.6573 <- CHOSEN
u: 1423498.3652
i: 1893205.2189
o: 7683588.3600 (CORRECT)
e: 9941399.9506

Test: ji - a 2.wav
Person: Angie
Actual: a | Predicted: a | CORRECT
Best Distance: 3424.5106
All distances:
a: 3424.5106 <- CHOSEN (CORRECT)
u: 1287116.0921
i: 2639839.8919
o: 9619740.3053
e: 11826466.3964

Test: ji - i 2.wav
Person: Angie
Actual: i | Predicted: a | WRONG
Best Distance: 2992.9879
All distances:
a: 2992.9879 <- CHOSEN
i: 1229523.6089 (CORRECT)
u: 1644399.0437
o: 6105476.6973
e: 13183036.8717

Test: ji - u 2.wav
Person: Angie
Actual: u | Predicted: a | WRONG
Best Distance: 2406.7696
All distances:
a: 2406.7696 <- CHOSEN
u: 1042001.9403 (CORRECT)
i: 1882661.7829
o: 6474788.8057
e: 11989483.8105

Test: ji - e 2.wav

Person: Angie
Actual: e | Predicted: a | WRONG
Best Distance: 2980.7909
All distances:
a: 2980.7909 <- CHOSEN
i: 1990447.4907
u: 2679103.6196
o: 8147405.5504
e: 8937419.3761 (CORRECT)

Test: ji - o 2.wav
Person: Angie
Actual: o | Predicted: a | WRONG
Best Distance: 2530.9865
All distances:
a: 2530.9865 <- CHOSEN
u: 1289451.1103
i: 3551946.9246
o: 6816085.4155 (CORRECT)
e: 10361445.4688

Test: nan - a 2.wav
Person: Adnan
Actual: a | Predicted: a | CORRECT
Best Distance: 3711.8978
All distances:
a: 3711.8978 <- CHOSEN (CORRECT)
i: 2183949.4730
u: 2226385.7630
o: 8367509.8178
e: 11134806.7591

Test: nan - i 2.wav
Person: Adnan
Actual: i | Predicted: a | WRONG
Best Distance: 2825.2558
All distances:
a: 2825.2558 <- CHOSEN
u: 945518.9548
i: 1929108.8160 (CORRECT)
o: 7300547.4808
e: 10206065.7337

Test: nan - u 2.wav
Person: Adnan
Actual: u | Predicted: a | WRONG
Best Distance: 3276.2274
All distances:
a: 3276.2274 <- CHOSEN
u: 467007.5235 (CORRECT)
i: 1364971.2838
o: 5837780.0791
e: 9217355.4614

Test: nan - e 2.wav

Person: Adnan
Actual: e | Predicted: a | WRONG
Best Distance: 2707.1067
All distances:
a: 2707.1067 <- CHOSEN
i: 1937291.7382
u: 2027447.1650
o: 7170259.4896
e: 10329895.1628 (CORRECT)

Test: nan - o 2.wav
Person: Adnan
Actual: o | Predicted: a | WRONG
Best Distance: 2360.9635
All distances:
a: 2360.9635 <- CHOSEN
u: 728853.8451
i: 2181284.0685
o: 6517116.4862 (CORRECT)
e: 11100338.9302

Test: zul - a 2.wav
Person: Azul
Actual: a | Predicted: a | CORRECT
Best Distance: 3430.9444
All distances:
a: 3430.9444 <- CHOSEN (CORRECT)
u: 649001.3688
i: 2201962.9045
o: 6850537.5736
e: 10821978.4975

Test: zul - i 2.wav
Person: Azul
Actual: i | Predicted: a | WRONG
Best Distance: 2548.5997
All distances:
a: 2548.5997 <- CHOSEN
u: 638911.6546
i: 2239649.4005 (CORRECT)
o: 8760791.1320
e: 10297268.8947

Test: zul - u 2.wav
Person: Azul
Actual: u | Predicted: a | WRONG
Best Distance: 3301.1932
All distances:
a: 3301.1932 <- CHOSEN
u: 641491.9008 (CORRECT)
i: 2356318.2730
o: 6903720.3846
e: 11410829.1340

Test: zul - e 2.wav

```
Person: Azul
Actual: e | Predicted: a | WRONG
Best Distance: 3859.5236
All distances:
  a: 3859.5236 <- CHOSEN
  u: 1247455.4737
  i: 2743991.6996
  o: 7795897.8645
  e: 10354471.1021 (CORRECT)
```

```
Test: zul - o 2.wav
Person: Azul
Actual: o | Predicted: a | WRONG
Best Distance: 3630.6758
All distances:
  a: 3630.6758 <- CHOSEN
  u: 677125.5304
  i: 2474988.7517
  o: 6482250.4526 (CORRECT)
  e: 10883416.4059
```

OPEN SCENARIO SUMMARY: 10/50 correct = 20.00%

```
--- SUMMARY FOR DENSU ---
Closed Accuracy: 80.00%
Open Accuracy: 20.00%
Average Accuracy: 50.00%
```

IV.1.5 Hasil Tes dengan *Template* 'Naufal'

```
=====
=== EVALUATION WITH TEMPLATES FROM NAUFAL ===
=====

Total test files (templates_us): 25
Total test files (templates_other): 50

[CLOSED SCENARIO] (all templates vs test from templates_us)

--- DETAILED CLOSED SCENARIO RESULTS ---

Test: densu - a 3.wav
Person: densu
Actual: a | Predicted: a | CORRECT
Best Distance: 70.2069
All distances:
  a: 70.2069 <- CHOSEN (CORRECT)
  o: 74.3216
  u: 76.2315
  e: 96.8769
  i: 101.2702

Test: densu - i 3.wav
```

Person: densu
Actual: i | Predicted: i | CORRECT
Best Distance: 53.9492
All distances:
i: 53.9492 <- CHOSEN (CORRECT)
u: 75.7062
e: 78.5462
a: 95.7307
o: 100.7484

Test: densu - u 3.wav
Person: densu
Actual: u | Predicted: u | CORRECT
Best Distance: 55.6178
All distances:
u: 55.6178 <- CHOSEN (CORRECT)
o: 72.8369
a: 75.8265
i: 92.3293
e: 105.2425

Test: densu - e 3.wav
Person: densu
Actual: e | Predicted: e | CORRECT
Best Distance: 30.6416
All distances:
e: 30.6416 <- CHOSEN (CORRECT)
i: 64.1177
u: 65.0695
a: 73.9451
o: 74.4279

Test: densu - o 3.wav
Person: densu
Actual: o | Predicted: a | WRONG
Best Distance: 54.9172
All distances:
a: 54.9172 <- CHOSEN
o: 61.5603 (CORRECT)
u: 65.1129
e: 101.6156
i: 108.8738

Test: hira - a 3.wav
Person: hira
Actual: a | Predicted: a | CORRECT
Best Distance: 63.2745
All distances:
a: 63.2745 <- CHOSEN (CORRECT)
u: 72.9438
o: 77.6447
i: 89.6612
e: 108.5596

Test: hira - i 3.wav

Person: hira
Actual: i | Predicted: u | WRONG
Best Distance: 80.0211
All distances:
u: 80.0211 <- CHOSEN
e: 81.3058
a: 89.3321
i: 90.1186 (CORRECT)
o: 94.0986

Test: hira - u 3.wav
Person: hira
Actual: u | Predicted: u | CORRECT
Best Distance: 52.2315
All distances:
u: 52.2315 <- CHOSEN (CORRECT)
o: 56.9199
a: 82.7891
i: 87.1686
e: 94.8692

Test: hira - e 3.wav
Person: hira
Actual: e | Predicted: e | CORRECT
Best Distance: 31.6649
All distances:
e: 31.6649 <- CHOSEN (CORRECT)
u: 67.2318
i: 67.2943
o: 82.4711
a: 93.1578

Test: hira - o 3.wav
Person: hira
Actual: o | Predicted: a | WRONG
Best Distance: 52.5049
All distances:
a: 52.5049 <- CHOSEN
o: 57.9618 (CORRECT)
u: 64.8192
e: 81.7117
i: 104.5562

Test: naufal - a 4.wav
Person: naufal
Actual: a | Predicted: a | CORRECT
Best Distance: 57.6231
All distances:
a: 57.6231 <- CHOSEN (CORRECT)
u: 72.1626
e: 78.4054
o: 83.8841
i: 89.6910

Test: naufal - i 4.wav

Person: naufal
Actual: i | Predicted: i | CORRECT
Best Distance: 59.4456
All distances:
i: 59.4456 <- CHOSEN (CORRECT)
e: 75.9879
u: 78.6571
a: 93.3244
o: 142.4249

Test: naufal - u 4.wav
Person: naufal
Actual: u | Predicted: u | CORRECT
Best Distance: 52.9969
All distances:
u: 52.9969 <- CHOSEN (CORRECT)
o: 68.3638
e: 83.6060
a: 83.9304
i: 103.4871

Test: naufal - e 4.wav
Person: naufal
Actual: e | Predicted: e | CORRECT
Best Distance: 71.9031
All distances:
e: 71.9031 <- CHOSEN (CORRECT)
u: 76.3323
a: 79.4679
i: 83.3870
o: 100.6227

Test: naufal - o 4.wav
Person: naufal
Actual: o | Predicted: o | CORRECT
Best Distance: 64.2112
All distances:
o: 64.2112 <- CHOSEN (CORRECT)
u: 65.4305
a: 67.4120
e: 71.1539
i: 94.5248

Test: wiga - a 3.wav
Person: wiga
Actual: a | Predicted: a | CORRECT
Best Distance: 53.1276
All distances:
a: 53.1276 <- CHOSEN (CORRECT)
o: 70.0291
u: 71.0580
e: 84.9119
i: 104.1089

Test: wiga - i 3.wav

Person: wiga
Actual: i | Predicted: i | CORRECT
Best Distance: 80.4008
All distances:
i: 80.4008 <- CHOSEN (CORRECT)
u: 83.4983
a: 88.8513
e: 114.4101
o: 126.2325

Test: wiga - u 3.wav
Person: wiga
Actual: u | Predicted: u | CORRECT
Best Distance: 52.9289
All distances:
u: 52.9289 <- CHOSEN (CORRECT)
o: 71.6992
a: 89.0054
e: 105.4107
i: 108.4208

Test: wiga - e 3.wav
Person: wiga
Actual: e | Predicted: u | WRONG
Best Distance: 86.2586
All distances:
u: 86.2586 <- CHOSEN
i: 90.7028
a: 109.7578
e: 143.0313 (CORRECT)
o: 162.3077

Test: wiga - o 3.wav
Person: wiga
Actual: o | Predicted: u | WRONG
Best Distance: 63.5333
All distances:
u: 63.5333 <- CHOSEN
o: 94.7962 (CORRECT)
a: 95.8924
e: 123.4073
i: 126.2884

