PEMODELAN SISTEM *SPEAKER RECOGNITION* BERBASIS *MEL FREQUENCY CEPSTRAL COEFFICIENTS* (MFCC) DAN *SUPPORT VECTOR MACHINE* (SVM) UNTUK AUTENTIKASI SUARA

(Skripsi)

Oleh

M. RYANDI FATRIQ

2015031017

PROGRAM STUDI TEKNIK ELEKTRO

JURUSAN TEKNIK ELEKTRO

FAKULTAS TEKNIK

UNIVERSITAS LAMPUNG

BANDAR LAMPUNG

2025

ABSTRAK

PEMODELAN SISTEM *SPEAKER RECOGNITION* BERBASIS *MEL FREQUENCY CEPSTRAL COEFFICIENTS* (MFCC) DAN *SUPPORT VECTOR MACHINE* (SVM) UNTUK AUTENTIKASI SUARA

Oleh

M. RYANDI FATRIQ

Teknologi pengenalan suara berkembang pesat dan menjadi salah satu metode autentikasi biometrik yang menjanjikan karena kemudahan dan keakuratannya. Namun, masih terdapat tantangan dalam hal akurasi sistem dan pengaruh lingkungan seperti kebisingan. Penelitian ini bertujuan untuk membangun dan mengevaluasi sistem autentikasi suara berbasis *Mel Frequency Cepstral Coefficients* (MFCC) dan *Support Vector Machine* (SVM). Proses diawali dengan perekaman suara tiga pembicara yang mengucapkan perintah "buka" dan "tutup", kemudian dilakukan ekstraksi fitur menggunakan metode MFCC. Model klasifikasi dilatih menggunakan algoritma SVM dengan kernel Radial Basis Function (RBF). Pengujian dilakukan dalam dua kondisi lingkungan: ruangan senyap dan ruangan dengan noise. Hasil menunjukkan bahwa akurasi sistem mencapai rata-rata 63,3% pada kondisi ideal, dan menurun menjadi 41,6% pada kondisi dengan noise. Sistem menunjukkan performa lebih baik dalam mengenali perintah "buka" dibandingkan "tutup". Penelitian ini menyimpulkan bahwa kombinasi MFCC dan SVM dapat membentuk sistem speaker recognition yang cukup andal pada lingkungan senyap, namun masih perlu perbaikan untuk pengenalan suara dalam kondisi berisik.

Kata Kunci: *Speaker Recognition*, MFCC, SVM, Autentikasi Suara, *Noise*.

ABSTRACT

SPEAKER RECOGNITION SYSTEM MODELING BASED ON MEL FREQUENCY CEPSTRAL COEFFICIENTS (MFCC) AND SUPPORT VECTOR MACHINE (SVM) FOR VOICE AUTHENTICATION

By

M. RYANDI FATRIQ

Voice recognition technology has rapidly advanced and emerged as one of the most promising biometric authentication methods due to its convenience and accuracy. However, challenges remain in achieving high system accuracy and maintaining performance under varying environmental conditions, particularly in the presence of noise. This study aims to develop and evaluate a voice authentication system based on Mel Frequency Cepstral Coefficients (MFCC) and Support Vector Machine (SVM). The process begins with recording voice samples from three speakers uttering the commands “buka” and “tutup” (meaning “open” and “close”), followed by feature extraction using the MFCC method. The classification model is trained using the SVM algorithm with a Radial Basis Function (RBF) kernel. The system is tested under two environmental conditions: a quiet room and a noisy room. Experimental results show that the system achieves an average accuracy of 63% in ideal (quiet) conditions, which decreases to 41.6% in the presence of noise. The system demonstrates higher accuracy in recognizing the command “buka” compared to “tutup.” This study concludes that the combination of MFCC and SVM can form a reliable speaker recognition system in noise-free environments but requires further improvement to maintain performance in noisy conditions.

