



## PENERAPAN METODE *MEL FREQUENCY CEPSTRAL COEFFICIENTS* PADA SISTEM PENGENALAN SUARA BERBASIS *DESKTOP*

Galih Ajinurseto\*, La Ode Bakrim, Nur Islamuddin

Program Studi Sistem Komputer, Sekolah Tinggi Manajemen Informatika Dan Komputer Bina Bangsa, Indonesia

**Abstrak:** Teknologi biometrik sedang menjadi tren teknologi dalam berbagai bidang kehidupan. Teknologi biometrik memanfaatkan bagian tubuh manusia sebagai alat ukur sistem yang memiliki keunikan disetiap individu. Suara merupakan bagian tubuh manusia yang memiliki keunikan dan cocok dijadikan sebagai alat ukur dalam sistem yang mengadopsi teknologi biometrik. Sistem pengenalan suara adalah salah satu penerapan teknologi biometrik yang fokus kepada suara manusia. Sistem pengenalan suara memerlukan metode ekstraksi fitur, salah satu metode ekstraksi fitur adalah metode *Mel Frequency Cepstral Coefficients*. Metode *Mel Frequency Cepstral Coefficients* merupakan metode ekstraksi fitur suara yang mengadopsi prinsip indra pendengaran manusia dengan tujuan mendapatkan hasil yang semirip mungkin sebagaimana indra pendengaran manusia. Metode ini dimulai dari tahap *pre-emphasis*, *frame blocking*, *windowing*, *fast fourier transform*, *mel frequency wrapping* dan *cepstrum*. Berdasarkan hasil pengujian, metode *Mel Frequency Cepstral Coefficients* pada pengujian dengan kondisi ideal, persentase keberhasilan sistem mencapai 90% dan persentase kegagalan sistem sebesar 10% dengan *top 5 error rate* sebesar 0%, sedangkan pada pengujian dengan kondisi tidak ideal, persentase keberhasilan sistem sebesar 76.6667% dan persentase kegagalan sistem sebesar 23.3333% dengan *top 5 error rate* sebesar 0%.

**Kata kunci:** *Mel Frequency Cepstral Coefficients*, Pengenalan, Sistem, Suara

### I. PENDAHULUAN

Tahun demi tahun perkembangan teknologi informasi menjadi sangat pesat dan tak terhentikan sehingga menciptakan berbagai jenis teknologi yang memiliki manfaat dan tujuan yang berbeda-beda. Salah satu teknologi populer saat ini adalah teknologi biometrik. Teknologi biometrik adalah teknologi yang menjadikan tubuh manusia sebagai ukuran dalam memutuskan sesuatu. Tubuh manusia yang sering dijadikan ukuran

biometrik adalah mata, sidik jari, wajah, suara dan lain-lain (Yang et al., 2021). Suara merupakan salah satu ukuran biometrik yang memiliki banyak potensi dalam implementasi teknologi biometrik. Suara adalah sekumpulan sinyal yang terbentuk dari getaran yang dapat didengar oleh manusia dan dapat diukur menggunakan satuann waktu dan ruang (Boles & Rad, 2017).

Teknologi biometrik terutama suara belum banyak dimanfaatkan oleh kalangan luas sehingga membuat jumlah penelitian yang sedikit dan kurang populer. Hal tersebut berbanding terbalik dengan data dari *Capgemini Research Institute* yang menyatakan bahwa keperluan teknolgi

<sup>\*)</sup> galihajinurseto@stimikbinsa.ac.id

berbasis suara meningkat akibat pandemi sebesar 77% dan 62% pasca pandemi dan sebesar 59% masyarakat ingin mengakses aplikasi fasilitas publik menggunakan suara agar terhindar dari keperluan melakukan kontak fisik(Capgemini, 2020).

Pengenalan suara atau yang dikenal dengan Speech Recognition adalah pengembangan teknik dan sistem teknologi yang diimplementasikan pada mesin pintar untuk dapat menerima input berupa suara yang disusun menjadi kata maupun kalimat yang diucapkan(Purwins et al., 2019). Teknologi ini memungkinkan sebuah mesin dapat belajar dan mengenali sumber suara dengan cara mengolah suara yang berupa kata atau kalimat menjadi sinyal suara dan, lalu sinyal suara tersebut akan diubah kedalam bentuk data digital untuk kemudian dilakukan pencocokan sinyal tersebut dengan suatu pola tertentu yang telah tersimpan dalam mesin pembelajaran(Ridho, 2019).