Test: zya - a 3.wav
Person: zya
Actual: a | Predicted: a | CORRECT
Best Distance: 52.8507
All distances:
a: 52.8507 <- CHOSEN (CORRECT)
u: 63.6856
o: 68.4934
e: 70.3881
i: 100.6245

Test: zya - i 3.wav

Person: zya
Actual: i | Predicted: i | CORRECT
Best Distance: 47.4606
All distances:
i: 47.4606 <- CHOSEN (CORRECT)
u: 66.4740
e: 66.7781
o: 94.1389
a: 100.8038

Test: zya - u 3.wav
Person: zya
Actual: u | Predicted: u | CORRECT
Best Distance: 49.1872
All distances:
u: 49.1872 <- CHOSEN (CORRECT)
o: 59.0017
a: 76.4671
e: 92.9694
i: 109.2422

Test: zya - e 3.wav
Person: zya
Actual: e | Predicted: e | CORRECT
Best Distance: 51.3478
All distances:
e: 51.3478 <- CHOSEN (CORRECT)
u: 64.8213
a: 73.6815
o: 75.4071
i: 90.7024

Test: zya - o 3.wav
Person: zya
Actual: o | Predicted: o | CORRECT
Best Distance: 61.4085
All distances:
o: 61.4085 <- CHOSEN (CORRECT)
a: 68.8325
u: 70.5674
e: 83.1020
i: 90.7356

CLOSED SCENARIO SUMMARY: 20/25 correct = 80.00%

[OPEN SCENARIO] (templates from naufal vs test from templates_other)
Test files: 50

=== BUILDING GENERALIZED TEMPLATES ===

Building generalized template for vowel: a
- Created 3 segments with mean and covariance parameters
Building generalized template for vowel: i
- Created 3 segments with mean and covariance parameters
Building generalized template for vowel: u
- Created 3 segments with mean and covariance parameters

```
Building generalized template for vowel: e
- Created 3 segments with mean and covariance parameters
Building generalized template for vowel: o
- Created 3 segments with mean and covariance parameters
Generalized templates built successfully!
```

```
--- DETAILED OPEN SCENARIO RESULTS ---
Using templates from: ['naufal']
```

```
Test: dab - a 2.wav
Person: Dabbir
Actual: a | Predicted: u | WRONG
Best Distance: 297.4818
All distances:
  u: 297.4818 <- CHOSEN
  e: 541.5259
  o: 953.2696
  a: 953.3399 (CORRECT)
  i: 1218.9822
```

```
Test: dab - i 2.wav
Person: Dabbir
Actual: i | Predicted: u | WRONG
Best Distance: 379.0626
All distances:
  u: 379.0626 <- CHOSEN
  e: 429.0749
  i: 575.9560 (CORRECT)
  o: 1209.0239
  a: 1213.4682
```

```
Test: dab - u 2.wav
Person: Dabbir
Actual: u | Predicted: u | CORRECT
Best Distance: 186.7188
All distances:
  u: 186.7188 <- CHOSEN (CORRECT)
  e: 374.9186
  o: 758.9780
  i: 809.4668
  a: 924.7832
```

```
Test: dab - e 2.wav
Person: Dabbir
Actual: e | Predicted: e | CORRECT
Best Distance: 256.6179
All distances:
  e: 256.6179 <- CHOSEN (CORRECT)
  u: 265.7017
  o: 634.0604
  a: 743.2674
  i: 1087.0756
```

```
Test: dab - o 2.wav
Person: Dabbir
```

Actual: o | Predicted: u | WRONG
Best Distance: 231.0991
All distances:
u: 231.0991 <- CHOSEN
e: 358.1805
o: 802.5039 (CORRECT)
a: 997.3550
i: 1176.6469

Test: Evelyn - a 2.wav
Person: Evelyn
Actual: a | Predicted: u | WRONG
Best Distance: 321.8758
All distances:
u: 321.8758 <- CHOSEN
e: 708.6280
i: 1626.2727
a: 1651.0516 (CORRECT)
o: 1835.9864

Test: Evelyn - i 2.wav
Person: Evelyn
Actual: i | Predicted: u | WRONG
Best Distance: 417.8222
All distances:
u: 417.8222 <- CHOSEN
e: 663.9215
i: 741.8623 (CORRECT)
a: 1873.3303
o: 2351.9697

Test: Evelyn - u 2.wav
Person: Evelyn
Actual: u | Predicted: u | CORRECT
Best Distance: 402.5364
All distances:
u: 402.5364 <- CHOSEN (CORRECT)
e: 586.3279
i: 1125.5393
a: 1390.7504
o: 1393.8698

Test: Evelyn - e 2.wav
Person: Evelyn
Actual: e | Predicted: u | WRONG
Best Distance: 324.8466
All distances:
u: 324.8466 <- CHOSEN
e: 639.7614 (CORRECT)
i: 996.5938
a: 1381.3900
o: 2276.0968

Test: Evelyn - o 2.wav
Person: Evelyn

Actual: o | Predicted: u | WRONG
Best Distance: 289.5587
All distances:
u: 289.5587 <- CHOSEN
e: 636.9797
o: 968.6360 (CORRECT)
a: 1050.6303
i: 1470.5644

Test: Fed - a 2.wav
Person: Fed
Actual: a | Predicted: u | WRONG
Best Distance: 498.7976
All distances:
u: 498.7976 <- CHOSEN
e: 527.8904
i: 980.8558
a: 1142.0266 (CORRECT)
o: 1691.2306

Test: Fed - i 2.wav
Person: Fed
Actual: i | Predicted: u | WRONG
Best Distance: 384.3996
All distances:
u: 384.3996 <- CHOSEN
e: 495.5632
i: 865.9186 (CORRECT)
a: 977.0855
o: 1172.4329

Test: Fed - u 2.wav
Person: Fed
Actual: u | Predicted: u | CORRECT
Best Distance: 246.5038
All distances:
u: 246.5038 <- CHOSEN (CORRECT)
e: 367.0489
i: 827.5216
o: 1061.6971
a: 1268.3068

Test: Fed - e 2.wav
Person: Fed
Actual: e | Predicted: u | WRONG
Best Distance: 448.7425
All distances:
u: 448.7425 <- CHOSEN
e: 516.0046 (CORRECT)
i: 945.7009
a: 2492.5946
o: 2619.0420

Test: Fed - o 2.wav
Person: Fed

Actual: o | Predicted: u | WRONG
Best Distance: 338.0453
All distances:
u: 338.0453 <- CHOSEN
e: 742.5764
i: 839.7717
a: 1153.7812
o: 1575.7775 (CORRECT)

Test: Akbar - a 2.wav
Person: Akbar
Actual: a | Predicted: u | WRONG
Best Distance: 411.1318
All distances:
u: 411.1318 <- CHOSEN
e: 484.6355
i: 935.7096
a: 1466.1317 (CORRECT)
o: 1760.5632

Test: Akbar - i 2.wav
Person: Akbar
Actual: i | Predicted: u | WRONG
Best Distance: 450.1409
All distances:
u: 450.1409 <- CHOSEN
e: 473.3880
i: 618.2280 (CORRECT)
o: 1069.8329
a: 1310.2879

Test: Akbar - u 2.wav
Person: Akbar
Actual: u | Predicted: u | CORRECT
Best Distance: 246.9421
All distances:
u: 246.9421 <- CHOSEN (CORRECT)
e: 478.1374
i: 637.0976
o: 844.0804
a: 1218.3414

Test: Akbar - e 2.wav
Person: Akbar
Actual: e | Predicted: u | WRONG
Best Distance: 485.6153
All distances:
u: 485.6153 <- CHOSEN
e: 626.1810 (CORRECT)
i: 1116.3379
o: 1483.4527
a: 1590.0574

Test: Akbar - o 2.wav
Person: Akbar

Actual: o | Predicted: u | WRONG
Best Distance: 429.4364
All distances:
u: 429.4364 <- CHOSEN
e: 464.1951
i: 881.7799
a: 1276.9594
o: 1282.7668 (CORRECT)

Test: Justin - a 2.wav
Person: Justin
Actual: a | Predicted: u | WRONG
Best Distance: 291.0576
All distances:
u: 291.0576 <- CHOSEN
e: 753.0489
i: 1297.8606
a: 1808.2484 (CORRECT)
o: 1874.4556

Test: Justin - i 2.wav
Person: Justin
Actual: i | Predicted: u | WRONG
Best Distance: 296.8492
All distances:
u: 296.8492 <- CHOSEN
e: 642.7748
i: 948.1201 (CORRECT)
a: 2041.8597
o: 2246.8703

Test: Justin - u 2.wav
Person: Justin
Actual: u | Predicted: u | CORRECT
Best Distance: 231.2381
All distances:
u: 231.2381 <- CHOSEN (CORRECT)
e: 797.3333
i: 1537.3965
o: 2453.7144
a: 3190.2177

Test: Justin - e 2.wav
Person: Justin
Actual: e | Predicted: u | WRONG
Best Distance: 334.7408
All distances:
u: 334.7408 <- CHOSEN
e: 592.0287 (CORRECT)
i: 1160.1768
a: 2057.1106
o: 2476.4867

Test: Justin - o 2.wav
Person: Justin

Actual: o | Predicted: u | WRONG
Best Distance: 186.3943
All distances:
u: 186.3943 <- CHOSEN
e: 609.3763
i: 1503.1772
a: 1569.9326
o: 1630.3420 (CORRECT)

Test: Ucup - a 2.wav
Person: Ucup
Actual: a | Predicted: u | WRONG
Best Distance: 374.4898
All distances:
u: 374.4898 <- CHOSEN
e: 606.2242
a: 1450.0895 (CORRECT)
i: 1629.4734
o: 2009.5755

Test: Ucup - i 2.wav
Person: Ucup
Actual: i | Predicted: u | WRONG
Best Distance: 363.9116
All distances:
u: 363.9116 <- CHOSEN
e: 639.4093
i: 1014.1122 (CORRECT)
a: 1694.0484
o: 2012.1704

Test: Ucup - u 2.wav
Person: Ucup
Actual: u | Predicted: u | CORRECT
Best Distance: 208.2197
All distances:
u: 208.2197 <- CHOSEN (CORRECT)
e: 625.5063
i: 1531.5456
o: 1544.8176
a: 1827.9142

Test: Ucup - e 2.wav
Person: Ucup
Actual: e | Predicted: e | CORRECT
Best Distance: 398.7391
All distances:
e: 398.7391 <- CHOSEN (CORRECT)
u: 501.1666
i: 875.4419
a: 997.3702
o: 1536.4719

Test: Ucup - o 2.wav
Person: Ucup

Actual: o | Predicted: u | WRONG
Best Distance: 346.2196
All distances:
u: 346.2196 <- CHOSEN
e: 471.6753
i: 700.1379
o: 1206.0005 (CORRECT)
a: 1408.6667