Keywords: *Speaker Recognition, MFCC, SVM, Voice Authentication, Noise.*

PEMODELAN SISTEM *SPEAKER RECOGNITION* BERBASIS *MEL FREQUENCY CEPSTRAL COEFFICIENTS* (MFCC) DAN *SUPPORT VECTOR MACHINE* (SVM) UNTUK AUTENTIKASI SUARA

Oleh

M. RYANDI FATRIQ

SKRIPSI

Sebagai Salah Satu Syarat untuk Mencapai Gelar

SARJANA TEKNIK

Pada

Program Studi S1 Teknik Elektro
Jurusan Teknik Elektro
Fakultas Teknik

PROGRAM STUDI TEKNIK ELEKTRO
JURUSAN TEKNIK ELEKTRO
FAKULTAS TEKNIK
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG
2025

Judul Skripsi	:	**PEMODELAN SISTEM *SPEAKER RECOGNITION* BERBASIS *MEL FREQUENCY CEPSTRAL COEFFICIENTS* (MFCC) DAN *SUPPORT VECTOR MACHINE* (SVM) UNTUK AUTENTIKASI SUARA**
Nama Mahasiswa	:	M. RYANDI FATRIQ
No. Pokok Mahasiswa	:	2015031017
Program Studi	:	Teknik Elektro
Fakultas	:	Teknik

MENYETUJUI

1. Komisi Pembimbing

Pembimbing Utama	Pembimbing Pendamping
Syaiful Alam, S.T., M.T	**Herlinawati, S.T., M.T.**
NIP. 196904161998031004	NIP. 197103141999032001

2. Mengetahui

Ketua Jurusan Teknik Elektro	Ketua Program Studi Teknik Elektro
Herlinawati, S.T., M.T.	**Sumadi, S.T., M.T.**
NIP. 197103141999032001	NIP. 197311042000031001

MENGESAHKAN

1. Tim Penguji

Ketua : **Syaiful Alam, S.T., M.T** ----------

Sekretaris : **Herlinawati, S.T., M.T.** ----------

Penguji Utama : **Dr. Sri Purwiyanti, S.T.,M.T., Ph.D** ----------

2. Dekan Fakultas Teknik

Dr. Eng. Ir. Helmy Fitriawan, S.T, M.Sc.

NIP. 197509282001121002

Tanggal Lulus Ujian Skripsi : **1 Agustus 2025**

SURAT PERNYATAAN

Dengan ini saya menyatakan bahwa dalam skripsi ini tidak terdapat karya yang pernah dilakukan orang lain dan sepanjang sepengetahuan saya tidak terdapat karya atau pendapat yang ditulis atau diterbitkan orang lain, kecuali secara tertulis diacu dalam naskah ini sebagaimana yang disebutkan di dalam daftar Pustaka. Selain itu, saya menyatakan pula bahwa skripsi ini dibuat oleh saya sendiri.

Apabila pernyataan saya tidak benar, maka saya bersedia dikenakan sanksi sesuai hukum yang berlaku.

Bandar Lampung, 28 Oktober 2025

M. Ryandi Fatriq

NPM. 2015031017

RIWAYAT HIDUP

M. Ryandi Fatriq lahir di Baturaja pada tanggal 12 November 2002, anak dari Bapak Toriq dan Ibu Fatmawati sebagai anak pertama dari tiga bersaudara, Pendidikan formal dimulai dari TK Nurul Falah, SD Negeri 20 OKU, SMP Negeri 13 OKU, dan SMA Negeri 4 OKU, dengan tahun lulus sekolah yaitu pada tahun 2020, pada tahun tersebut saya resmi terdaftar sebagai mahasiswa baru Program Studi S1 Teknik Elektro, Fakultas Teknik, Universitas Lampung, diterima melalui jalur Seleksi Nasional Masuk Perguruan Tinggi Negeri (SNMPTN).

Selama menjadi mahasiswa, saya aktif terlibat dalam berbagai kegiatan di kampus dari bidang akademik maupun non akademik, antara lain menjadi anggota maupun pengurus di Himpunan Mahasiswa Teknik Elektro (HIMATRO) selama 2 periode kepengurusan, dan menjadi Asisten Laboratorium Teknik Elektronika pada tahun 2022 hingga 2025. Saya juga melakukan kerja praktik (KP) di PT. Juang Jaya Abdi Alam pada tahun 2023 dan melakukan magang di PT. PLN UID Rajabasa Bandarlampung.