Sebelumnya telah ada beberapa penelitian yang berhubungan dengan topik pengenalan suara yang dijadikan acuan oleh peneliti yaitu, penelitian mengenali suara yang dilakukan oleh Adhinata dkk dengan judul Pengenalan Jenis Kelamin Manusia Berbasis Suara Menggunakan MFCC dan GMM. Berdasarkan pengujian pada penelitian tersebut. Didapatkan hasil berupa metode MFCC-GMM memiliki tingkat akurasi sebesar 81.18% dalam mengenali kelamin dari suara inputan(Adhinata et al., 2021).

Penelitian oleh Permana dkk dengan judul Implementasi Metode Mfcc Dan Dtw Untuk Pengenalan Jenis Suara Pria Dan Wanita menghasilkan bahwa tingkat akurasi yang didapatkan pada penelitian ini adalah 80% untuk suara alto, 90% untuk suara mezzosopran, 80% untuk suara sopran, 80% untuk suara bass, 70% untuk suara bariton

dan 60% untuk suara tenor(Permana et al., 2018).

Penelitian oleh Moch Wisuda Sardjono dengan judul Pengenalan Suara Pembicara Menggunakan Ekstraksi MFCC dengan Metode *Gaussian Mixture Model* (GMM). Berdasarkan hasil penelitian dengan menggunakan koefisien MFCC tanpa energi sebanyak 12 dan 20 didapatkan tingkat akurasi sebesar 73,185% dan 77,621%. Sedangkan pengujian dengan menggunakan energi, menghasilkan tingkat akurasi sebesar 72,177% dan 78,831%(Sardjono, 2017).

Penelitian Purwono Prasetyawan dengan judul Perbandingan Identifikasi Pembicara Menggunakan Mfcc Dan Sbc Dengan Ciri Pencocokan Lbg-Vq. identifikasi pembicara menggunakan ekstrasi ciri MFCC dengan ciri pencocokan LBG-VQ dapat diimplementasikan pada speaker identification secara *realtime*, tetapi masih belum cukup memuaskan akurasinya, masih dibawah 70%(Prasetyawan, 2016).

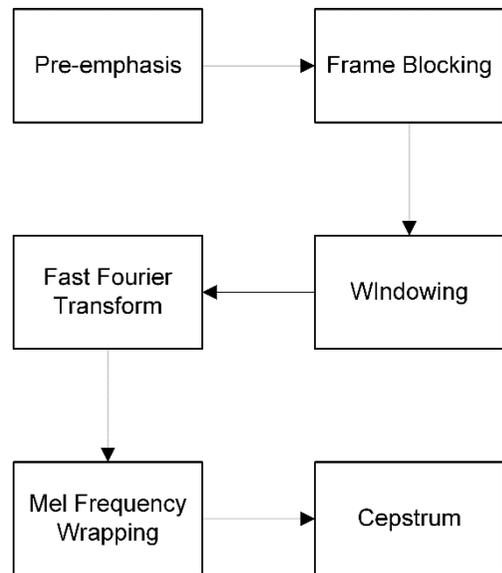
Berdasarkan kajian yang telah penulis jelaskan, dalam penelitian ini penulis akan meneliti Penerapan Metode *Mel Frequency Cepstral Coefficients* Pada Sistem Pengenalan Suara Berbasis *Desktop*. Tujuan yang diharapkan pada penelitian ini penulis mampu meneliti sistem pengenalan suara dan menganalisis peforma metode *Mel Frequency Cepstral Coefficients* berdasarkan tingkat akurasi sistem dalam mengenali suara. Alasan metode *Mel Frequency Cepstral Coefficients* dipilih karena memiliki persentase paling baik dibanding metode yang lain dan didukung oleh literatur yang dimiliki penulis. Dengan adanya penelitian ini penulis berharap agar teknologi biometrik suara lebih banyak diteliti sehingga dapat sepopuler dengan ukuran biometrik lain.