Test: gon - a 2.wav
Person: Gonza
Actual: a | Predicted: u | WRONG
Best Distance: 437.4954
All distances:
u: 437.4954 <- CHOSEN
i: 692.4686
e: 720.1099
a: 1077.7378 (CORRECT)
o: 1319.9287

Test: gon - i 2.wav
Person: Gonza
Actual: i | Predicted: u | WRONG
Best Distance: 446.8275
All distances:
u: 446.8275 <- CHOSEN
i: 472.9783 (CORRECT)
e: 547.9523
a: 1116.8641
o: 1230.0472

Test: gon - u 2.wav
Person: Gonza
Actual: u | Predicted: u | CORRECT
Best Distance: 259.3919
All distances:
u: 259.3919 <- CHOSEN (CORRECT)
e: 339.3974
a: 971.0249
i: 1058.0573
o: 1253.9095

Test: gon - e 2.wav
Person: Gonza
Actual: e | Predicted: e | CORRECT
Best Distance: 312.2973
All distances:
e: 312.2973 <- CHOSEN (CORRECT)
u: 354.0166
o: 914.5955
a: 977.9398
i: 1040.0768

Test: gon - o 2.wav
Person: Gonza

Actual: o | Predicted: u | WRONG
Best Distance: 262.8113
All distances:
u: 262.8113 <- CHOSEN
e: 405.7040
i: 551.5076
o: 946.8669 (CORRECT)
a: 1362.2952

Test: ji - a 2.wav
Person: Angie
Actual: a | Predicted: u | WRONG
Best Distance: 334.5482
All distances:
u: 334.5482 <- CHOSEN
e: 899.5124
i: 959.2211
a: 1101.8870 (CORRECT)
o: 1306.8689

Test: ji - i 2.wav
Person: Angie
Actual: i | Predicted: u | WRONG
Best Distance: 438.0598
All distances:
u: 438.0598 <- CHOSEN
i: 589.4035 (CORRECT)
e: 686.3901
o: 816.8200
a: 1235.0250

Test: ji - u 2.wav
Person: Angie
Actual: u | Predicted: u | CORRECT
Best Distance: 323.5029
All distances:
u: 323.5029 <- CHOSEN (CORRECT)
e: 675.3123
o: 828.7492
i: 1002.9233
a: 1355.9097

Test: ji - e 2.wav
Person: Angie
Actual: e | Predicted: e | CORRECT
Best Distance: 362.8654
All distances:
e: 362.8654 <- CHOSEN (CORRECT)
u: 401.6165
i: 645.3798
a: 1270.7879
o: 1453.8195

Test: ji - o 2.wav
Person: Angie

Actual: o | Predicted: u | WRONG
Best Distance: 267.2401
All distances:
u: 267.2401 <- CHOSEN
e: 518.8127
o: 837.0462 (CORRECT)
i: 953.8060
a: 1063.1603

Test: nan - a 2.wav
Person: Adnan
Actual: a | Predicted: u | WRONG
Best Distance: 327.3626
All distances:
u: 327.3626 <- CHOSEN
i: 532.3376
e: 708.2905
a: 757.7720 (CORRECT)
o: 1561.2669

Test: nan - i 2.wav
Person: Adnan
Actual: i | Predicted: u | WRONG
Best Distance: 198.0844
All distances:
u: 198.0844 <- CHOSEN
e: 488.9121
i: 703.5614 (CORRECT)
o: 1098.0349
a: 1106.4871

Test: nan - u 2.wav
Person: Adnan
Actual: u | Predicted: u | CORRECT
Best Distance: 274.1925
All distances:
u: 274.1925 <- CHOSEN (CORRECT)
e: 339.0280
i: 888.7020
o: 1018.8466
a: 1160.7483

Test: nan - e 2.wav
Person: Adnan
Actual: e | Predicted: u | WRONG
Best Distance: 258.3859
All distances:
u: 258.3859 <- CHOSEN
e: 360.7460 (CORRECT)
i: 638.6496
a: 1192.0171
o: 1450.6990

Test: nan - o 2.wav
Person: Adnan

Actual: o | Predicted: u | WRONG
Best Distance: 287.3019
All distances:
u: 287.3019 <- CHOSEN
e: 380.3445
a: 849.3978
o: 1042.4193 (CORRECT)
i: 1169.7487

Test: zul - a 2.wav
Person: Azul
Actual: a | Predicted: u | WRONG
Best Distance: 370.5822
All distances:
u: 370.5822 <- CHOSEN
e: 746.8069
o: 880.0587
a: 1091.6273 (CORRECT)
i: 1442.1191

Test: zul - i 2.wav
Person: Azul
Actual: i | Predicted: u | WRONG
Best Distance: 247.1584
All distances:
u: 247.1584 <- CHOSEN
e: 402.2415
i: 745.6788 (CORRECT)
o: 779.0746
a: 1085.3109

Test: zul - u 2.wav
Person: Azul
Actual: u | Predicted: u | CORRECT
Best Distance: 163.2853
All distances:
u: 163.2853 <- CHOSEN (CORRECT)
e: 473.9900
o: 741.0533
a: 1183.7386
i: 1325.1718

Test: zul - e 2.wav
Person: Azul
Actual: e | Predicted: u | WRONG
Best Distance: 217.2870
All distances:
u: 217.2870 <- CHOSEN
e: 369.1541 (CORRECT)
o: 909.3446
a: 931.9653
i: 1095.2444

Test: zul - o 2.wav
Person: Azul

Actual: o | Predicted: u | WRONG
 Best Distance: 199.6131
 All distances:
 u: 199.6131 <- CHOSEN
 e: 438.1294
 o: 632.4461 (CORRECT)
 a: 954.5275
 i: 1072.4106

OPEN SCENARIO SUMMARY: 14/50 correct = 28.00%

--- SUMMARY FOR NAUFAL ---
 Closed Accuracy: 80.00%
 Open Accuracy: 28.00%
 Average Accuracy: 54.00%

Overall Results

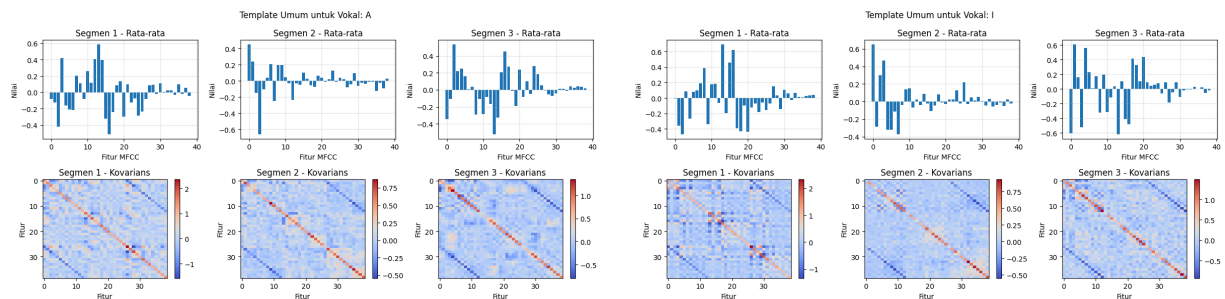
=====
 === OVERALL RESULTS ACROSS ALL TEMPLATE PERSONS ===
 =====

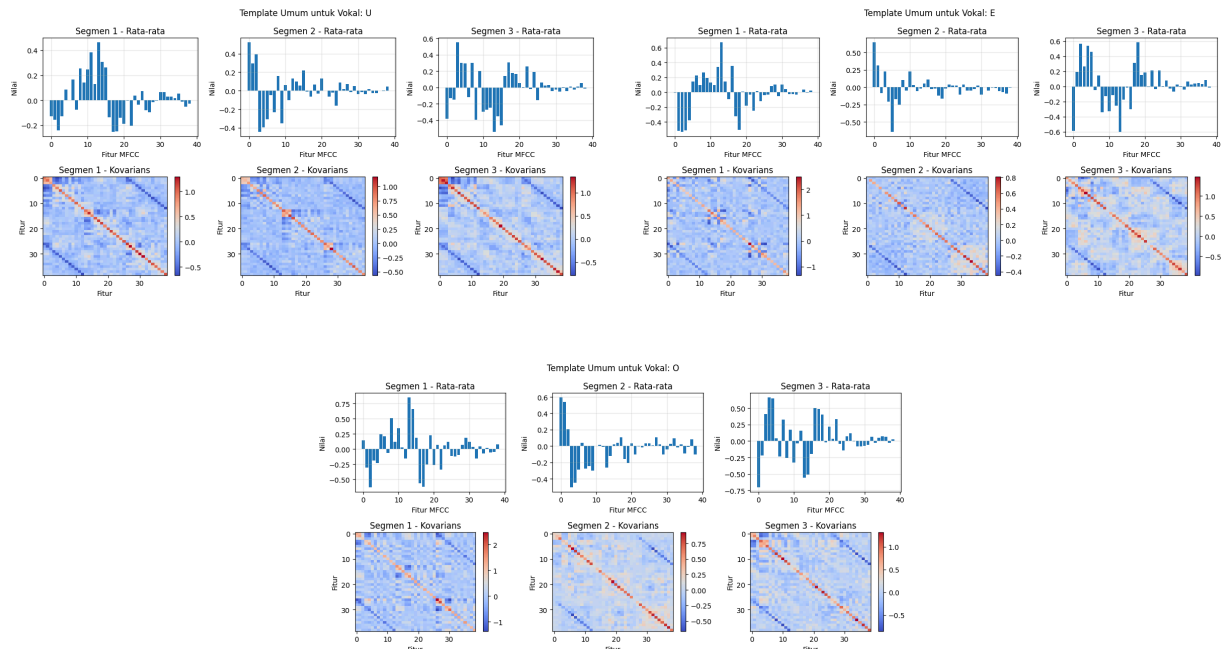
Overall Closed Accuracy: 80.00%
 Overall Open Accuracy: 21.60%
 Overall Average Accuracy: 50.80%

