



Klasifikasi Jenis Suara Wanita Berdasarkan Register Suara Dalam Teknik Bernyanyi Secara Real Time Menggunakan Algoritma BAM dan Algoritma Viterbi

Rini Meiyanti¹⁾, Cut Lika Mestika Sandy²⁾

^{1 2)}Jurusan Teknik Informatika, Fakultas Komputer dan Multimedia,
Universitas Islam Kebangsaan Indonesia
e-mail: rinimeiyanti522@gmail.com

Abstract

[Classification of Female Voices Based on Voice Registers in Real Time Singing Techniques Using the BAM Algorithm and Viterbi Algorithm] In singing technique, women have more diverse voice registers than men. There are many human voice registers, but the voice registers commonly used when singing include Chest Voices, Head Voices, Falsetto, and Vocals. So for the recognition of the type of female voice based on the register of the human voice in singing techniques, a system is needed so that later it can be used for clearer recognition of singing techniques. To get maximum results, it is necessary to perform the two methods that will be used in this research, namely the BAM algorithm and the Viterbi algorithm. Voice recognition is done by inputting recorded sound samples and also in real time. After inputting the voice, the system will get the results according to the method to be used, namely the BAM algorithm and the Viterbi algorithm. The BAM algorithm can process incomplete inputs, because of the reciprocal relationship between the output layer to the input layer. With the Viterbi algorithm, the processing of statuses in the voice recognition system can be optimized. In the BAM algorithm, testing the sound value and the value of the training sample obtained will be searched for the vector value using the weight assessment which is carried out by converting the binary matrix into a bipolar matrix. Meanwhile, in the Viterbi algorithm, the value of the training sound sample will be searched for Viterbi using the Mellin transformation. The Mellin transform is used to convert the time domain to the frequency domain. This research was conducted to produce a system that can consist of female voices based on human voice registers in singing techniques, which shows the performance of the two methods so that we can see which method is more optimal. It is hoped that this research will provide accurate results between the two methods in a sound which will give hope that gives a recruitment decision between the BAM algorithm and the Viterbi algorithm.

Keywords: Voice; BAM; Viterbi; Mellin; Classification.

Abstrak

Dalam teknik bernyanyi, wanita lebih mempunyai register suara yang beragam dibandingkan pria. Banyak sekali register suara manusia, namun register suara yang biasa digunakan saat bernyanyi antara lain yaitu *Chest Voice*, *Head Voice*, *Falsetto*, dan *Vocal fry*. Maka untuk pengenalan jenis suara wanita berdasarkan register suara manusia dalam teknik bernyanyi diperlukan sebuah sistem sehingga nantinya berguna untuk pengenalan teknik-teknik bernyanyi secara lebih jelas. Untuk mendapatkan hasil yang maksimal, maka diperlukan unjuk kerja dua metode yang akan digunakan dalam penelitian kali ini yaitu algoritma BAM dan algoritma *Viterbi*. Pengenalan suara dilakukan dengan menginputkan sampel suara hasil rekaman dan juga secara *real time*. Setelah diinput suara maka sistem akan mendapatkan hasil sesuai dengan metode yang akan digunakan yaitu algoritma BAM dan algoritma *Viterbi*. Algoritma BAM dapat memproses input yang tidak lengkap, karena adanya hubungan timbal balik antara dari lapisan *output* ke lapisan *input*. Dengan algoritma *Viterbi* maka pengolahan status-status dalam sistem pengenalan suara dapat lebih optimal. Pada algoritma BAM, nilai suara pengujian dan nilai sampel suara pelatihan yang diperoleh akan dicari nilai vektornya menggunakan pencarian nilai bobot yang dilakukan dengan cara mengubah matriks biner ke

dalam matriks bipolar. Sedangkan pada algoritma *Viterbi*, nilai sampel suara pelatihan akan dicari pencarian *Viterbi* dengan menggunakan transformasi *Mellin*. Transformasi *Mellin* digunakan untuk merubah domain waktu ke domain frekuensi. Penelitian ini dilakukan untuk menghasilkan sistem yang dapat mengenali jenis suara wanita berdasarkan register suara manusia dalam teknik bernyanyi, yaitu dengan menunjukkan unjuk kerja dari kedua metode sehingga kita dapat mengetahui metode mana yang lebih optimal. Diharapkan dari penelitian ini akan memberikan hasil seberapa akurasi antara kedua metode tersebut dalam mendeteksi suara yang akan dilatih dengan harapan dapat memberikan sebuah keputusan perekomendasi antara algoritma BAM dengan algoritma *Viterbi*.

Kata Kunci: Suara; BAM; *Viterbi*; *Mellin*; Klasifikasi.

1. Pendahuluan

Suara manusia pada dasarnya memiliki register masing-masing yaitu seperti suara asli, suara palsu dan whistle register. Dalam hal bernyanyi, berbagai penyanyi memiliki jenis register suara yang berbeda-beda seperti register suara *falsestto*, *head voice*, *chest voice*, *vocal fry* dan beberapa istilah lainnya. Maka dari itu, supaya tidak membingungkan kita dalam jenis suara menyanyi maka diperlukan sebuah sistem yang dapat mendeteksi jenis suara.