## II. METODOLOGI

### 2.1. Metode *Mel Frequency Cepstral Coefficients*

Pengenalan suara mempunyai berbagai algoritma yang sering digunakan untuk proses ekstraksi ciri data suara, salah satunya adalah algoritma *Mel Frequency Cepstral Coefficients* (MFCC). MFCC merupakan Algoritma atau algoritma yang mengubah ciri suara menjadi sebuah koefisien yang dapat mewakili ciri suara tersebut dan didasarkan pada perhitungan frekuensi suara yang dapat didengar oleh manusia atau audiosonik. MFCC merupakan metode ekstraksi suara yang populer dan salah satu yang terbaik. MFCC dikenalkan kepada dunia oleh peneliti Davis dan Mermelstein pada akhir tahun 1980-an. Ciri khas metode MFCC ketika ekstraksi suara adalah dengan mengkonversi data suara menjadi data berupa grafik gelombang spektrum berdasarkan frekuensi data suara (Permana et al., 2018).

MFCC mengadopsi alur pemrosesan ekstraksi ciri suara yang mirip seperti cara kerja indra pendengaran manusia dalam menangkap dan mengelola suara sehingga hasil MFCC mirip atau mendekati hasil pendengaran indra pendengaran manusia. Kelebihan MFCC dalam ekstraksi ciri suara adalah MFCC memiliki kemampuan untuk memproduksi data dari ciri-ciri suara yang tidak menghilangkan atau mengurangi informasi penting yang terdapat pada data suara dengan jumlah yang minimal. MFCC terdiri atas tahap *pre-emphasis*, *frame blocking*, *windowing*, *fast fourier transform*, *mel frequency wrapping* dan *cepstrum* (Maurya et al., 2018). Ilustrasi tahapan *Mel Frequency Cepstral Coefficients* dapat dilihat pada Gambar 1



**Gambar 1.** Ilustrasi Tahapan *Mel Frequency Cepstral Coefficients*.

*Pre-emphasis* adalah tahapan paling awal MFCC dan berfungsi sebagai filter yang mengurangi *noise* atau derau pada data suara. *Frame Blocking* adalah tahapan setelah *pre-emphasis* dimana data suara akan dibagi menjadi beberapa bagian yang disebut *frame* sehingga lebih mudah untuk diproses oleh system. *Windowing* adalah tahap dimana bagian-bagian suara atau *frame* suara diperbaiki apabila terdapat diskontinuitas pada setiap ujung *frame* suara. *Fast Fourier Transform* (FFT) adalah tahapan untuk mengubah setiap *frame* dari data suara yang telah melalui tahapan *windowing* dari domain berbasis waktu menjadi domain berbasis frekuensi. Tahapan *Mel Frequency Wrapping* adalah tahapan yang mengaplikasikan skala mel pada data suara yang telah diterima dari proses sebelumnya dengan cara membuat *filter bank*. Tahapan *Cepstrum* adalah tahapan yang bertujuan untuk mendapatkan informasi atau ciri dari suatu data suara yang diterima oleh sistem. Sebagai tahapan akhir dari algoritma MFCC, hasil dari tahapan

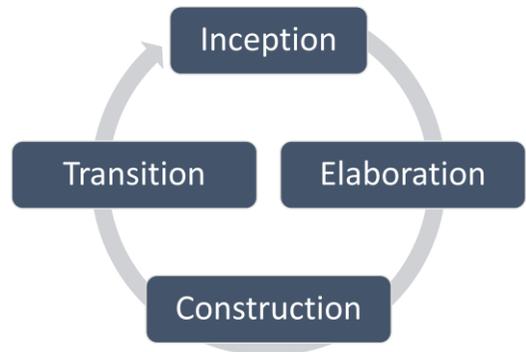
sebelumnya akan diubah menjadi *cepstrum* menggunakan metode DCT atau *discrete cosine transform*. Hasil dari dct inilah yang disebut dengan MFCC. Hasil inilah yang memiliki data berupa ciri-ciri dari data suara yang dapat digunakan sebagai pola untuk mengenali suara(Putra et al., 2017).