SI Closed Results

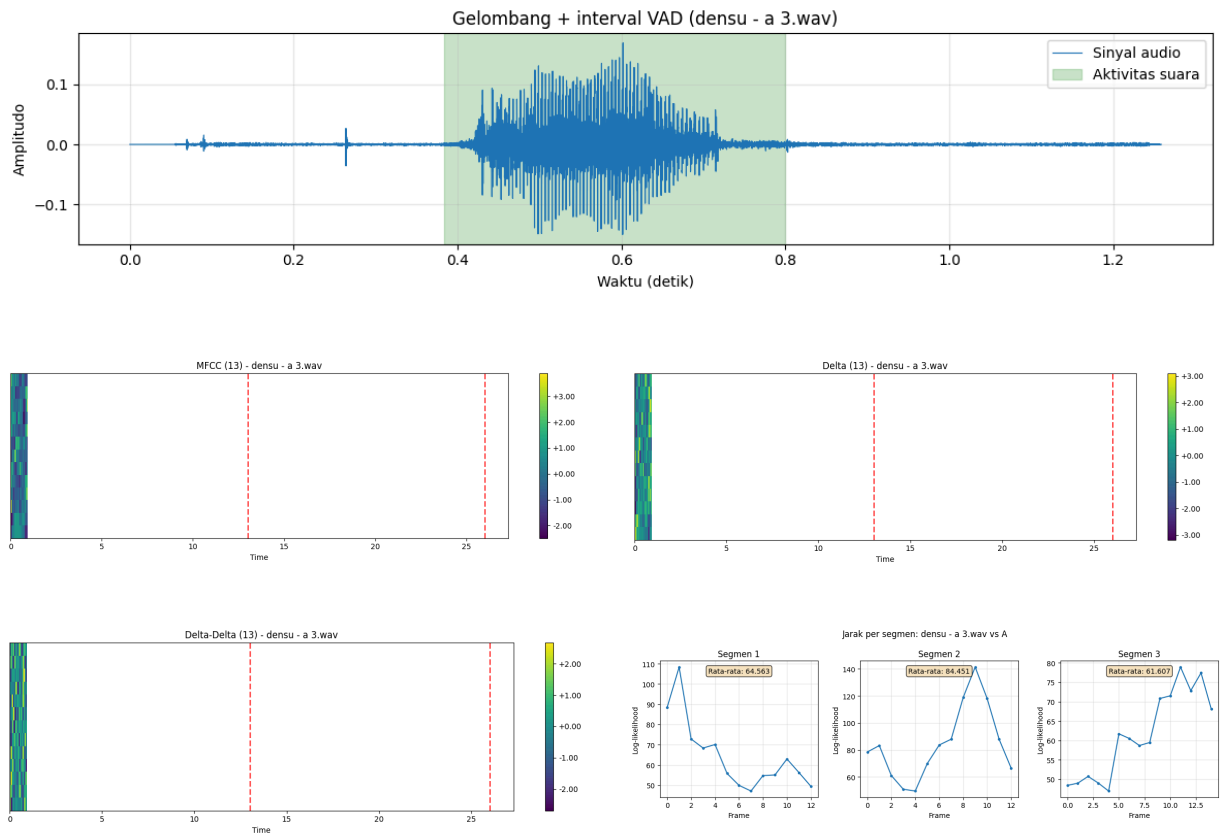
SI-Closed Results: 13/25 = 52.00%
 SI-Closed Accuracy: 52.00%

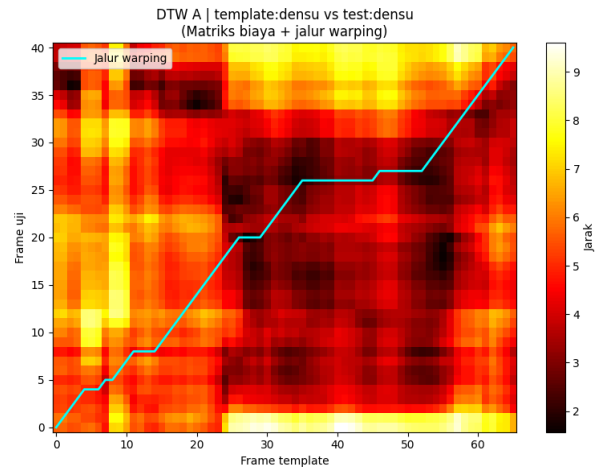
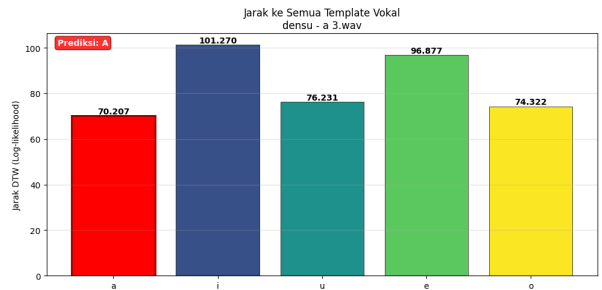
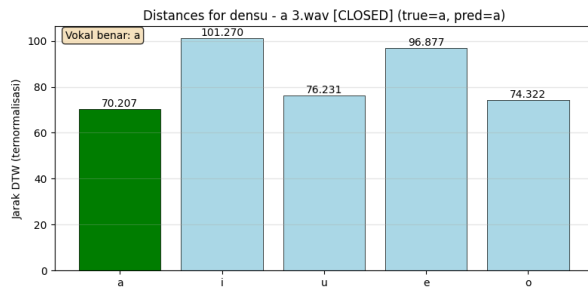
Visualisasi *generalized template*



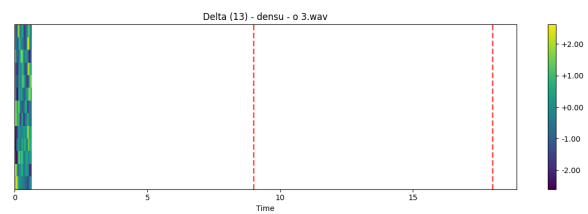
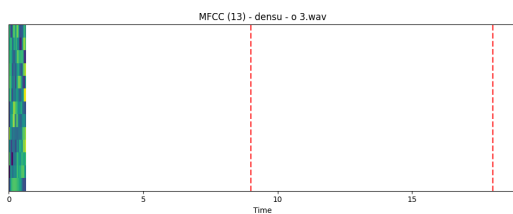
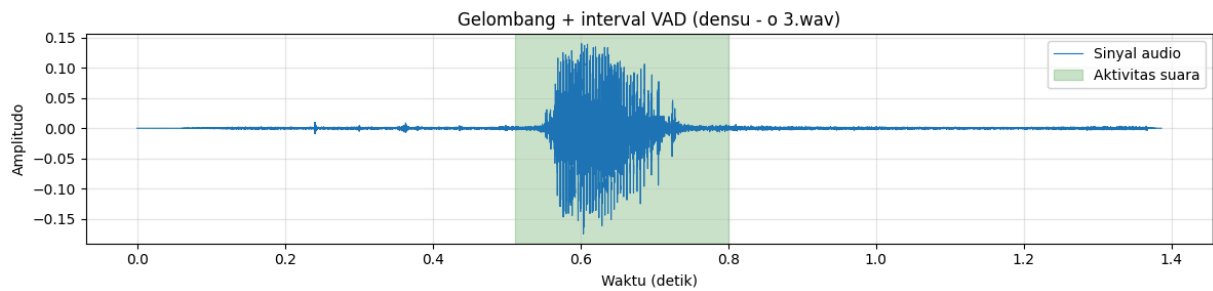


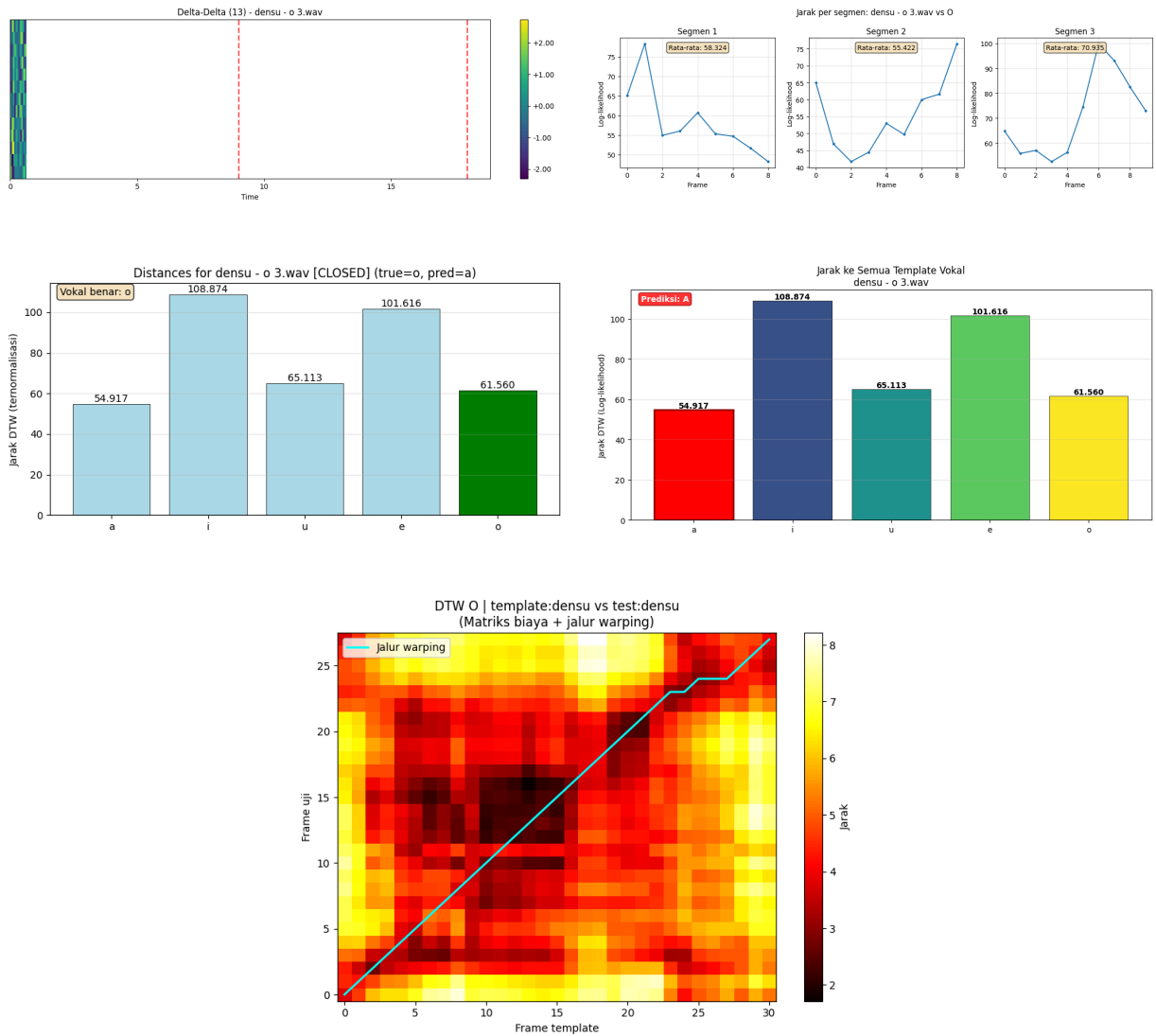
IV.1.6 Hasil Tes Sampel 1 (*Template Kelompok*)