Algoritma *Bidirectional Associative Memory* (BAM) telah banyak digunakan dalam berbagai bidang terutama dalam jaringan syaraf tiruan. BAM mampu mengekstrak bentuk yang sempurna dengan bentuknya yang tidak sempurna. Sistem ini adalah cabang pengolahan suara yang dapat digunakan dalam mengidentifikasi beberapa jenis register suara. Sedangkan Algoritma *Viterbi* adalah algoritma dynamic programming untuk menemukan kemungkinan rangkaian status tersembunyi yang biasa disebut *Viterbi path* dapat dihasilkan pada rangkaian pengamatan kejadian. Dalam algoritma *Viterbi*, dengan menggunakan transformasi *Mellin* yang merupakan transformasi integral yang dapat dianggap sebagai versi perkalian dari dua sisi tranformasi *Laplace*. Dalam hal ini, dapat membantu pembentukan pola dalam mendeteksi pengenalan jenis suara.

Metode algoritma BAM dapat digunakan dalam pengolahan suara dan algoritma *Viterbi* sering digunakan dalam mendeteksi suara. Algoritma BAM dapat memproses input yang tidak lengkap, karena adanya hubungan timbal balik antara dari lapisan output ke lapisan input. Dengan algoritma *Viterbi* maka pengolahan status-status dalam sistem pengenalan suara dapat lebih optimal. Pada algoritma BAM, nilai suara pengujian dan nilai sampel suara pelatihan yang diperoleh akan dicari nilai vektornya menggunakan pencarian nilai bobot yang dilakukan dengan cara mengubah matriks biner ke dalam matriks bipolar. Sedangkan pada algoritma *Viterbi*, nilai sampel suara pelatihan akan dicari pencarian *Viterbi* dengan menggunakan transformasi *Mellin*. Transformasi *Mellin* digunakan untuk merubah domain waktu ke domain frekuensi.

Bustami, Fadlisyah, and Mauliza (2017) melakukan penelitian pada pendeteksi kesalahan dalam membaca Al-Qur'an menggunakan algoritma *Viterbi*. Dalam proses sistem pendeteksi kesalahan membaca Al-Qur'an diperlukan sebuah metode untuk mendapatkan nilai-nilai segmen dalam mendeteksi kesalahan bacaan setiap ayat. Dengan menggunakan metode *Viterbi* proses perhitungan tidak begitu kompleks dan pencarian nilai-nilai segmen dapat dilakukan dengan membandingkan state terbaik. Nilai-nilai tersebut digunakan untuk membandingkan ciri deteksi suara yang benar pada setiap bacaan. Pada proses pelatihan masing-masing bacaan dilatih dengan 10 sample suara kemudian dilakukan proses pengujian untuk memperoleh hasil deteksi berupa segmen. Berdasarkan hasil kompleksitas algoritma, sistem pendeteksi kesalahan dalam membaca Al-Qur'an menggunakan algoritma *Viterbi* adalah efektif. Hasil unjuk kerja sistem mampu mendeteksi kesalahan dalam membaca Al-Qur'an menggunakan algoritma *Viterbi* hingga mencapai 89%.

Bustami., et al (2015), melakukan penelitian mengenai pendeteksi kebohongan pada suara menggunakan algoritma BAM. Penelitian ini melakukan proses pengenalan dan pelatihan sampel suara bohong pada kata pergi, jalan dan pindah lalu diuji dengan mensimulasikannya pada data latihan dan data uji untuk menghasilkan persentase pengenalan dan pengklasifikasian suara kebohongan tersebut. Hasil penelitian menunjukkan bahwa sistem pendeteksian kebohongan menggunakan algoritma BAM memiliki kisaran true detection tertinggi sebesar 90% pada kata pergi.

Berdasarkan permasalahan tersebut, maka perlu dikembangkan suatu aplikasi sistem pengenalan jenis suara wanita berdasarkan register suara dalam teknik yang digunakan ketika sedang menyanyi. Dengan menggunakan beberapa sampel suara teknik tertentu, maka suatu sistem dapat membedakan jenis suara yang dikenali dengan berbagai resister suara sehingga akan mudah digunakan dalam teknik bernyanyi.

2. Metode

A. Pengolahan Suara

Pengenalan suara (juga dikenal sebagai Automatic Speech Recognition (ASR), atau pengenalan suara komputer adalah proses mengkonversi sinyal suara ke urutan kata-kata, melalui algoritma yang diimplementasikan kedalam program komputer (Chamidy, 2016). Pengenalan suara (*speech recognition*) adalah suatu proses untuk mengenali huruf, kata atau kalimat yang diucapkan. Pengenalan suara lebih dikenal dengan istilah *Automatic Speech Recognition* atau *Computer Speech Recognition* dimana penggunaan sebuah mesin/komputer untuk mengenali sebuah suara atau identitas seseorang dari suara yang diucapkan. Umumnya pengucap berbicara di depan komputer/mesin kemudian komputer/mesin mengenali suara/identitas seseorang dengan tepat sesuai yang diucapkan. Pengenalan pola suara dikenali ke dalam berbagai level tugas, pengenalan dalam tingkat sinyal akustik berupa uji tingkatan dalam susunan unit sub kata berupa fonem, kata, frase dan kalimat.