PERSEMBAHAN

Dengan Ridho Allah SWT
Teriring shalawat kepada Nabi Muhammad SAW
Karya Tulis ini ku persembahkan untuk:

Ayah dan Bunda tercinta

A.Toriq Mandaki dan Fatmawati

Serta Adik-adikku

Nadhira Fatriq dan Hafiz Al Fatriq

Terima kasih atas segala pengorbanan, nasihat dan doa baik yang tidak pernah berhenti kalian berikan kepadaku. Aku selamanya bersyukur dengan keberadaan kalian sebagai keluargaku.

MOTTO

“Lebih baik pernah salah, daripada tidak pernah melakukan kesalahan”

“Yang pergi mengajarkan arti menghargai, yang tersisa mengajarkan arti bertahan”

"Karena sesungguhnya bersama kesulitan itu ada kemudahan. Sesungguhnya bersama kesulitan itu ada kemudahan."
(QS. Al-Insyirah [94]: 5–6)

" Yang berat bukan jalannya, tapi ketika kita lupa tujuan"

SANWACANA

Puji Syukur Alhamdulillah penulis ucapkan kepada Allah SWT yang telah memberikan nikmat, Rahmat dan hidayah-Nya sehingga saya diberikan kelancaran dan kemudahan untuk menjalankan perkuliahan, penelitian, dan penulis naskah skripsi yang berjudul "**Pemodelan Sistem *Speaker Recognition* Berbasis *Mel Frequency Cepstral Coefficients* (MFCC) Dan *Support Vector Machine* (SVM) Untuk Autentikasi Suara**" hingga selesai.

Selama proses menyelesaikan skripsi ini, penulis meyakini banyak mendapatkan bimbingan, masukan, bantuan, dukungan, kritik dan saran dari berbagai pihak. Untuk itu, penulis ingin mengucapkan rasa terima kasih kepada:

1. Ibu Prof. Dr. Ir Lusmeilia Afriani, D.E.A., I.P.M., selaku Rektor Universitas Lampung.
2. Bapak Dr. Eng. Ir. Helmy Fitriawan, S.T., M.Sc., selaku Dekan Fakultas Teknik Universitas Lampung.
3. Ibu Herlinawati, S.T., M.T. selaku Ketua Jurusan Teknik Elektro Fakultas Teknik Universitas Lampung dan juga sekaligus dosen pembimbing bagi penulis, terima kasih yang sebesar-besarnya atas bimbingan, masukan, dan arahan yang telah Ibu berikan selama proses penyusunan skripsi ini. Dukungan dan saran yang Ibu berikan turut membantu penulis dalam menyelesaikan tugas akhir ini dengan baik
4. Bapak Sumadi, S.T., M.T. selaku Ketua Program Studi Teknik Elektro Fakultas Teknik Universitas Lampung.
5. Bapak Syaiful Alam, S.T., M.T. selaku Pembimbing Utama dalam penyusunan Tugas Akhir dan selaku Kepala Laboratorium Elektronika, terima kasih yang sebesar-besarnya atas segala ilmu, bimbingan, dan arahan yang diberikan dengan penuh kesabaran dan ketulusan. Terima kasih atas waktu yang Bapak luangkan, motivasi yang senantiasa menguatkan, serta tantangan yang Bapak berikan yang mendorong penulis untuk berkembang. Kesediaan Bapak membimbing di setiap tahap penyusunan skripsi ini menjadi dorongan berharga hingga penulis dapat menyelesaikannya dengan baik.

6. Ibu Dr. Sri Purwiyanti, S.T.,M.T., Ph.D, selaku dosen penguji, terimakasih sebesar-besarnya atas bimbingan bagi penulis yang telah membantu dalam penyusunan skirpsi ini.
7. Kak Yudi Eka Putra, S.T., M.T. selaku PLP Laboratorium Elektronika. Terima kasih atas ilmu dan diberikannya tempat untuk menyelesaikan tugas akhir ini.
8. Seluruh Dosen dan karyawan Jurusan Teknik Elektro Universitas Lampung, berkat ilmu yang telah diajarkan kepada penulis selama penulis menjalani masa studi di perkuliahan.
9. Keluarga rekan-rekan di Laboratorium Elektronika yang selalu memberikan dukungan, pertolongan, canda tawa, dalam setiap proses apapun selama menjadi asisten Laboratorium Elektronika.
10. Keluarga besar Angkatan 2020, terima kasih atas kebersamaan, dan segala bantuan yang diberikan.
11. Keluarga besar HIMATRO UNILA, yang telah menjadi tempat penulis belajar dan berkembang dalam berorganisasi.
12. Tim Manchester United (1999 – 2008) yang menjadi cinta pertama, yang mengajarkan arti setia, hingga saat ini "*Once Red, always Red*".
13. Seluruh teman-teman, saudara, dan kerabat lainnya yang telah memberikan doa, dukungan, serta semangat, sehingga penulis dapat melalui proses ini dan menyelesaikan skripsi dengan baik.
14. Terakhir kepada seseorang yang pernah bersama penulis pernah berarti banyak dalam perjalanan ini terima kasih atas kebaikan, kesabaran, dan semua hal indah yang pernah kamu bagikan. Meski waktu telah membawa kita pada jalan yang berbeda, doaku selalu baik untukmu.