## 2.2. Metode *Rational Unified Process*(RUP)

*Rational Unified Process* atau sering disebut dengan RUP merupakan suatu metode pengembangan perangkat lunak dengan cara mengumpulkan berbagai informasi dan kebutuhan sistem dan pengguna terbaik dalam proses pengembangan perangkat lunak. RUP memiliki keunikan dimana menggunakan pendekatan dengan bahasa pemodelan UML dan iteratif untuk menjalankan siklus pengembangan perangkat lunak(Rambe et al., 2020). RUP mempunyai empat tahapan yaitu *inception*, *elaboration*, *construction*, dan *transition*.

Tahap pertama adalah *Inception* (permulaan). Pada tahap *inception* dilakukan pendefinisian ruang lingkup dan perkiraan jadwal serta mencari tahu apa saja kebutuhan pengguna serta membuat rancangan awal sistem. Tahap selanjutnya adalah *Elaboration* (perencanaan) Tahapan ini berguna untuk memulai perencanaan arsitektur sistem. Pada perencanaan arsitektur sistem dimulai dari melakukan spesifikasi fitur perangkat lunak yang akan digunakan dan membuat *prototype* sistem. Selanjutnya adalah *Construction* (konstruksi). Tahapan ini adalah tahap pengimplementasian rancangan sistem dari tahap sebelumnya. Tahapan ini akan berfokus pada implementasi metode dan rancangan sistem pada kode program serta melakukan pengujian sistem. Tahap terakhir adalah *Transition* (transisi). Tahapan ini adalah tahapan terakhir dari RUP. Tahapan ini bertujuan untuk melakukan pelatihan pada

*user* yang akan menggunakan sistem serta melakukan pengujian sistem apakah sudah memenuhi harapan *user*(Pratama, 2019).



**Gambar 2.** Ilustrasi Metode *Rational Unified Process*

## III. HASIL DAN PEMBAHASAN

### 3.1. Implementasi Sistem

Tahapan implementasi sistem merupakan tahap pembangunan menggunakan bahasa pemrograman dan berdasarkan rancangan sistem yang telah direncanakan pada bab sebelumnya dengan tujuan untuk memenuhi tujuan penelitian ini. Sistem pengenalan suara membutuhkan kebutuhan perangkat lunak dan perangkat keras. Adapun kebutuhan tersebut bertujuan agar sistem yang dibangun dapat berjalan dengan lancar dan mampu memenuhi tujuan dan keperluan penelitian.

Guna memenuhi tujuan dan rancangan penelitian seperti yang telah dijelaskan sebelumnya, dibutuhkan beberapa perangkat keras dan perangkat lunak sebagai roda penggerak sekaligus implementasi sistem yang dibangun. Berikut spesifikasi minimum perangkat keras yang dibutuhkan dapat dilihat pada Tabel 1. Dan kebutuhan perangkat lunak dapat dilihat pada Tabel 2.

**Tabel 1.** Kebutuhan Perangkat Keras

No	Nama Perangkat	Spesifikasi
1.	Laptop	Lenovo Ideapad 320-IKB
2.	Processor	Intel(R) Core(TM) i5-7200U CPU @ 2.50GHz 2.71 GHz
3.	Monitor	14 Inch
4.	Memori	RAM 8 GB
5.	Harddisk	1 TB HDD dan 120 GB SSD
6.	Sound Meter	30-110 dB

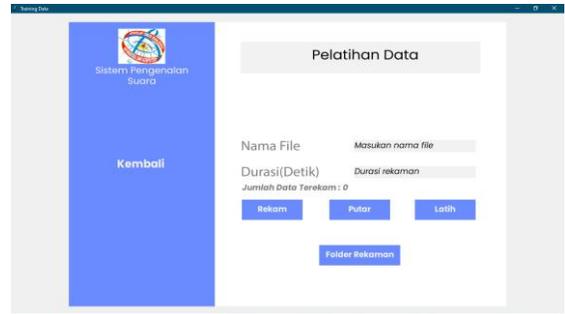
**Tabel 2.** Kebutuhan Perangkat Lunak

No	Nama Perangkat	Spesifikasi
1	Windows	Windows 10 Pro 21H2
2	Pycharm	Pycharm Pro 2020.2
3	Python	Python 3.6

Sistem yang digunakan dalam penelitian ini dibangun menggunakan bahasa pemrograman python. Adapun tampilan sistem dapat dilihat pada Gambar 3., Gambar 4., Gambar 5.