B. Jaringan Syaraf Tiruan

Jaringan saraf tiruan (JST) adalah sistem komputer yang telah dirancang sebagai bagian dari penelitian tentang kecerdasan buatan (AI). Meskipun sejarah Jaringan Syaraf Tiruan berasal dari tahun 1940-an, baru muncul komputer elektronik, dan terutama sampai era ketika komputer tersedia, yang terlihat lonjakan minat terhadap JST (Lesnussa, Latuconsina, & Persulesy, 2015).

JST mempelajari sampel yang disebut training set. Karena mempelajari sampel, JST memiliki potensi membangun sistem komputasi sebagai hasil input dan output memetakan hubungan yang ada dalam sistem (Sudarsono, 2016). Perlengkapan latihan dikenal sebagai pola pelatihan dalam bentuk vektor dan diperoleh dari sumber seperti gambar, sinyal suara, dan berbagai informasi lainnya (Pangaribuan & Sagala, 2017).

C. Bidirectional Associative Memory (BAM)

Jaringan Syaraf Tiruan (JST) Bidirectional Associative Memory (BAM) memiliki kemampuan sebagai associative memory atau content addressable memory yaitu memori yang dapat dipanggil menggunakan bagian dari informasi yang tersimpan di dalamnya (Arifin, Asfani, & Handayani, 2018). Menurut Windarto (2017) selain itu terdapat dua lapisan neuron, yaitu lapisan input dan lapisan output. Tidak ada hubungan lateral, yaitu tidak ada dua neuron di lapisan yang sama terhubung. Sambungan berulang yang merupakan koneksi umpan balik ke neuron itu sendiri, mungkin atau mungkin tidak ada, arsitektur ini cukup sederhana. Sifat kestabilan dan pengkodean jaringan saraf umpan balik nonlinier dua lapis diperiksa. Bidirectionality diperkenalkan di jaringan saraf untuk menghasilkan pencarian asosiatif dua arah untuk asosiasi yang disimpan. Memori asosiatif dua arah (BAM) adalah jaringan umpan balik nonlinier dua lapis.

D. Viterbi

Algoritma *Viterbi* diperkenalkan oleh Andrew J. *Viterbi* pada tahun 1967. Algoritma *Viterbi* adalah algoritma dynamic programming untuk menemukan kemungkinan rangkaian status yang tersembunyi (biasa disebut *Viterbi path*) yang dihasilkan pada rangkaian pengamatan kejadian (Hidayat, 2015). Untuk menemukan sebuah rangkaian status terbaik, $q = (q_1, q_2, \dots, q_r)$, untuk rangkaian observasi $O = (O_1, O_2, \dots, O_r)$, perlu didefinisikan kuantitas:

$$\delta_t(i) = \max_{q_1, q_2, \dots, q_{t-1}} P[q_1, q_2, \dots, q_{t-1}, q_t = i, o_1, q_1, q_2, \dots, q_{t-1}, o_2, \dots, o_t | \lambda] \quad (1)$$

$\delta_t(i)$ = rangkaian terbaik

t = waktu

i = status i

Dengan menginduksi, didapat:

$$\delta_{t+1}(j) = [\max_i \delta_t(i) a_{ij}] \cdot b_j(o_{t+1})$$

E. Transformasi Mellin

Dalam matematika, Transformasi *Mellin* adalah transformasi integral yang dapat dianggap sebagai versi perkalian dari dua sisi transformasi Laplace (Fadlisyah, Fuadi, & Kamilaini, 2019). integral tranformasi ini adalah berkaitan erat dengan teori seni Dirichlet dan sering digunakan di nomor teori dan teori ekspansi asimtotik, itu berkaitan erat dengan Trasformasi *Laplace*, Tranformasi Fouries, teori fungsi gamma dan fungsi khusus bersekutu.

Transformasi *Mellin* adalah :

$$\varphi(s) = \sum_{x=0}^{n-1} x^{s-1} f(x) \quad (2)$$

F. Data Yang Digunakan

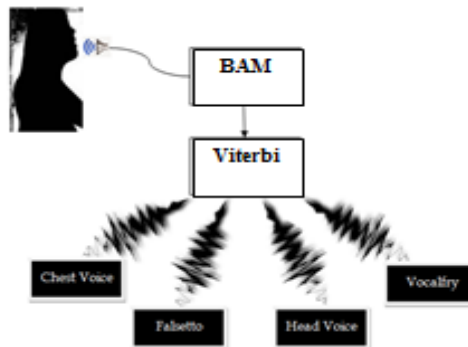
Setelah merencanakan penelitian, yang dilakukan terlebih dahulu adalah mengumpulkan referensi tentang Pengolahan Sinyal dan Suara serta data yang dibutuhkan dalam pembuatan aplikasi. Data yang digunakan dalam penelitian ini terbagi dua, yaitu data secara langsung (*realtime*), dan data tidak langsung yang berupa sampel rekaman suara wanita dengan 4 register dalam teknik bernyanyi.