DAFTAR ISI

Halaman

ABSTRAK ..... **i**

DAFTAR ISI ..... **xii**

DAFTAR GAMBAR ..... **xiv**

DAFTAR TABEL ..... **xvi**

BAB I PENDAHULUAN **1**

1.1 Latar Belakang 1
1.2 Tujuan Penelitian 2
1.3 Manfaat Penelitian 2
1.4 Rumusan Masalah 3
1.5 Batasan Masalah 3
1.6 Hipotesis 3
1.7 Sistematika Penulisan 4

BAB II TINJAUAN PUSTAKA **5**

2.1 Penelitian Terdahulu 5
2.2 *Speaker Recognition* 6
2.3 Ekstraksi Fitur 6
2.3.1 *Pre-emphasis* 7
2.3.2 *Framing* 8
2.3.3 *Windowing* 9
2.3.4 *Mel Filterbank* 10
2.3.5 *Fast Fourirer Transform* 10
2.3.6 *Discrete Consine Transform* 11
2.4 *Support Vector Machine* 11
2.5 *MATLAB* 15
2.6 *Principal Component Analysis* (PCA) 16
2.7 *Mic Fifine K669B* 16

BAB III METODE PENELITIAN **18**

3.1 Waktu dan Tempat Penelitian 18
3.2 Alat dan Bahan Penelitian 18

3.3 Diagram Alir Penelitian 19
3.4 Diagram Blok Perancangan Sistem 22
3.5 Diagram Alir Sistem Data *Train* 23
3.6 Diagram Alir Pemodelan SVM 26
BAB IV HASIL DAN PEMBAHASAN 28
4.1 Perancangan Model Sistem 28
4.1.1 Ekstraksi Fitur Mel Frequency Cepstral Coefficients (MFCC) 30
4.1.2 Model *Support Vector Machine* (SVM) 49
4.2 Hasil Pengujian 51
4.2.1 Evaluasi Kinerja Model 51
4.2.1.1 Confusion Matrix dengan subjek Yandi 51
4.2.1.2 Confusion Matrix dengan subjek Gibran 53
4.2.1.3 Confusion Matrix dengan subjek Faris 55
4.2.2 Hasil Pengujian *Real time* Nilai Akurasi Pada Ruangan tanpa *Noise* 57
4.2.3 Hasil Pengujian *Real time* Nilai Akurasi Pada Ruangan dengan *Noise*.. 59
4.3 Analisis Hasil Pengujian 61
4.3.1 Pengaruh Kondisi Lingkungan terhadap Akurasi Pengenalan Suara 61
4.3.2 Evaluasi Akurasi Sistem pada Kondisi tanpa Noise dan dengan Noise 62
4.3.3 Visulisasi Klasifikasi Kelas SVM menggunakan Principal Component Analysis (PCA) 64
4.3.4 Visualisasi Data Train Suara Yandi 75
4.3.5 Visualisasi Data Train Suara Gibran 77
4.3.6 Visualisasi Data Train Suara Faris 79
BAB V KESIMPULAN DAN SARAN 81
5.1 Kesimpulan 81
5.2 Saran 81
DAFTAR PUSTAKA 83
LAMPIRAN 85