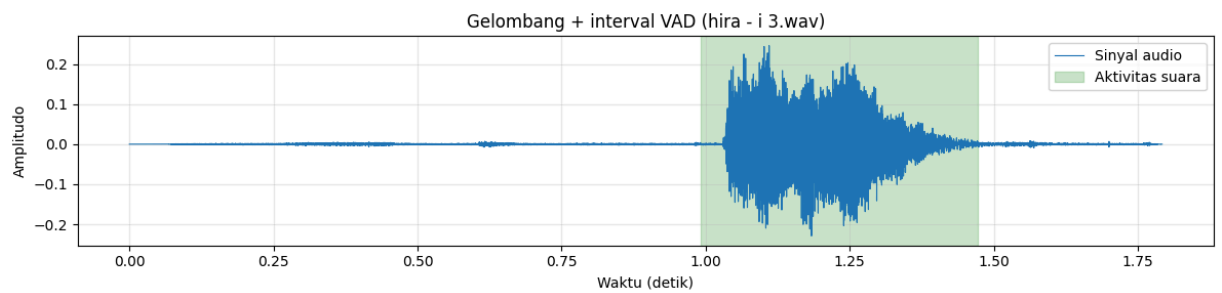


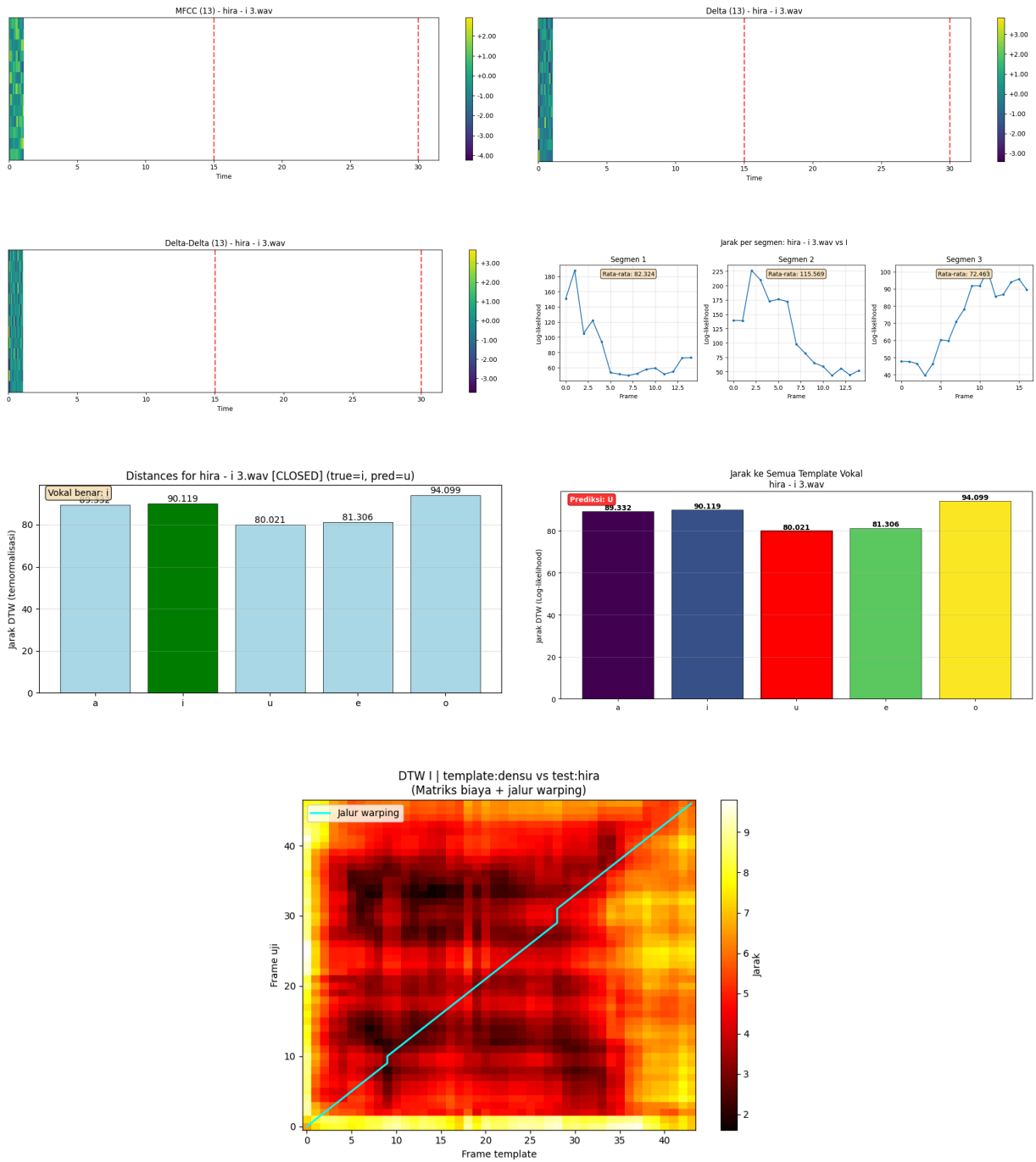
IV.1.7 Hasil Tes Sampel 2 (*Template Kelompok*)



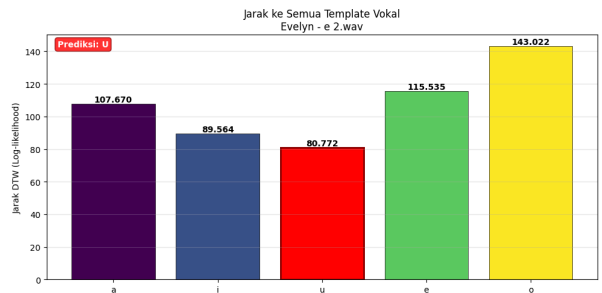
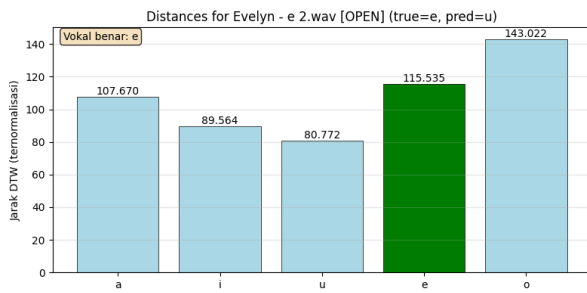
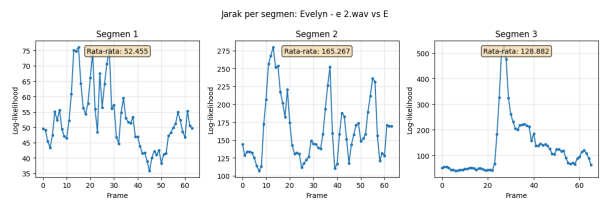
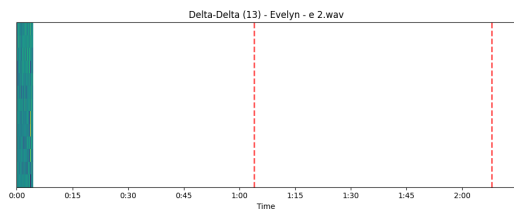
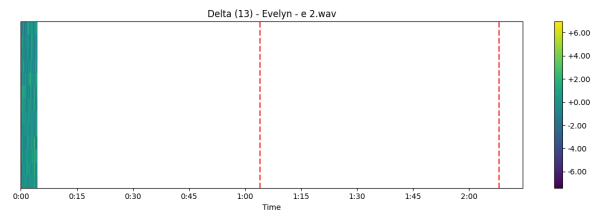
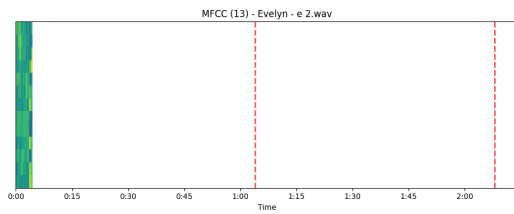
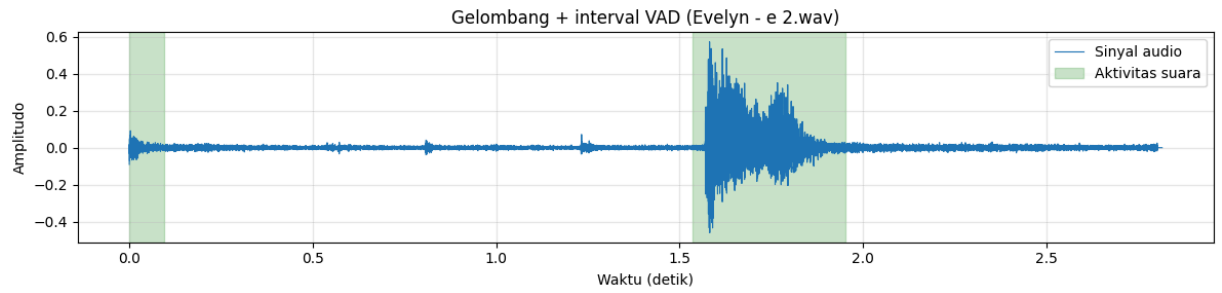


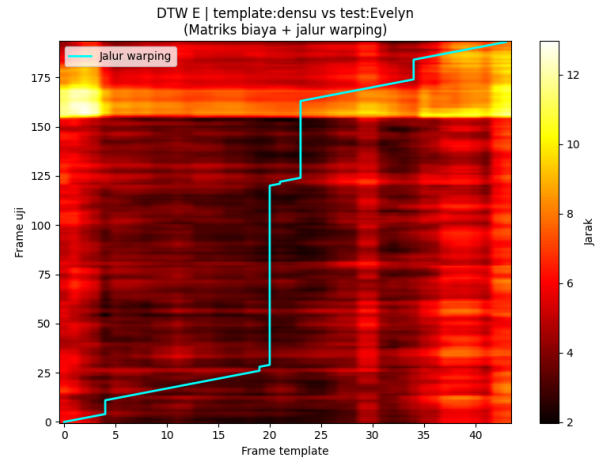
IV.1.8 Hasil Tes Sampel 3 (*Template Kelompok*)