G. Analisa Data

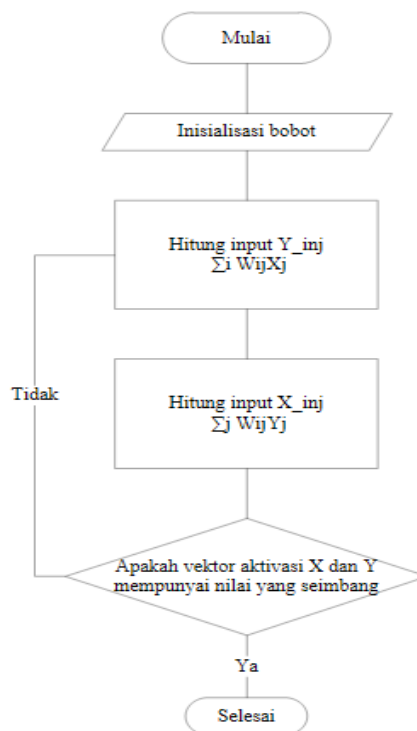
Penelitian yang dilakukan terlebih dahulu adalah mengumpulkan referensi tentang Pengolahan Sinyal dan Suara serta data yang dibutuhkan dalam pembuatan aplikasi. Data yang digunakan dalam penelitian ini terbagi dua, yaitu data secara langsung (*realtime*), dan data tidak langsung yang berupa sampel rekaman suara wanita dengan 4 register dalam teknik bernyanyi. Dengan rincian sebagai berikut:

1. Rincian Sampel Suara Rekaman (*Record*), yaitu mengumpulkan sampel suara wanita yang sudah direkam dengan 4 register suara dalam teknik bernyanyi yaitu chest voice, falsetto, head voice dan vocal fry.
2. Rincian Sampel Suara Langsung (*Real Time*), Untuk sampel suara realtime dilafalkan secara langsung oleh beberapa responden dengan bantuan microphone. Sampel suara realtime ini digunakan sebagai perbandingan untuk melihat keakuratan sistem antara proses pengenalan suara secara langsung dengan proses pengenalan suara yang diambil dari sampel rekaman.

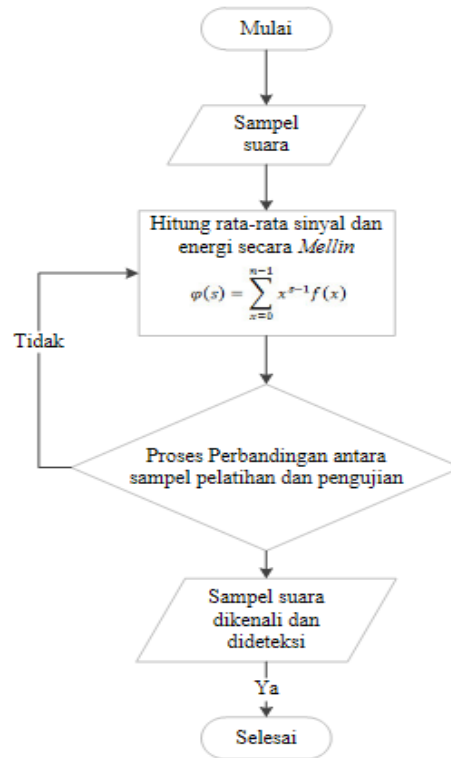
H. Skema Sistem Secara Keseluruhan



Gambar 1. Skema sistem secara keseluruhan.



Gambar 2. Diagram Metode Algoritma Bidirectional Associative Memory (BAM).



Gambar 3. Diagram Metode Algoritma Viterbi

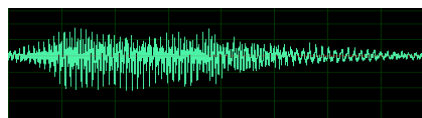
Pada tahap ini, jaringan menerima inputan struktur sampel suara. Struktur suara inputan merupakan sebarisan nilai sinyal suara tepi-tepi objek yang akan mendeteksi jenis suara, dan selanjutnya jaringan melakukan tahapan pelatihan yaitu pengenalan sinyal suara dan tandai ciri, dan uji pengenalan suara melalui algoritma *Viterbi*. Pada proses utama, komputasi menggunakan transformasi *Mellin*, sampel suara akan dilatih untuk mendapatkan sebuah nilai energi dan nilai segmen.

Adapun pelaksanaan yang akan dilakukan dalam membangun sistem ini adalah dengan menggunakan suara wanita dengan ekstensi file *.wav* sebagai suara yang di input dalam program. Pada proses pengumpulan data (sampel suara) dilakukan dengan langkah-langkah sebagai berikut :

1. Mengumpulkan responden dari keempat register suara yaitu *Chest Voice*, *Falsetto*, *Head Voice*, dan *Vocal fry*.
2. Merekam suara responden yang merupakan contoh suara dengan teknik bernyanyi yang dilafalkan oleh masing-masing responden dari tiap-tiap register suara.
3. Durasi rekaman maksimal 10 detik menggunakan perekam catatan suara dengan suara berformat *.amr*.
4. Merubah (*convert*) file suara sampel rekaman dari format *amr/midi* menjadi format *.wav*.
5. Mengklarifikasi jenis register suara lalu membagi sampel menjadi 2 jenis sampel, yaitu : **sampel pelatihan** dan **sampel pengujian**.
6. Untuk proses pengenalan suara secara *real time*, sampel suara langsung diuji dari beberapa responden dengan aplikasi yang sudah berjalan.