DAFTAR GAMBAR

Halaman

Gambar 2.1 *Time Domain dan Frequency Domain* 7
Gambar 2.2 Alur MFCC..... 7
Gambar 2.3 *Frame Blocking* 8
Gambar 2.4 *Windowing* 9
Gambar 2.5 *Mel-Filterbank*..... 10
Gambar 2.6 SVM Data Terpisah Secara *Linear*..... 12
Gambar 2.7 SVM Data *Non-Linier* 13
Gambar 2.8 MATLAB 15
Gambar 2.9 *Microphone Fifine* K669B..... 17
Gambar 3.1 Diagram Alir Penelitian..... 19
Gambar 3.2 Diagram Blok Sistem..... 22
Gambar 3.3 Diagram Alir Sistem Data Train..... 23
Gambar 3.4 DataTrain.....24
G**ambar 3.5** Diagram Alir Pemodelan SVM..... 26
G**ambar 4.1** Gelombang Suara dari salah satu data audio.....29
Gambar 4.2 Program Modular *Pre Emphasis*.....30
Gambar 4.3 Gelombang Suara ketika telah melewati proses *pre emphasis*.....30
Gambar 4.4 Program Modular *Frame*.....31
Gambar 4.5 Visual 1 *Frame*, bentuk gelombang suara ketika dipotong oleh proses frame dengan durasi 25 ms dan overlap 10ms.....32
Gambar 4.6 Visual 2 *frame*.....33
Gambar 4.7 Visual 3 *frame*.....34
Gambar 4.8 Program Modular *Windowing*.....35
Gambar 4.9 Hasil visual 1 *frame* setelah melewati proses *windowing*.....36
Gambar 4.10 Hasil visual 2 *frame* setelah melewati proses *windowing*.....37
Gambar 4.11 Hasil visual 3 *frame* setelah melewati proses *windowing*.....38
Gambar 4.12 Program Modular *Fast Fourier Transform*.....39
Gambar 4.13 Hasil visual *window* pada *frame* pertama yang telah diproses FFT..40
Gambar 4.14 Hasil visual *window* pada *frame* kedua yang telah diproses FFT.....41

Gambar 4.15 Hasil visual *window* pada *frame* ketiga yang telah diproses FFT.....42
Gambar 4.16 Program Modular *Mel Filter Bank*.....43
Gambar 4.17 Visualisasi Filter Bank Mel (*Mel Filterbank*).....44
Gambar 4.18 Hasil Energi Logaritmik dari Filter Mel pada Frame Pertama.....45
Gambar 4.19 Program Modular *Discrate Consine Transform*.....46
Gambar 4.20 Nilai fitur (13 koefisien).....46
Gambar 4.21 Program untuk menyimpan seluruh nilai ekstraksi dalam satu dataset.....47
Gambar 4.22 Hasil ekstraksi fitur MFCC.....47
Gambar 4.23 Program Pelatihan SVM.....49
Gambar 4.24 *Confusion Matirx* data uji suara Yandi.....51
Gambar 4.25 *Confusion Matirx* data uji suara Gibran.....53
Gambar 4.26 *Confusion Matirx* data uji suara Faris.....55
Gambar 4.27 10 sample data nilai ekstraksi fitur.....65
Gambar 4.28 Hasil dari proses mean centering.....65
Gambar 4.29.a Hasil proyeksi visual antara fitur x1 dengan x2.....67
Gambar 4.29.b Hasil proyeksi visual antara fitur x1 dengan x3.....67
Gambar 4.30 Program Matriks kovarians untuk tahap *eigenvalue*.....69
Gambar 4.31 Hasil Visual Proyeksi 13 Dimensi ke 2 Dimensi hasil dari *Eigenvalue*.....70
Gambar 4.32 Visualisasi PCA dari Data Train.....71
Gambar 4.33 Visual Data yang tumpang tindih dari suara Faris.....72
Gambar 4.34 Visual Data yang tumpang tindih.....73
Gambar 4.35 Visualisasi Data Train Suara Yandi jika hanya data suara Yandi saja pada pemodelan SVM.....75
Gambar 4.36 Visualisasi Data Train Suara Gibran jika hanya data suara Yandi saja pada pemodelan SVM.....77
Gambar 4.37 Visualisasi Data Train Suara Faris jika hanya data suara Yandi saja pada pemodelan SVM.....79