**Gambar 3.** Tampilan Awal



**Gambar 4.** Menu Pelathan Data



**Gambar 5.** Menu Pengujian Data

1. Tombol 'Training Data' pada Gambar 3. digunakan untuk menampilkan menu Training Data.
2. Tombol 'Pengujian Sistem' pada Gambar 3. digunakan untuk menampilkan menu Pengujian Sistem.
3. Tombol 'Keluar' pada Gambar 3. digunakan untuk keluar atau menutup aplikasi.
4. Tombol 'Rekam' pada Gambar 4. digunakan untuk merekam suara yang akan disimpan pada database dan menjadi data latih sistem.
5. Tombol 'Putar' pada Gambar 4. dan Gambar 5. digunakan untuk memutar rekaman suara terakhir yang telah direkam sebelumnya.
6. Tombol "Folder Rekaman" pada Gambar 4. digunakan untuk menampilkan seluruh rekaman data latih yang telah direkam oleh sistem.

7. Tombol ‘Latih’ pada Gambar 4. digunakan untuk membuat model suara.

8. Tombol ‘Rekam’ pada Gambar 5. digunakan untuk merekam suara yang akan dijadikan data uji sistem.

9. Tombol ‘Hasil Pengujian’ pada Gambar 5. digunakan untuk menampilkan hasil pengenalan suara.

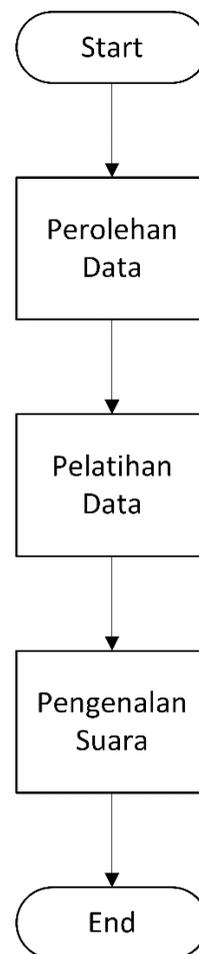
### 3.2. Alur Kerja Sistem

Sistem ini memiliki tiga tahapan utama yaitu tahapan perolehan data, tahapan pelatihan data dan tahapan pengenalan suara. Ilustrasi alur kerja sistem dapat dilihat pada Gambar 6.

Tahapan perolehan data ini bertujuan untuk memperoleh data berupa suara untuk kebutuhan sistem. Sumber data suara berasal dari rekaman suara mahasiswa dengan durasi selama 5 detik dengan mengucapkan kalimat “Nama” sebanyak 27 rekaman untuk tahapan pelatihan dan 6 rekaman untuk tahapan pengujian. Rekaman akan dilakukan pada laptop tempat sistem terpasang dan berformat “.WAV”. Tahapan ini akan dilakukan pada sebelum tahap pelatihan dan pengujian.