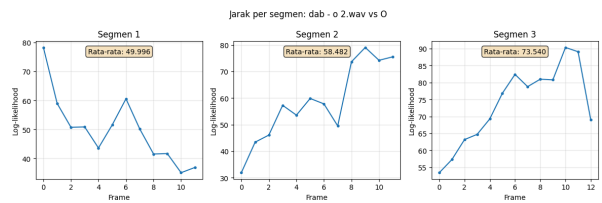
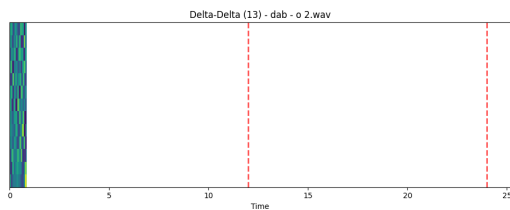
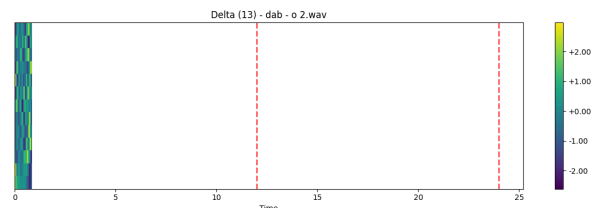
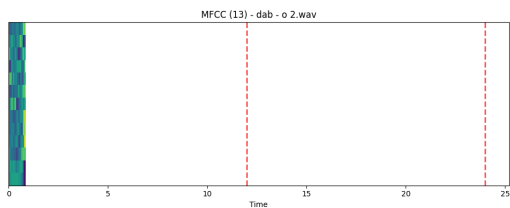
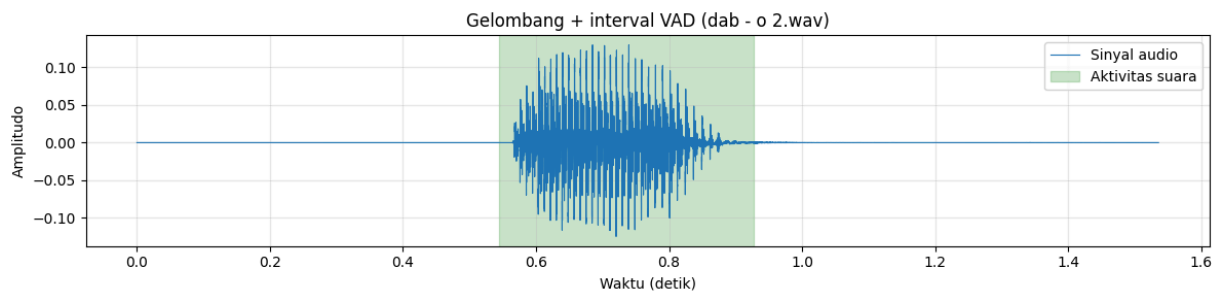


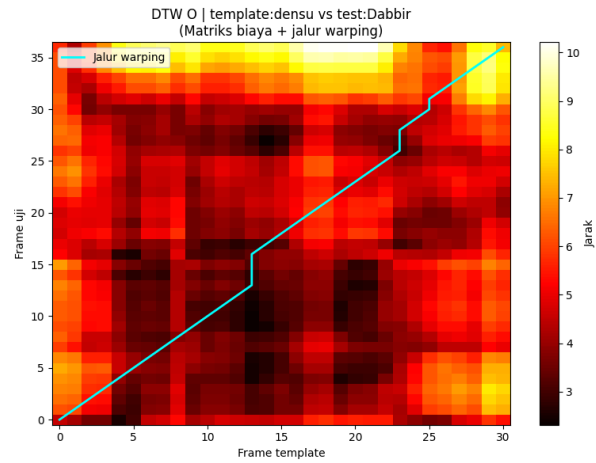
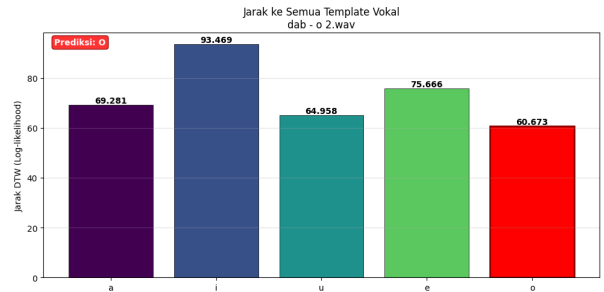
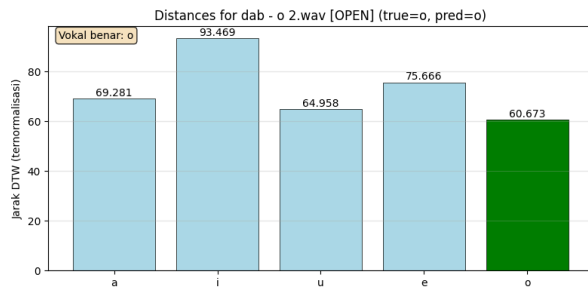
IV.1.9 Hasil Tes Sampel 1 (*Template* Kelompok Lain)



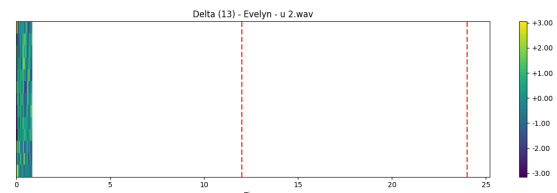
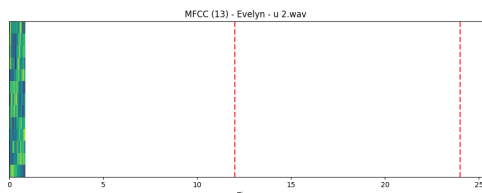
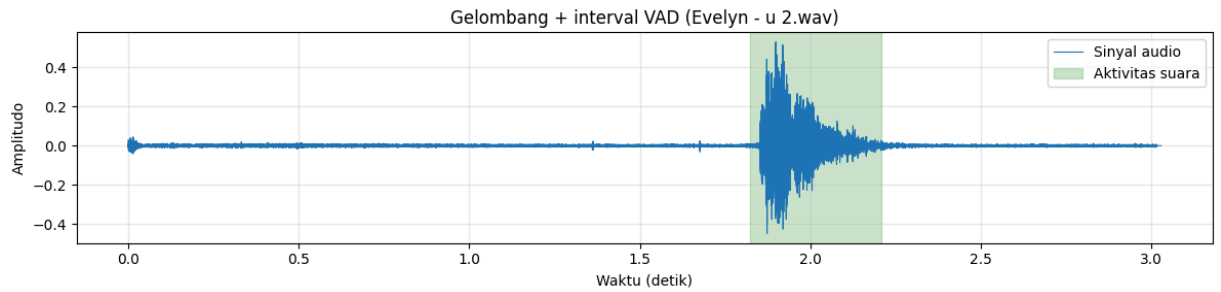


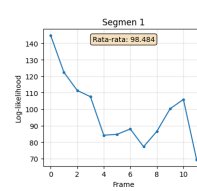
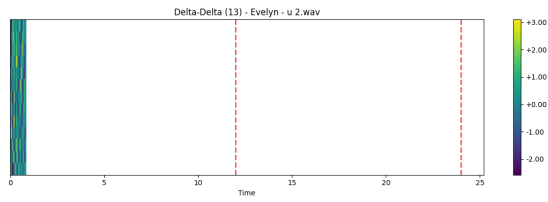
IV.1.10 Hasil Tes Sampel 2 (*Template Kelompok Lain*)



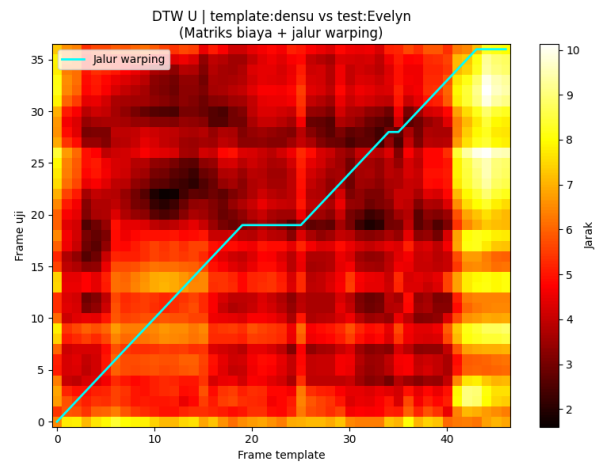
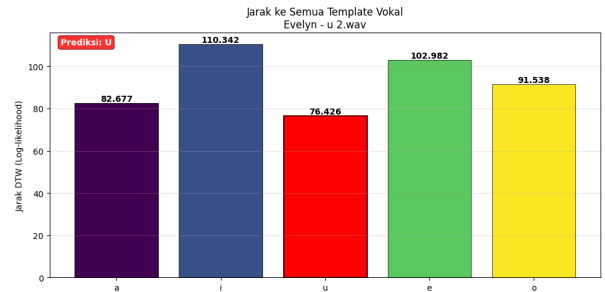
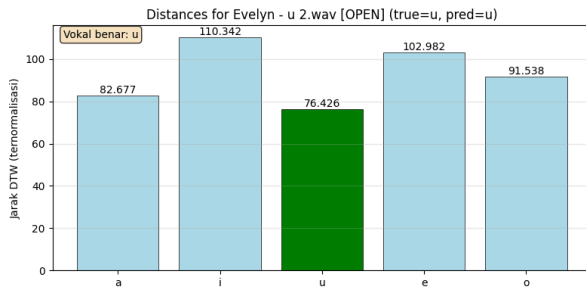
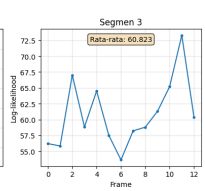
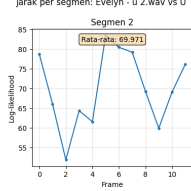


IV.1.11 Hasil Tes Sampel 3 (*Template Kelompok Lain*)