3. Hasil dan Pembahasan

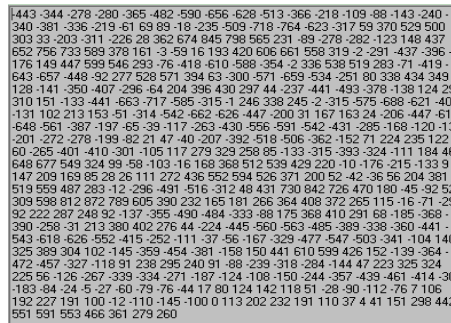
Sampel training dan testing berupa sampel suara secara real-time dan record. Data training yang digunakan pada penelitian ini berupa sampling suara wanita pada jenis *head voice*, *falsetto*, *chest voice* dan *vocal fry*.



Gambar 4. Mengilustrasikan gambar sinyal suara pada *falsetto*

Setiap suara memiliki sampling suara yang spesifik atau berbeda antara satu suara dengan suara lainnya. Sampling suara ialah nilai sinyal suara asli berupa nilai frekuensi dari suara. Dimana gambar dibawah ini

menunjukkan ciri-ciri nilai sampling suara wanita pada jenis *head voice*, *false*to, *chest voice* dan *vocal fry*. Setelah nilai sampling suara diketahui barulah kita dapat memproses suara tersebut untuk mendapatkan informasi yang lebih spesifik sebagai acuan penelitian.



Gambar 5. Nilai sampling suara wanita pada jenis *false*to

Pada gambar diatas menunjukkan nilai sampling suara wanita pada jenis *false*to. Itu merupakan salah satu sampel yang digunakan untuk sinyal suara pelatihan pada *false*to. Untuk suara wanita jenis *false*to memiliki banyak ciri vektor dari nilai sinyal suara. Untuk menentukan normalisasi sinyal suara yaitu jika nilai $f > 0$ maka $f = 1$, jika $f < 0$ maka $f = -1$ dan jika $f = 0$ maka $f = 0$.

Pelatihan dilakukan dengan menginputkan sampel dari berbagai macam jenis suara wanita berdasarkan register suara dalam teknik bernyanyi dengan jumlah sampel pengujian sebanyak 100 sampel suara untuk proses pengenalan suara secara *record*, dan 50 sampel suara pengujian untuk proses pengenalan suara secara *real time*. Hal ini dimaksudkan untuk melihat sejauh mana tingkat keberhasilan saat melakukan pengujian pada masing-masing proses pengenalan suara. Berikut adalah contoh proses pelatihan yang dilakukan pada setiap jenis register suara, baik pada proses pengenalan suara secara *real time* maupun proses pengenalan suara *record* :



Gambar 6. Proses Pengujian Jenis Register Suara *False*to Secara *Real Time*



Gambar 7. Proses Pengujian Jenis Register Suara *False*to Secara *Record*

Pada gambar di atas, *user* kembali menginput sampel suara, kali ini jenis register suara *false*to. Selanjutnya program akan melakukan proses pengenalan yang sama dengan cara menghitung nilai energi yang ada pada sampel suara tersebut, dan kemudian muncul jenis register suara sebagai hasil akhir. Hasil eksekusi beberapa sampel dalam bentuk tabel sebagai data statistik yang lebih spesifik dan detail yang dapat menggambarkan sejauh mana keefektifan aplikasi dan persentase atau rasio kebenaran dari program

yang telah dibuat juga untuk mendapatkan bukti dan hasil yang lebih nyata tentang pengenalan jenis register suara.

Tabel berikut menggabungkan empat jenis register suara sekaligus dalam satu tabel, dengan pelatihan sebanyak satu kali untuk masing-masing jenis register dan menggunakan dua sampel pelatihan yang berbeda dengan sampel suara pengujian yang sama sebanyak 100 sampel suara *record* dan 50 sampel suara *real time* dengan melihat unjuk kerja dari algoritma BAM dan algoritma Viterbi. Untuk lebih detailnya dapat kita lihat pada tabel berikut :

Tabel 1. Unjuk Kerja Sistem Untuk Proses Pengenalan Suara *Record* pada BAM

Sampel Suara Pelatihan	Jumlah Pengujian				Identifikasi Benar						Identifikasi Salah					
	H	F	C	V	H	F	C	V	Tot	%	H	F	C	V	Tot	%
Latih 1	100	100	100	100	87	85	79	78	329	78,3	13	15	21	22	71	21,7
Latih 2	100	100	100	100	73	76	87	84	320	74,2	27	24	13	16	80	25,8

Tabel 2. Unjuk Kerja Sistem Untuk Proses Pengenalan Suara *Record* pada Viterbi

Sampel Suara Pelatihan	Jumlah Pengujian				Identifikasi Benar						Identifikasi Salah					
	H	F	C	V	H	F	C	V	Tot	%	H	F	C	V	Tot	%
Latih 1	100	100	100	100	43	44	32	45	164	77,8	7	6	18	5	36	22,2
Latih 2	100	100	100	100	43	42	47	45	177	82,3	7	8	3	5	23	11,7