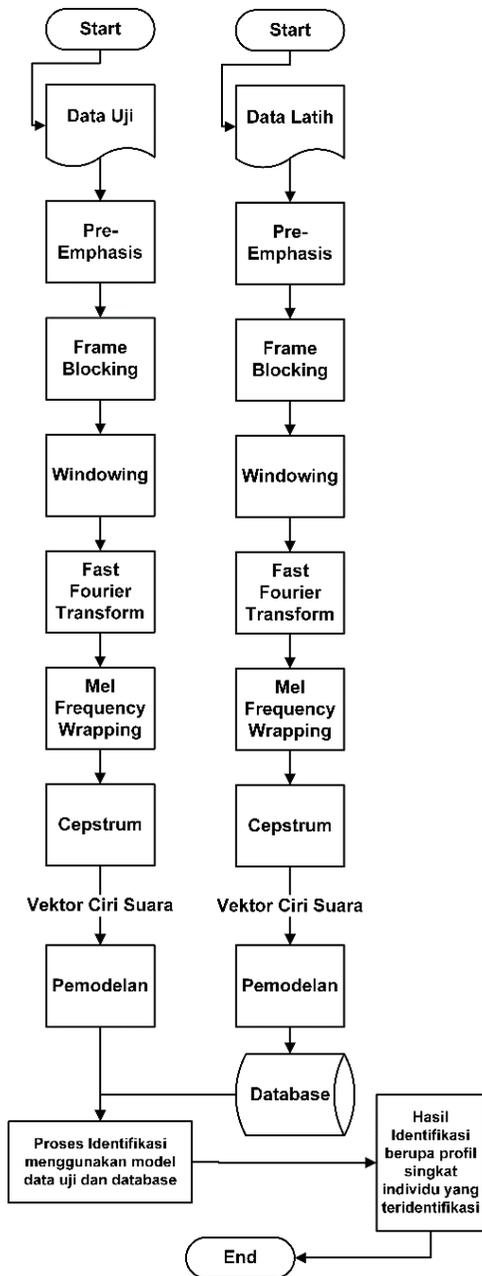
Tahapan Pelatihan dan Pengujian ini bertujuan untuk mendapat model vektor *mel frequency* dari data latih. Tahapan pada proses data latih memiliki kemiripan dengan proses data uji. Yang membedakan kedua proses ini adalah model yang dibuat berdasarkan vektor *mel frequency*.

Data latih akan disimpan ke dalam basis data sistem. Sedangkan pada proses data uji, model yang dibuat berdasarkan vektor *mel frequency* data uji akan dilakukan perhitungan *log likelihood* untuk mengukur tingkat kemiripan model data uji dan model dari basisdata sistem.



**Gambar 6.** Alur Kerja Sistem

Rekaman suara yang digunakan sebagai data latih dan data uji berdurasi 5 detik dengan ucapan yang direkam adalah Nama Mahasiswa. Ilustrasi alur kerja proses pelatihan dan pengujian data dapat dilihat pada Gambar 7.



Gambar 7. Alur Proses Pelatihan dan Pengujian Data

### 3.3. Pengujian *Black Box*

Pengujian ini dilakukan dengan menguji perangkat lunak dari segi fungsionalitasnya.

Pada fungsionalitas perangkat lunak ini diuji sesuai dengan skenario pada tahap desain sistem. Tujuan *black box* yaitu untuk mengetahui bagian-bagian dalam sistem aplikasi telah benar menampilkan pesan-pesan kesalahan jika terjadi kesalahan pada penginputan oleh *user*. Adapun hasil pengujian *blackbox* pada penelitian ini dapat dilihat pada Tabel 3.

Tabel 3. Pengujian *Black Box*

No	Input	Detail Uji	Output	Hasil
1.	Pengujian Menu <i>User</i>	Menu Training	Menampilkan halaman training data	Valid
		Menu Pengujian	Menampilkan halaman pengujian data	Valid
		Tombol Keluar	Keluar dari aplikasi	Valid
		Tombol Kembali	Kembali ke halaman sebelumnya	Valid
		Tombol Rekam	Merekam suara data latihan	Valid
		Tombol Play	Memutar Rekaman data latihan Terakhir	Valid
2.	Pengujian Menu Pelatihan Data	Tombol Latihan	Membuat Model Suara	Valid
		Record Counter	Menampilkan jumlah rekaman yang telah direkam untuk <i>user</i>	Valid
		Tombol Folder Rekaman	Menampilkan folder rekaman suara data latihan	Valid
3.	Pengujian Menu Pengujian Sistem	Tombol Rekam	Merekam suara data uji	Valid
		Tombol Play	Memutar Rekaman data uji	Valid
		Tombol Hasil Pengujian	Menampilkan hasil pengenalan suara sistem	Valid

### 3.3 Pengujian Akurasi Pengenalan Suara

Hasil pengujian pada penelitian ini terbagi atas dua yaitu pengujian dengan kondisi ideal dan pengujian dengan kondisi tidak ideal. Pengujian dengan kondisi ideal adalah pengujian dimana tidak terdapat gangguan berupa noise atau derau saat proses pengambilan data uji sedangkan pengujian dengan kondisi tidak ideal adalah pengujian dimana terdapat gangguan saat pengambilan data uji berupa *noise* dan derau. Tujuan pengujian ini adalah untuk mengetahui seberapa baik peforma algoritma *Mel Frequency Cepstral Coefficients* untuk mengenali suara dalam kondisi ideal dan tidak ideal. Hasil pengujian kondisi ideal dapat dilihat pada Tabel 4, dan hasil pengujian kondisi tidak ideal dapat dilihat pada Tabel 5.