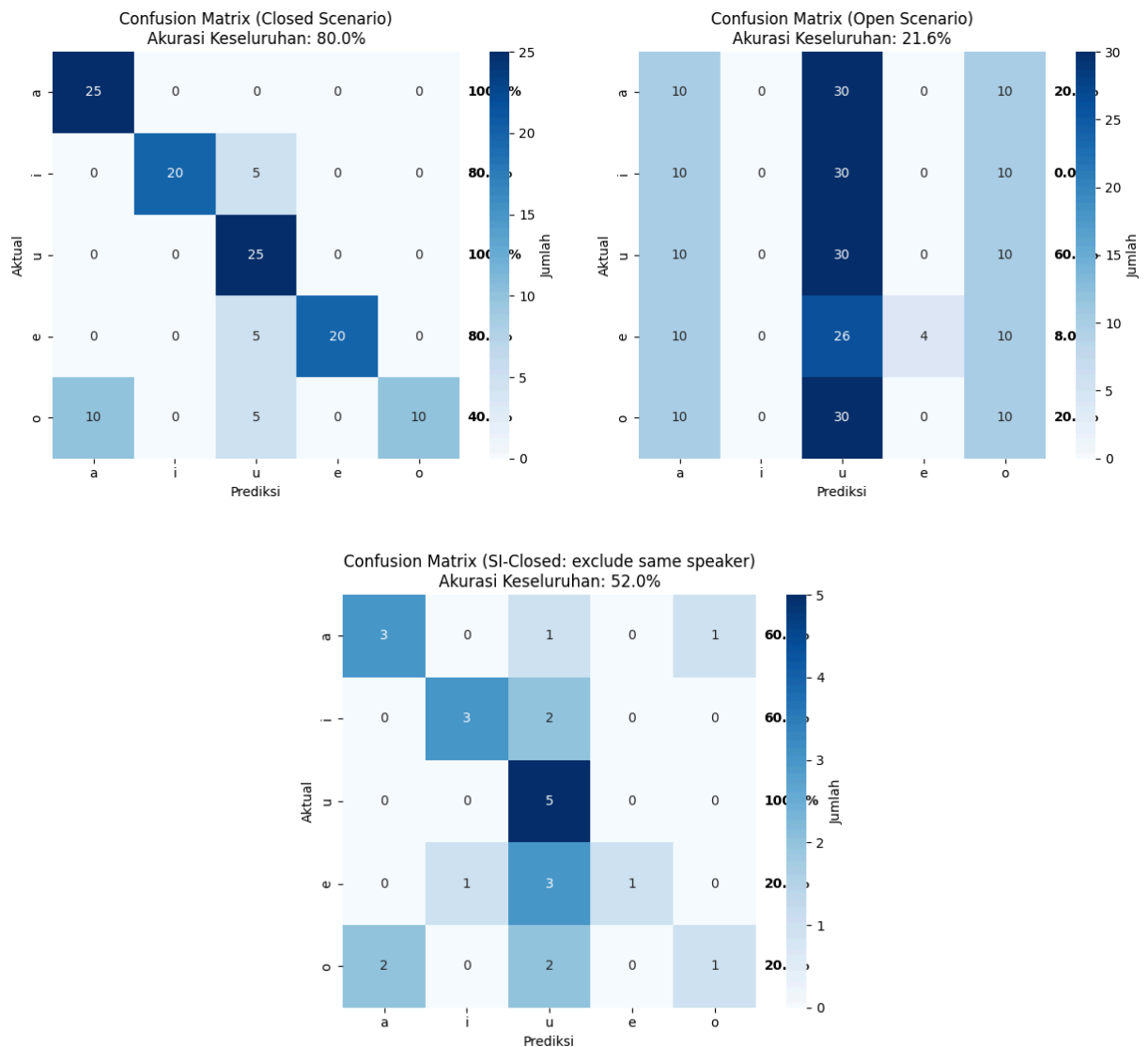




Jarak per segmen: Evelyn - u 2.wav vs U



IV.1.12 Hasil *Confusion Matrix*



IV.2 Analisis

Dari hasil akurasi dari *open* dan *closed scenario*, dapat dilihat bahwa adanya kesenjangan akurasi 60 poin persentase menunjukkan bahwa model sangat bergantung pada karakteristik *speaker* (*speaker-dependent*), dan generalisasi antar penutur masih rendah.

IV.2.1 Pola Fitur MFCC (Vektor Rata-rata)

- Vokal A: Menunjukkan energi kuat pada koefisien MFCC rendah (0-15) dengan pola karakteristik di semua segmen. Evolusi temporal dari segmen 1→3 menunjukkan penurunan amplitudo pada koefisien tinggi.

- Vokal I: Memiliki pola paling khas dengan puncak energi signifikan sekitar koefisien 10-15 pada segmen 1 dan 3. Ini menunjukkan potensi diskriminatif yang baik untuk /i/.
- Vokal U: Menunjukkan distribusi energi konsisten di seluruh segmen dengan aktivitas *notable* pada rentang MFCC tengah (koefisien 8-12). Polanya tampak lebih stabil di seluruh segmen temporal.
- Vokal E: Menunjukkan tingkat energi sedang dengan beberapa kesamaan pada pola /a/, yang mungkin menjelaskan kebingungan antara vokal-vokal ini dalam klasifikasi.
- Vokal O: Menampilkan puncak energi tertinggi pada segmen 1 (koefisien 12-15) tetapi energi lebih rendah pada segmen selanjutnya, menunjukkan peluruhan temporal yang cepat.

IV.2.2 Analisis Matriks Kovarians

- Dominasi Diagonal: Semua matriks kovarians menunjukkan pola diagonal yang kuat, mengindikasikan bahwa koefisien MFCC relatif tidak berkorelasi dalam setiap kelas vokal.
- Varians Tinggi pada Koefisien Rendah: Koefisien 0-15 menunjukkan varians tertinggi (daerah merah), yang mengandung informasi paling diskriminatif.
- Varians Rendah pada Koefisien Tinggi: Fitur delta-delta (koefisien 26-39) menunjukkan varians minimal, menunjukkan bahwa mereka mungkin tidak berkontribusi signifikan pada diskriminasi vokal.
- Kesamaan Antar-Vokal: Struktur kovarians sangat mirip di seluruh vokal, yang menjelaskan generalisasi yang buruk - sistem sulit membedakan antara kelas vokal.

IV.2.3 Analisis *Confusion Matrix*

1. *Closed Scenario* (80% akurasi):
 - Vokal /a/: Klasifikasi sempurna (100% *recall*)
 - Vokal /i/: 80% *recall* - 5 *instance* salah sebagai /u/
 - Vokal /u/: Klasifikasi sempurna (100% *recall*)
 - Vokal /e/: 80% *recall* - 5 *instance* salah sebagai /u/
 - Vokal /o/: 40% *recall* - 10 *instance* salah sebagai /a/, 5 sebagai /u/

Vokal /o/ adalah yang paling bermasalah dengan tingkat kesalahan 60%. Konfusi /o/→/a/ menunjukkan *overlap* dalam ruang fitur akustik.

2. *Open Scenario* (21.6% akurasi):

Bias Ekstrem terhadap /u/:

- Semua vokal mengalami misklasifikasi masif ke /u/ (30 *instance each*)
- Hanya vokal /u/ yang menunjukkan klasifikasi benar signifikan (60% *recall*)
- Vokal /e/: Satu-satunya yang kadang diklasifikasi sebagai /e/ (4 *instance*)

Sistem hampir sepenuhnya gagal mengenali /a/, /i/, dan /o/ pada speaker asing.

3. *SI-Closed* (52% akurasi):

- Vokal /a/ dan /i/: 60% *recall each*
- Vokal /u/: Tetap *perfect* (100% *recall*)
- Vokal /e/ dan /o/: Performa terburuk (20% *recall each*)

Vokal /u/ tetap *robust* bahkan tanpa *template speaker* sendiri, mengindikasikan karakteristik akustik yang lebih universal.

IV.2.4 Analisis Detail dari Visualisasi Per-Sampel

1. Sampel *Template* Kelompok

Contoh Kasus: *densu-a 3.wav*

- *Voice Activity Detection* (VAD)
 - Deteksi yang baik: VAD berhasil mengisolasi bagian vocal (0.4-0.7 detik)
 - Durasi optimal: ~300ms durasi vokal, cukup untuk analisis
 - Clean segmentation: Minimal noise di awal/akhir
- Pola Fitur MFCC
 - Temporal stability: MFCC menunjukkan pola yang relatif stabil di bagian tengah
 - Transisi halus: Tidak ada diskontinuitas yang mencolok
 - Segmentasi bermasalah: Garis merah (*segment boundaries*) membagi secara kasar, tidak mengikuti pola akustik natural
- *DTW Alignment Quality*
 - *Path* suboptimal: Jalur *warping* menunjukkan *alignment* yang tidak *smooth*

- *Cost matrix*: Menunjukkan beberapa *region* dengan *mismatch* tinggi (warna gelap)
 - *Template mismatch*: Beberapa bagian *template* tidak *align* dengan baik ke test sample
 - Jarak Klasifikasi
 - Margin kecil: Jarak ke /a/ (70.2) vs /u/ (76.2) - margin hanya 6 poin
 - *Separability* buruk: Jarak antar kelas terlalu dekat untuk *robust classification*
 - *Correct but risky*: Meskipun benar, margin error sangat tipis
2. Sampel *Template* Kelompok Lain (Perbandingan dengan hasil *template* kelompok)

Contoh Kasus: Evelyn-a 2.wav

- Kualitas Audio & VAD
 - Evelyn: Audio lebih panjang (~2.5 detik vs 1.2 detik), ada *multiple voice activity regions*
 - VAD Detection: Lebih kompleks dengan beberapa segmen terdeteksi, berpotensi *noise/artifacts*
 - Signal quality: Tampak lebih *noisy* dibanding densu yang *clean*
- *Feature Pattern Differences*
 - MFCC *patterns*: Evelyn menunjukkan pola yang sangat berbeda (lebih *sparse* dan *irregular*)
 - Temporal *structure*: Tidak ada *clear steady-state region* seperti pada densu
 - *Normalization issues*: Range nilai MFCC berbeda signifikan
- *DTW Catastrophic Failure*
 - *Cost matrix*: Hampir seluruhnya merah gelap (*high cost*) - menunjukkan *fundamental mismatch*
 - *Warping path*: Sangat suboptimal dengan banyak *vertical/horizontal movements*
- *Distance Scale*
 - Densu (template kelompok): 70-101 *range*
 - Evelyn (template kelompok lain): 83-151 *range*
- Margin

- *Best choice /u/*: 83.172
- *True /a/*: 91.185
- Margin hanya 8 poin yang menandakan bahwa ini *virtually random classification*
- Segmentation
 - *Segment 1*: Relatif *reasonable* (53.936)
 - *Segment 2*: Mulai bermasalah (84.989)
 - *Segment 3*: Total *breakdown* (134.631)

Dari analisis per sampel ditemukan bahwa:

1. Segmentasi *Uniform* Bermasalah

Garis merah menunjukkan pembagian yang tidak mengikuti struktur akustik natural vokal. Vokal memiliki *onset*, *steady-state*, dan *offset* yang bervariasi durasi.

2. *DTW Cost Matrix Suboptimal*

Pattern warping menunjukkan *forced alignment* daripada *natural acoustic correspondence*. Ini mengindikasikan *template* dan test *sample* memiliki karakteristik temporal yang berbeda.

3. *Feature Space Overlap*

Jarak yang terlalu dekat antar kelas (margin <10 poin) menunjukkan *feature space* tidak cukup *discriminative* untuk *robust classification*.

4. *Template Overfitting Evidence*

Fakta bahwa sistem “benar” dengan margin tipis menunjukkan *memorization* daripada *true pattern recognition*. Untuk speaker yang sama, sistem mengandalkan *similarity* artifak daripada *vowel characteristics*.

Analisis ini menjelaskan mengapa sistem gagal di *open scenario*:

- *Template* terlalu spesifik untuk *speaker* tertentu
- Segmentasi tidak mengikuti *phonetic boundaries*
- *Feature extraction* tidak *robust* terhadap variasi *speaker*
- *DTW alignment* tidak dapat menangani *speaker variability*

Visualisasi per-sampel ini mengkonfirmasi bahwa sistem mengandalkan *speaker-specific patterns* daripada *vowel-invariant features*.