Tabel 3. Unjuk Kerja Sistem Untuk Proses Pengenalan Suara *Real Time* pada BAM

Sampel Suara Pelatihan	Jumlah Pengujian				Identifikasi Benar						Identifikasi Salah					
	H	F	C	V	H	F	C	V	Tot	%	H	F	C	V	Tot	%
Latih 1	50	50	50	50	78	88	86	87	339	80,2	22	12	14	13	61	19,8
Latih 2	50	50	50	50	65	85	79	84	313	73,1	35	15	21	16	87	26,9

Tabel 4. Unjuk Kerja Sistem Untuk Proses Pengenalan Suara *Real Time* pada Viterbi

Sampel Suara Pelatihan	Jumlah Pengujian				Identifikasi Benar						Identifikasi Salah					
	H	F	C	V	H	F	C	V	Tot	%	H	F	C	V	Tot	%
Latih 1	50	50	50	50	46	38	49	41	164	77,8	4	12	1	9	26	22,2
Latih 2	50	50	50	50	49	46	42	47	184	86,5	1	4	8	3	16	13,5

Keterangan : H = *Head Voice*, F = *Falsetto*, C = *Chest Voice*, V = *Vocalfry*

Tabel tersebut menunjukkan bahwa dari hasil eksekusi program akan didapatkan persentase keberhasilan pengenalan untuk jenis register suara dari beberapa sampel yang diinput,. Kolom ‘Sampel Suara Pelatihan’ memuat beberapa sampel suara yang berbeda yang dijadikan sebagai pelatihan. Kolom ‘Jumlah Pengujian’ pun demikian, berisi jumlah pengujian yang dilakukan *user*, kemudian ada kolom ‘Identifikasi Benar’, yang berisi data tentang berapa jumlah pengenalan suara yang berhasil disertai dengan persentasenya. Terakhir yaitu kolom ‘Identifikasi Salah’, kebalikan dari kolom ‘Identifikasi Benar’, yang berisi data tentang jumlah pengenalan suara yang tidak berhasil beserta persentasenya. Pada metode *Bidirectional Associative Memory* (BAM) pencarian nilai bobot dilakukan dengan cara mengubah matriks biner ke dalam matriks bipolar, yang selanjutnya akan dilakukan pencarian bobot. Contoh pencarian bobot pada algoritma *Bidirectional Associative Memory* (BAM) yang diambil dari sembarang vektor fitur.

$$x_1 = (1,0,0,1) \quad y_1 = (0,1,1)$$

$$x_2 = (0,1,1,0) \quad y_2 = (1,0,1)$$

Matriks biner tersebut selanjutnya diubah kedalam matriks bipolar menjadi:

$$\begin{aligned} \text{Cari nilai bipolar } 0 &\xrightarrow{-1} \rightarrow \\ (x1)^+ &= (1, -1, -1, 1) \quad (y1)^+ = (-1, 1, 1, 1) \\ (x2)^- &= (-1, 1, 1, -1) \quad (y2)^- = (1, -1, 1, 1) \end{aligned}$$

Maka untuk pencarian bobot dilakukan dengan menggunakan rumus $W = \sum_{i=1} x_i y_i$

$$\begin{aligned} W &= \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix} \begin{bmatrix} -1 & 1 & 1 \end{bmatrix} + \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix} \begin{bmatrix} -1 & 1 & 1 \end{bmatrix} \\ W &= \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & -1 \\ 1 & -1 & -1 \\ -1 & 1 & 1 \end{bmatrix} + \begin{bmatrix} -1 & 1 & 1 \\ 1 & -1 & -1 \\ 1 & -1 & -1 \\ -1 & 1 & 1 \end{bmatrix} \\ W &= \begin{bmatrix} -2 & 2 & 0 \\ 2 & -2 & 0 \\ 2 & -2 & 0 \\ -2 & 2 & 0 \end{bmatrix} \quad \text{dan } W^T = \begin{bmatrix} -2 & 2 & 2 & -2 \\ 2 & -2 & -2 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix} \end{aligned}$$

Pencarian algoritma *Viterbi* dalam sistem ini menggunakan transformasi *Mellin*. Contoh penjabaran rumus transformasi *Mellin* dan *Viterbi* untuk perhitungan manual yang diterapkan pada sistem ini : Rumus *Mellin*

$$\varphi(s) = \sum_{x=0}^{n-1} x^{s-1} f(x)$$

Mencari energi kata pergi

$$\begin{aligned} \text{Dik : } x &= 0, 1, 2, 3 \\ S &= 0, 1, 2, 3 \\ N &= 4 \\ f(x) &= 0, 2, 0, 0 \end{aligned}$$