**Tabel 4.** Hasil Pengujian Kondisi Ideal

No	Nama	dB	Benar	Salah	ER
1	Muhammad risal padli	68-71	3	0	0%
2	Muhammad Andi wahyudi	70-72	2	1	0%
3	Fadrisal Ainun Ahyat	70-73	3	0	0%
4	I Made Agus Mahendra	69-71	3	0	0%
5	Jusmin	67-70	3	0	0%
6	Diki Herbiansyah	68-72	2	1	0%
7	Irene Hedry Saudi	70-73	3	0	0%
8	Komang Sindi Sapitri	69-74	3	0	0%
9	Komang Doni Iswara	70-74	3	0	0%
10	Resmin	71-74	2	1	0%
Total		30	27	3	0%

**Tabel 5.** Hasil Pengujian Kondisi Tidak Ideal

No	Nama	dB	Benar	Salah	ER
1	Muhammad risal padli	68-71	3	0	0%
2	Muhammad Andi wahyudi	70-72	2	1	0%
3	Fadrisal Ainun Ahyat	70-73	1	2	0%
4	I Made Agus Mahendra	69-71	3	0	0%
5	Jusmin	67-70	3	0	0%
6	Diki Herbiansyah	68-72	2	1	0%
7	Irene Hedry Saudi	70-73	3	0	0%
8	Komang Sindi Sapitri	69-74	1	2	0%
9	Komang Doni Iswara	70-74	3	0	0%
10	Resmin	71-74	2	1	0%
Total		30	23	7	0%

Dari hasil-hasil Tabel diatas dapat disimpulkan persentase keberhasilan pengenalan sistem dan persentase kegagalan pengenalan sebagai berikut:.

Persentase pengujian kondisi Ideal yang berhasil dikenali sistem :

$$\text{Akurasi} = (27/(27+3)) * 100\% = 90\%$$

Persentase pengujian kondisi Tidak Ideal yang berhasil dikenali sistem :

$$\text{Akurasi} = (23/(23+7)) * 100\% = 76.66667\%$$

Metode MFCC memiliki persentase terbaik pada kondisi ideal dengan tingkat persentase keberhasilan sistem dalam mengenali suara sebesar 90%, sedangkan pada kondisi tidak ideal didapatkan hasil persentase sistem dalam mengenali suara sebesar 76.6667%.

#### IV. KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan Penerapan Metode *Mel Frequency Cepstral Coefficients* Pada Sistem Pengenalan Suara Berbasis *Desktop*, maka diperoleh kesimpulan bahwa Pada pengujian dengan kondisi ideal, persentase keberhasilan sistem mencapai 90% dan persentase kegagalan sistem sebesar 10% dengan *top 5 error rate* sebesar 0%, sedangkan pada pengujian dengan kondisi tidak ideal, persentase keberhasilan sistem sebesar 76.6667% dan persentase kegagalan sistem sebesar 23.3333% dengan *top 5 error rate* sebesar 0%.

Berdasarkan hasil pengujian pada penelitian ini, performa metode *Mel Frequency Cepstral Coefficients* untuk pengenalan suara memiliki hasil yang baik dengan persentase mencapai 90% dibanding penelitian lain yang dijadikan penulis sebagai acuan dalam penelitian ini. Dapat disimpulkan bahwa metode *mel frequency cepstral coefficients* menjadi metode yang terbaik untuk ekstraksi dan pengenalan suara.