DAFTAR TABEL

Halaman

Tabel 3.1 Alat dan Bahan..... 18
Tabel 4.1 Akurasi SVM pada ruangan tanpa *noise*.....57
Tabel 4.2 Akurasi SVM pada ruangan dengan *noise*.....59

BAB I
PENDAHULUAN

1.1 Latar Belakang

Teknologi pengenalan suara telah berkembang pesat sebagai alat penting dalam sistem keamanan, karena mudah digunakan dan memiliki tingkat keakuratan yang tinggi. Teknologi ini mengubah data digitial menjadi suara yang dikenali, sehingga dapat digunakan untuk mengidentifikasi seseorang berdasarkan karakteristik suaranya yang unik. Saat ini, teknologi pengenalan suara semakin popular karena kepraktisannya dan sering digunakan dalam berbagai bidang, seperti asisten virtual dan sistem keamanan.

Meski memiliki banyak manfaat, masih ada beberapa tantangan dalam mengembangkan sistem pengenalan suara, terutama untuk keamanan. Beberapa masalah yang perlu diatasi meliputi tingkat akurasi dan keandalan sistem, pengaruh lingkungan seperti kebisingan di sekitar, serta resiko penyalahgunaan dengan teknologi pemalsuan suara. Selain itu, masalah privasi juga menjadi perhatian penting, karena data suara yang dikumpulkan harus dijaga keamanannya agar tidak disalahgunakan.

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[1]. Berdasarkan pengujian pada penelitian tersebut. Didapatkan hasil berupa metode MFCC-GMM memiliki tingkat akurasi sebesar 81.18% dalam mengenali kelamin dari suara inputan[1]. Selanjutnya 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[2]. Selanjutnya penelitian yang dilakukan Purwono Prasetyawan dengan judul Perbandingan Identifikasi Pembicara menggunakan MFCC Dan SBC Dengan Ciri Pencocokan LBG-VQ mengidentifikasi 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%[3]

Berdasarkan kajian yang telah penulis jelaskan, dalam penelitian ini penulis akan meneliti dengan membuat Pemodelan Sistem *Speaker Recognition* Berbasis *Mel Frequency Cepstral Coefficients* (MFCC) Dan *Support Vector Machine* (SSVM) Untuk Autentikasi Suara. Tujuan yang diharapkan pada penelitian ini penulis mampu meneliti dan membuat autentikasi 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 penilitian ini penulis berharap agar teknologi biometrik suara lebih banyak diteliti.

1.2 Tujuan Penelitian

Adapun tujuan penelitian ini adalah sebagai berikut:

1. Bagaimana merancang sistem pengenalan pembicara dan perintah suara menggunakan metode MFCC dan SVM?
2. Seberapa akurat sistem dalam mengenali pembicara dan perintah suara pada kondisi ruangan tanpa *noise* dan dengan *noise*?
3. Bagaimana pengaruh kondisi lingkungan (*noise*) terhadap perfoma klasifikasi SVM yang dilatih hanya pada data yang ideal (tanpa *noise*)?

1.3 Manfaat Penelitian

Manfaat penelitian ini adalah mengembangkan sistem autentikasi suara yang dapat mengenali identitas pembicara dan kata perintah secara otomatis,

memanfaatkan karakteristik unik suara untuk autentikasi non-fisik yang efisien dan akurat.

1.4 Rumusan Masalah

Adapun rumusan masalah pada penelitian yang dibahas adalah sebagai berikut:

1. Bagaimana merancang sistem pengenalan pembicara dan perintah suara menggunakan metode MFCC dan SVM?
2. Seberapa akurat sistem dalam mengenali pembicara dan perintah suara pada kondisi ruangan tanpa *noise* dan dengan *noise*?
3. Bagaimana pengaruh kondisi lingkungan (*noise*) terhadap perfoma klasifikasi SVM yang dilatih hanya pada data yang ideal (tanpa *noise*)?

1.5 Batasan Masalah

Adapun batasan masalah pada penelitian ini sebagai berikut:

1. Penelitian ini menggunakan Bahasa Indonesia sebagai bahasa input suara.
2. Data latih diambil dari rekaman vokal manusia yang diambil dari rekaman suara tiga orang pembicara dalam kondisi ruangan senyap.
3. Data uji terdiri dari dua kondisi yaitu diuji dalam kondisi senyap dan berisik.
4. Sistem hanya melakukan proses pengenalan terhadap pembicara dan kata perintah (“buka” dan “tutup”)
5. Penelitian ini difokuskan pada tahap pemodelan dan pengujian sistem.