Penyelesaian :

$$\begin{aligned} \varphi(0) &= (0^{0-1} \cdot 0) + (0^{0-1} \cdot 1) + (0^{0-1} \cdot 2) + (0^{0-1} \cdot 3) \\ \varphi(0) &= 0 + 2 + 0 + 0 = 2 \\ \varphi(1) &= (0^{1-1} \cdot 0) + (0^{1-1} \cdot 1) + (0^{1-1} \cdot 2) + (0^{1-1} \cdot 3) \\ \varphi(1) &= 0 + 2 + 0 + 0 = 2 \\ \varphi(2) &= (0^{2-1} \cdot 0) + (0^{2-1} \cdot 1) + (0^{2-1} \cdot 2) + (0^{2-1} \cdot 3) \\ \varphi(2) &= 0 + 2 + 0 + 0 = 2 \\ \varphi(3) &= (0^{3-1} \cdot 0) + (0^{3-1} \cdot 1) + (0^{3-1} \cdot 2) + (0^{3-1} \cdot 3) \\ \varphi(3) &= 0 + 2 + 0 + 0 = 2 \end{aligned}$$

Energi Mellin Persegmen

$$\begin{aligned} \varphi &= \frac{\varphi(0) + \varphi(1) + \varphi(2) + \varphi(3)}{N^2} \\ \varphi &= \frac{\varphi(2) + \varphi(2) + \varphi(2) + \varphi(2)}{4^2} = \frac{8}{16} = 0,5 \end{aligned}$$

Mencari energi kata jalan

$$\begin{aligned} \text{Dik : } x &= 0, 1, 2, 3 \\ S &= 0, 1, 2, 3 \\ N &= 4 \\ f(x) &= 0, 1, 1, 0 \end{aligned}$$

Penyelesaian :

$$\begin{aligned} \varphi(0) &= (0^{0-1} \cdot 0) + (0^{0-1} \cdot 1) + (0^{0-1} \cdot 2) + (0^{0-1} \cdot 3) \\ \varphi(0) &= 0 + 1 + 0 + 0 = 1,5 \\ \varphi(1) &= (0^{1-1} \cdot 0) + (0^{1-1} \cdot 1) + (0^{1-1} \cdot 2) + (0^{1-1} \cdot 3) \\ \varphi(1) &= 0 + 1 + 1 + 0 = 2 \\ \varphi(2) &= (0^{2-1} \cdot 0) + (0^{2-1} \cdot 1) + (0^{2-1} \cdot 2) + (0^{2-1} \cdot 3) \\ \varphi(2) &= 0 + 1 + 2 + 0 = 3 \end{aligned}$$

$$\begin{aligned}\varphi(3) &= (0^{3-1} \cdot 0) + (0^{3-1} \cdot 1) + (0^{3-1} \cdot 2) + (0^{3-1} \cdot 3) \\ \varphi(3) &= 0 + 1 + 4 + 0 = 5\end{aligned}$$

Energi Mellin Persegmen

$$\begin{aligned}\varphi &= \frac{\varphi(0) + \varphi(1) + \varphi(2) + \varphi(3)}{N^2} \\ \varphi &= \frac{\varphi(1,5) + \varphi(2) + \varphi(3) + \varphi(5)}{4^2} = \frac{11,5}{16} = 0,7\end{aligned}$$

Mencari energi kata pindah

$$\begin{aligned}\text{Dik : } x &= 0,1,2,3 \\ S &= 0,1,2,3 \\ N &= 4 \\ f(x) &= 2,0,0,1\end{aligned}$$

Penyelesaian :

$$\begin{aligned}\varphi(0) &= (0^{0-1} \cdot 0) + (0^{0-1} \cdot 1) + (0^{0-1} \cdot 2) + (0^{0-1} \cdot 3) \\ \varphi(0) &= 0 + 0 + 0 + 0,33 = 0,33 \\ \varphi(1) &= (0^{1-1} \cdot 0) + (0^{1-1} \cdot 1) + (0^{1-1} \cdot 2) + (0^{1-1} \cdot 3) \\ \varphi(1) &= 2 + 0 + 0 + 1 = 3 \\ \varphi(2) &= (0^{2-1} \cdot 0) + (0^{2-1} \cdot 1) + (0^{2-1} \cdot 2) + (0^{2-1} \cdot 3) \\ \varphi(2) &= 0 + 0 + 0 + 3 = 3 \\ \varphi(3) &= (0^{3-1} \cdot 0) + (0^{3-1} \cdot 1) + (0^{3-1} \cdot 2) + (0^{3-1} \cdot 3) \\ \varphi(3) &= 0 + 0 + 0 + 9 = 9\end{aligned}$$

Energi Mellin Persegmen

$$\begin{aligned}\varphi &= \frac{\varphi(0) + \varphi(1) + \varphi(2) + \varphi(3)}{N^2} \\ \varphi &= \frac{\varphi(0,33) + \varphi(3) + \varphi(3) + \varphi(9)}{4^2} = \frac{15,33}{16} = 0,9\end{aligned}$$

4. Kesimpulan

Dari hasil penelitian yang dilakukan terhadap beberapa sampel suara, hasil pengenalan suara yang dilakukan secara realtime rata-rata memiliki persentase lebih tinggi daripada pengenalan suara dengan sampel hasil record, walaupun perbedaannya tidak terlalu signifikan. Hal ini disebabkan karena suara yang diinput secara real time lebih jernih dan terdengar lebih jelas daripada suara hasil rekaman. Alat perekam suara juga ikut mempengaruhi kinerja sistem dalam proses pengenalan suara. Tetapi tidak bisa disimpulkan bahwa hasil pengenalan suara secara realtime sepenuhnya lebih akurat dibandingkan pengenalan suara dengan hasil rekaman. Ada beberapa hal yang mempengaruhinya, antara lain :