IV.2.5 Masalah yang Diidentifikasi

1. Masalah Generalisasi *Template*

- Pemodelan Gaussian dengan hanya 3 segmen mungkin tidak memadai
- Matriks kovarians mungkin diestimasi dengan buruk dengan data terbatas
- Pendekatan segmentasi seragam mungkin tidak menangkap dinamika vokal secara efektif

2. Keterbatasan Representasi Fitur

- MFCC 39 dimensi mungkin tidak menangkap karakteristik *speaker-invariant*
- Kurangnya teknik normalisasi *speaker*
- Hilangnya normalisasi panjang saluran vokal (VTLN)

3. Masalah Metrik Jarak

- Gaussian *log-likelihood* mengasumsikan distribusi normal dari fitur
- Mungkin sensitif terhadap *outlier* dan matriks kovarians yang diestimasi dengan buruk

IV.2.6 Analisis Keterbatasan Dataset dan Bias Gender

- Komposisi Dataset
 - *Training speakers*: 5 orang (densu, hira, naufal, wiga, zya) → mayoritas perempuan
 - *Test speakers*: 10 orang yang berbeda dengan karakteristik vokal bervariasi
 - Ketidakseimbangan gender: *Template training* didominasi suara perempuan
- Implikasi Bias Gender pada Performa
 - Karakteristik Vokal *Gender-Specific*:
 - *Fundamental frequency (F0)*: Perempuan umumnya 150-250 Hz, laki-laki 80-150 Hz
 - *Formant frequencies*: Perempuan cenderung lebih tinggi karena *vocal tract* yang lebih pendek
 - *Spectral characteristics*: Perbedaan energi spektral yang signifikan antar gender
- *Template Overfitting* ke Karakteristik Perempuan
 - Generalized *templates* kemungkinan:

- Mengkodekan F0 *range* yang lebih tinggi
- Menangkap *formant patterns* spesifik perempuan
- Tidak dapat *generalize* ke karakteristik vokal laki-laki
- Bias Sistematis dalam *Open Scenario*
 - Preferensi klasifikasi ke /u/ mungkin terjadi karena:
 - *Template* /u/ perempuan paling "fleksibel" mengakomodasi variasi gender
 - Vokal /u/ memiliki *formant* yang relatif lebih stabil antar gender
 - *Template* lain terlalu *gender-specific* untuk *robust classification*

Keterbatasan Dataset Memperparah Masalah

1. Ukuran Dataset *Insufficient*
 - a. 95 *training files* untuk 5 kelas vokal dan 5 *speakers*
 - b. Hanya ~19 samples per vokal - tidak cukup untuk *robust statistical modeling*
 - c. *Template generalized* dibangun dari data yang sangat terbatas
2. Variabilitas *Speaker* Terbatas
 - a. Hanya 5 *speakers* untuk *training* (tidak representatif populasi)
 - b. *Missing critical demographic variations* (usia, aksen regional, kondisi vokal)
 - c. Tidak ada *stratified sampling* untuk *balanced representation*
3. Konteks Rekaman *Homogen*
 - a. Kemungkinan semua rekaman dalam kondisi *controlled/lab setting*
 - b. *Missing real-world variations* (*noise, emotion, speaking style*)
 - c. Tidak ada variasi prosodik atau *coarticulatory effects*

Sistem menunjukkan performa yang dapat diterima untuk pengenalan vokal bergantung-*speaker* (80% akurasi tertutup) tetapi gagal pada pengenalan *independent-speaker* (21,6% akurasi terbuka). Bias konsisten terhadap vokal "u" dalam *open scenario* dan sifat sistematis kesalahan menunjukkan keterbatasan mendasar dalam pendekatan *template* tergeneralisasi.

Faktor keterbatasan dataset yang signifikan memperparah masalah:

- Bias gender dalam *training set* (mayoritas perempuan)
- Ukuran dataset sangat terbatas (95 *files*, 5 *speakers*)
- Kurangnya variabilitas demografis dan akustik

Visualisasi *template* menunjukkan bahwa model Gaussian tergeneralisasi menangkap informasi diskriminatif yang tidak memadai untuk klasifikasi vokal. Matriks kovarians mengungkap bahwa ruang fitur MFCC 39 dimensi saat ini mengandung redundansi signifikan, sementara vektor rata-rata menunjukkan pemisahan yang tidak memadai antara kelas vokal.

Perbaikan yang paling dibutuhkan adalah:

- Ekspansi dataset dengan *balanced gender representation* dan variabilitas *speaker* yang lebih luas
- Normalisasi *speaker* dan gender dengan teknik VTLN dan *pitch normalization*
- Representasi fitur yang lebih *robust* yang menangkap karakteristik vokal *speaker-invariant*
- Arsitektur *multi-template* yang mengakomodasi variabilitas gender dan demografis

Pendekatan pemodelan Gaussian saat ini dengan segmentasi terbatas dan dataset homogen tidak memadai untuk menangani variabilitas speaker dalam produksi vokal, terutama variabilitas gender yang fundamental dalam karakteristik akustik vokal.

BAB V

KESIMPULAN DAN SARAN

V.1 Kesimpulan

Berdasarkan hasil implementasi dan pengujian sistem pengenalan vokal Bahasa Indonesia, dapat disimpulkan beberapa hal sebagai berikut:

1. **Keberhasilan Implementasi:** Sistem pengenalan vokal Bahasa Indonesia untuk lima vokal (a, i, u, e, o) telah berhasil diimplementasikan menggunakan bahasa Python. Sistem ini menerapkan ekstraksi fitur MFCC 39 dimensi dan metode pencocokan pola *Dynamic Time Warping* (DTW) dengan pendekatan *Generalized Template* berbasis model Gaussian (rata-rata dan kovarians).
2. **Kesenjangan Kinerja Signifikan:** Hasil pengujian menunjukkan kesenjangan kinerja yang sangat besar antara dua skenario. Sistem mencapai akurasi yang relatif baik pada *closed scenario* (pembicara terlihat dalam data latih) sebesar 80.00%. Namun, kinerjanya turun secara drastis pada *open scenario* (pembicara baru yang tidak dikenal) dengan akurasi rata-rata hanya 21.60%.
3. **Ketergantungan pada Pembicara (*Speaker-Dependent*):** Kesenjangan akurasi sebesar hampir 60 poin persentase membuktikan bahwa model yang dibangun bersifat sangat *speaker-dependent*. Model ini mampu mengenali pola vokal dari pembicara yang sudah dilatih, namun gagal total dalam melakukan generalisasi terhadap karakteristik vokal dari pembicara baru.
4. **Bias Model dan Kegagalan Generalisasi:** Kegagalan pada *Open Scenario* didominasi oleh bias ekstrem terhadap vokal /u/, di mana sistem hampir selalu memprediksi /u/ untuk data uji dari pembicara baru, terlepas dari vokal aslinya. Hal ini mengindikasikan bahwa *template* tergeneralisasi untuk /u/ mungkin terlalu "longgar" atau *template* vokal lainnya terlalu spesifik.
5. **Keterbatasan Dataset dan Bias Gender:** Analisis menunjukkan bahwa keterbatasan dataset latih menjadi akar masalah utama. Dengan hanya 5 pembicara (didominasi suara perempuan), *Generalized Template* yang dihasilkan menjadi bias terhadap karakteristik vokal perempuan. Model ini tidak mampu menangani variasi akustik fundamental

antar-gender (seperti perbedaan F0 dan frekuensi formant) yang ada pada data uji *Open Scenario*.

6. **Metodologi Kurang Robust:** Pendekatan *Generalized Template* yang dikombinasikan dengan segmentasi uniform (pembagian fitur menjadi 3 segmen kaku) terbukti tidak memadai. Metode ini gagal menangkap struktur akustik alami vokal (seperti *onset*, *steady-state*, dan *offset*) yang durasinya bervariasi, sehingga memperburuk masalah generalisasi.

V.2 Saran

Untuk mengatasi keterbatasan yang ditemukan dan meningkatkan kinerja sistem, terutama pada skenario independen-pembicara (*speaker-independent*), berikut adalah beberapa saran untuk pengembangan selanjutnya:

1. Pengembangan Dataset:

- **Eksansi Data:** Melakukan penambahan data latih. Jumlah pembicara harus ditingkatkan secara signifikan untuk mencakup variabilitas demografis yang lebih luas (berbagai usia, aksen, dan karakteristik fisiologis).
- **Keseimbangan Gender:** Memastikan keterwakilan gender yang seimbang (*balanced gender representation*) dalam dataset latih. Ini adalah langkah krusial untuk melatih model yang mampu mengenali vokal dari pembicara laki-laki maupun perempuan.

2. Normalisasi dan Ekstraksi Fitur:

- **Normalisasi Pembicara:** Menerapkan teknik normalisasi pembicara sebelum ekstraksi fitur untuk mengurangi variasi antar-pembicara. Metode yang sangat disarankan adalah *Vocal Tract Length Normalization* (VTLN) atau normalisasi berbasis F0 (*pitch normalization*) untuk menstandarkan karakteristik akustik.
- **Analisis Fitur:** Melakukan evaluasi terhadap kontribusi 39 dimensi fitur. Visualisasi kovarians menunjukkan bahwa fitur Delta-Delta (koefisien 26-39) memiliki varians rendah, mengindikasikan kemungkinan redundansi. Eksplorasi fitur yang lebih *speaker-invariant* mungkin diperlukan.

3. Perbaikan Metode Pemodelan:

- **Segmentasi Adaptif:** Mengganti metode segmentasi uniform yang kaku dengan pendekatan yang lebih adaptif, seperti *Segmental K-Means*. Metode ini dapat mengelompokkan frame audio berdasarkan kemiripan akustik alaminya, sehingga menghasilkan segmen yang lebih bermakna secara fonetis.
- **Arsitektur *Multi-Template*:** Daripada satu *Generalized Template* untuk semua pembicara, disarankan menggunakan pendekatan *multi-template*. Misalnya, membuat *template* tergeneralisasi yang terpisah untuk setiap gender (template laki-laki dan template perempuan) untuk menangani variasi vokal antar-gender secara eksplisit.
- **Model Probabilistik Lanjutan:** Mengeksplorasi penggunaan model probabilistik yang lebih canggih seperti *Gaussian Mixture Models* (GMM) penuh atau *Hidden Markov Models* (HMM). Model HMM secara khusus dirancang untuk memodelkan sinyal yang memiliki struktur temporal (seperti ucapan) dan lebih superior daripada DTW dalam menangani variasi durasi.

DAFTAR REFERENSI

- Lestari, D. P. (2025). *IF4071 – Feature Extraction (Mel Frequency Cepstrum Coefficient)* [Presentasi PowerPoint]. Sekolah Teknik Elektro dan Informatika, Institut Teknologi Bandung.
- Lestari, D. P. (2025). *IF4071 – Automatic Speech Recognition: Dynamic Time Warping (DTW)* [Presentasi PowerPoint]. Sekolah Teknik Elektro dan Informatika, Institut Teknologi Bandung.

LAMPIRAN

Source code:

<https://github.com/hiirrs/IF4071-Tubes1>