#### DAFTAR PUSTAKA

- Adhinata, F. D., Rakhmadani, D. P., & Segara, A. J. T. (2021). Pengenalan Jenis Kelamin Manusia Berbasis Suara Menggunakan MFCC dan GMM. *Jorunal of Data Science, IoT, Machine Learnings and Informatics*, 1(1), 1–12.
- Boles, A., & Rad, P. (2017). Voice biometrics: Deep learning-based voiceprint authentication system. *2017 12th System of Systems Engineering Conference, SoSE 2017*. <https://doi.org/10.1109/SYSOSE.2017.7994971>
- Capgemini. (2020). COVID-19 and the age of the contactless customer experience. In *Capgemini*. <https://www.capgemini.com/insights/research-library/covid-19-and-the-age-of-the-contactless-customer-experience/>
- Maurya, A., Kumar, D., & Agarwal, R. K. (2018). Speaker Recognition for Hindi Speech Signal using MFCC-GMM Approach. *Procedia Computer Science*, 125, 880–887. <https://doi.org/10.1016/j.procs.2017.12.112>
- Permana, I. S., Indrawaty, Y., & Zulkarnain, A. (2018). Implementasi Metode MFCC Dan DTW Untuk Pengenalan Jenis Suara Pria Dan Wanita. *MIND Journal*, 3(1), 49–63. <https://doi.org/10.26760/mindjournal.v3i1.61-76>
- Prasetyawan, P. (2016). Perbandingan Identifikasi Pembicara Menggunakan MFCC Dan SBC Dengan Ciri Pencocokan LBG-VQ. *Seminar Nasional Teknologi Informasi Dan Komunikasi*. <https://doi.org/10.31227/osf.io/85k9u>
- Pratama, A. R. (2019). *Belajar UML - Use Case Diagram*. <https://codepolitan.com/blog/mengenal-uml-diagram-use-case>
- Purwins, H., Li, B., Virtanen, T., Schlüter, J., Chang, S. Y., & Sainath, T. (2019). Deep Learning for Audio Signal Processing. *IEEE Journal on Selected Topics in Signal Processing*, 13(2), 206–219. <https://doi.org/10.1109/JSTSP.2019.2908700>
- Putra, D. K., Triasmoro, I. I., Atmaja, R. D., Iwut, I., & Atmaja, R. D. (2017). Simulasi Dan Analisis Speaker Recognition

Menggunakan Metode Mel Frequency Cepstrum Coefficient (MFCC) Dan Gaussian Mixture Model (GMM). *EProceedings of Engineering*, 4(2), 1766–1772.  
<http://libraryproceeding.telkomuniversity.ac.id/index.php/engineering/article/view/487/460>

Rambe, B. H., Pane, R., Irmayani, D., Nasution, M., Munthe, I. R., Ekonomi, F., & Bisnis, D. (2020). UML Modeling and Black Box Testing Methods in the School Payment Information System. *Jurnal Mantik*, 4(3), 1634–1640.  
<https://iocscience.org/ejournal/index.php/mantik>

Ridho, A. Z. (2019). *Coefficient Dan Dynamic Time Warping K Nearest Neighbour Dalam Rekognisi Aksentu Suku Di*

*Indonesia Skripsi Oleh: Asrori Zainur Ridho* [Universitas Islam Negeri Maulana Malik Ibrahim Malang]. <http://etheses.uin-malang.ac.id/17087/>

Sardjono, M. W. (2017). Pengenalan Suara Pembicara Menggunakan Ekstraksi MFCC dengan Metode Gaussian Mixture Model ( GMM ). In *JURNAL ILMIAH INFORMATIKA KOMPUTER*.  
[http://moch\\_wisuda.staff.gunadarma.ac.id/Downloads/files/81683/Penulisan-Pengenalan+Suara+Pembicara+-+MFCC+-+GMM.pdf](http://moch_wisuda.staff.gunadarma.ac.id/Downloads/files/81683/Penulisan-Pengenalan+Suara+Pembicara+-+MFCC+-+GMM.pdf)

Yang, W., Wang, S., Sahri, N. M., Karie, N. M., Ahmed, M., & Valli, C. (2021). Biometrics for internet-of-things security: A review. *Sensors*, 21(18), 1–26.  
<https://doi.org/10.3390/s21186163>