1. Ketepatan teknik suara dari responden yang dijadikan sampel. Tidak semua orang memiliki kemampuan teknik bernyanyi yang tepat dan jenis suara dalam berbagai register.
2. Program hanya dapat mengenali suara berdasarkan perbandingan nilai energi antara proses pelatihan dan proses pengujian. Ini merupakan salah satu kelemahan dari sistem ini. Program tidak bisa mendeteksi langsung asal sumber suara yang dilafalkan, karena register suara sesungguhnya tidak hanya dilihat dari tinggi rendahnya nada yang dihasilkan, tetapi juga dilihat dari sumber suara, sensasi ruang resonansi serta warna suara.
3. Berdasarkan kinerja sistem, proses pengenalan suara secara real time lebih efektif dibandingkan proses pengenalan suara secara record. Untuk proses pengenalan suara secara real time, sampel suara yang diinput memiliki durasi yang lebih panjang, sedangkan untuk proses pengenalan suara secara record sampel suara yang diinput berdurasi pendek, hanya beberapa detik saja. Hal ini dipengaruhi oleh memori penyimpanan. Jika diinputkan suara dalam durasi yang panjang, memori penyimpanan tidak sanggup menampung data dalam kapasitas yang besar sehingga terjadi floating point overflow.
4. Dari unjuk kerja kedua algoritma didapatkan hasil bahwa detection rate dengan menggunakan algoritma Viterbi memiliki hasil yang lebih optimal dibandingkan dengan algoritma BAM.
5. Berdasarkan data dan fakta yang muncul pada tabel dan beberapa penjabaran diatas menunjukkan bahwa sampel suara yang dijadikan pelatihan sangat berdampak terhadap tingginya persentase keberhasilan pengenalan suara dan jenis metode yang digunakan.
6. Dari hasil penelitian dapat disimpulkan bahwa persentase rata-rata keberhasilan pengenalan register suara secara *record* yang optimal diiperoleh dengan menggunakan algoritma *Viterbi* yaitu 80,2% sedangkan untuk register suara secara *real time* yang optimal juga diperoleh dengan menggunakan algoritma *Viterbi* yaitu 86,5%.

Daftar Pustaka

- Arifin, M., Asfani, K., & Handayani, A. N. (2018). Aplikasi Jaringan Saraf Tiruan Metode Perceptron Pada Pengenalan Pola Notasi. *Simetris: Jurnal Teknik Mesin, Elektro dan Ilmu Komputer*, 9(1), 77-86.
- Bustami, B., Fadlisyah, F., & Mauliza, I. (2017). SISTEM PENDETEKSI KESALAHAN DALAM MEMBACA AL-QUR'AN AYAT 1-5 MENGGUNAKAN METODE VITERBI. *TECHSI-Jurnal Teknik Informatika*, 9(1), 30-44.
- Chamidy, T. (2016). Metode Mel Frequency Cepstral Coeffisients (MFCC) Pada klasifikasi Hidden Markov Model (HMM) Untuk Kata Arabic pada Penutur Indonesia. *Matics*, 8(1), 36-39.
- Fadlisyah, F., Fuadi, W., & Kamilaini, K. (2019). SISTEM WAKTU NYATA PENGUJIAN HAFALAN AYAT-AYAT SUCI AL QURAN SECARA EKSPONENSIAL DAN NONSINUSOIDAL. *TECHSI-Jurnal Teknik Informatika*, 11(3), 457-567.
- Hidayat, I. (2015). *Desain dan Analisis Pendekode Viterbi Menggunakan Satu Butterfly Berbasis Bahasa Vhdl*. Paper presented at the Seminar Nasional Informatika (SEMNASIF).
- Lesnussa, Y., Latuconsina, S., & Persulesy, E. (2015). Aplikasi Jaringan Saraf Tiruan Backpropagation untuk Memprediksi Prestasi Siswa SMA (Studi kasus: Prediksi Prestasi Siswa SMAN 4 Ambon). *Jurnal Matematika Integratif, ISSN*, 1412-6184.
- Pangaribuan, Y., & Sagala, M. (2017). Menerapkan Jaringan Saraf Tiruan untuk Mengenali Pola Huruf Menggunakan Metode Perceptron. *Jurnal Teknik Informatika UNIKA Santo Thomas*, 2(2), 53-59.
- Sudarsono, A. (2016). Jaringan Syaraf Tiruan Untuk Memprediksi Laju Pertumbuhan Penduduk Menggunakan Metode Bacpropagation (Studi Kasus Di Kota Bengkulu). *Jurnal Media Infotama*, 12(1).
- Windarto, A. P. (2017). Implementasi Jst Dalam Menentukan Kelayakan Nasabah Pinjaman Kur Pada Bank Mandiri Mikro Serbelawan Dengan Metode Backpropogation. *J-SAKTI (Jurnal Sains Komputer dan Informatika)*, 1(1), 12-23.