1.6 Hipotesis

Pada penelitian sistem pemantauan parkir yang dikembangkan diharapkan pengguna untuk mendapatkan informasi slot parkir yang tersedia secara *real-time* melalui aplikasi Android. Dengan menggunakan teknologi *Internet of Things* (IoT) dan mikrokontroler ESP32 sebagai prosesor, informasi ketersediaan parkir bisa diakses dalam waktu kurang dari 10 detik dan bisa dipantau lebih dari 1 pengguna.

1.7 Sistematika Penulisan

BAB I PENDAHULUAN

Bab ini membahas latar belakang, rumusan masalah, batasan masalah, tujuan, manfaat penelitian, hipotesis, dan sistematika penulisan. Fokus utamanya pada perancangan sistem autentikasi suara yang dapat mengenali identitas pembicara dan kata perintah menggunakan metode MFCC dan SVM.

BAB II TINJAUAN PUSTAKA

Bab ini memuat penelitian terdahulu yang relevan, teori dasar mengenai speaker recognition, teknik ekstraksi fitur MFCC (termasuk *pre-emphasis*, *framing*, *windowing*, FFT, *mel filterbank*, dan DCT), konsep algoritma SVM, penggunaan MATLAB dalam penelitian, serta deskripsi perangkat keras seperti mikrofon yang digunakan.

BAB III METODOLOGI PENELITIAN

Bab ini menjelaskan waktu dan tempat penelitian, alat dan bahan yang digunakan, diagram alir penelitian, perancangan sistem, proses pembuatan dataset, tahapan ekstraksi fitur MFCC, serta pemodelan SVM dengan kernel RBF untuk klasifikasi pembicara dan perintah suara.

BAB IV HASIL DAN PEMBAHASAN

Bab ini memaparkan perancangan model sistem, hasil ekstraksi fitur MFCC, pembentukan dan pelatihan model SVM, hasil pengujian pada kondisi senyap dan berisik, evaluasi akurasi menggunakan confusion matrix, analisis pengaruh noise, serta visualisasi distribusi data menggunakan PCA.

BAB V PENUTUP

Bab ini berisi kesimpulan yang diperoleh dari hasil penelitian dan saran untuk pengembangan sistem autentikasi suara berbasis MFCC dan SVM di masa mendatang.

DAFTAR PUSTAKA

BAB II
TINJAUAN PUSTAKA

2.1 Penelitian Terdahulu

Dalam proses penelitian ini untuk memperdalam pemahaman maka dilakukan peninjauan penelitian yang berguna dalam penelitian ini. Pada penelitian yang berjudul "Sistem Keamanan Berbasis Pengenalan Suara Sebagai Pengakses Pintu Menggunakan Metode *Mel Frequency Cepstral Coefficient* (MFCC)" dengan metode *Mel Frequency Cepstral Coefficient* (MFCC) sebagai pengekstraksi ciri serta metode *Vector Quantization* (VQ) sebagai kuantisasi vektor. Dilakukan dengan melakukan dua tahap, yaitu tahap training dan tahap testing. Sistem buka tutup pintu secara otomatis dengan pengenalan suara menggunakan arduino uno dan pemrograman *python* ini memanfaatkan metode *Mel Frequency Cepstral Coeffcient* (MFCC) sebagai ekstaksi ciri dan metode *Vector Quantization* (VQ) sebagai kuantisasi vektor. Dilakukan pada jarak 10 cm dan 15 cm dalam keadaan hening dan ramai. Akurasi pada jarak dalam keadaan hening menghasilkan akurasi sebesar 82% dan pada keadaan ramai menghasilkan akurasi sebesar 74%[4].

Kemudian penelitian berikutnya yang berjudul "Penerapan Metode *Mel Frequency Cepstral Coefficient* Pada Sistem Pengenalan Suara Berbasis Desktop" dalam penelitian ini berfokus pada teknologi biometric. 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 23.333% dengan top 5 error rate sebesar 0%[5].

Penelitian ketiga dengan judul "Implementasi dan Analisis Simulasi Deteksi Emosi Melalui Pengenalan Suara Menggunakan *Mel-Frequency Cepstrum Coefficient* dan *Hidden Markov Model* Berbasis IOT" dalam penelitian ini berfokus pada bagaimana mesin mampu memahami emosi dari percakapan manusia, hasil pendeteksi emosi yang sudah ditentukan sebanyak 7 emosi telah berhasil menentukan 4 jenis emosi dengan benar yang berarti ketepatan data lebih dari 50%[6].

Berbeda dengan beberapa penelitian terdahulu yang umumnya hanya melakukan pengujian pada kondisi suara ideal atau terbatas pada klasifikasi jenis kata atau identitas pembicara secara terpisah, penelitian ini memfokuskan pada pemodelan sistem pengenalan suara yang secara simultan dapat mengenali siapa yang berbicara dan apa yang diucapkan, menggunakan kombinasi metode MFCC dan SVM. Selain itu, penelitian ini juga menguji performa sistem dalam dua kondisi lingkungan yang berbeda, yaitu ruangan senyap dan ruangan berisik, guna mengevaluasi tingkat keandalan sistem terhadap perubahan kondisi.

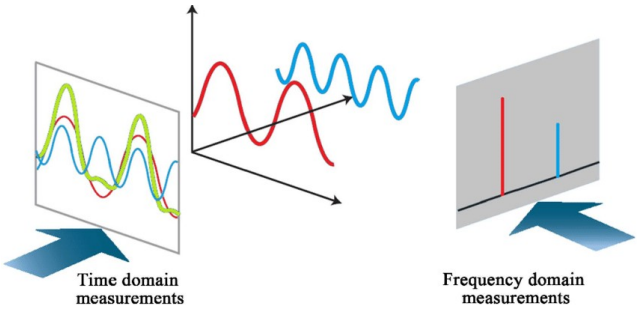
2.2 *Speaker Recognition*

Speaker recognition merupakan sebuah metode yang digunakan untuk mengenali atau memverifikasi identitas seseorang berdasarkan suara yang diucapkan. Proses ini dilakukan dengan menganalisis karakteristik suara dan dibagi menjadi dua pendekatan, yaitu *speaker identification* dan *speaker verification*. Pada *speaker identification*, sistem berupaya menentukan identitas pembicara dengan cara membandingkan fitur suara input dengan seluruh data suara yang tersimpan dalam *database*, guna menemukan kecocokan yang paling sesuai. Sementara itu, *speaker verification* bertujuan untuk memastikan apakah suara yang diberikan sesuai dengan identitas yang telah diketahui sebelumnya, berdasarkan data yang telah disimpan atau didaftarkan sebelumnya oleh pengguna[7].

2.3 Ekstraksi Fitur

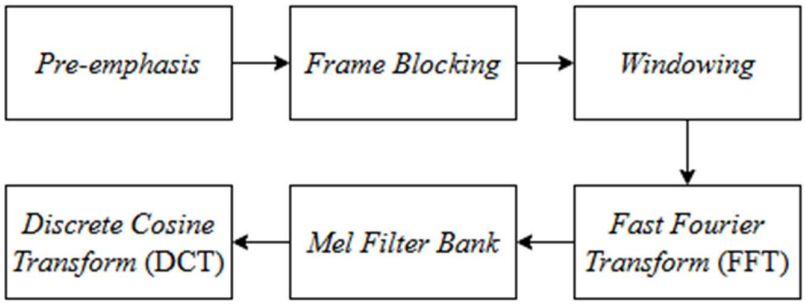
MFCC adalah salah satu teknik ekstraksi fitur paling populer yang digunakan dalam pengenalan suara berdasarkan *domain* frekuensi. MFCC menggunakan skala

Mel yang didasarkan pada skala telinga manusia[8]. MFCC yang dianggap sebagai fitur domain frekuensi, jauh lebih akurat daripada fitur domain waktu[9]. Pada gambar 2.1 dapat dilihat perbedaan antara *time domain* dengan *frequency domain.*



Gambar 2.1 *Time Domain* dan *Frequency Domain*

MFCC adalah satu-satunya pendekatan akustik yang mengambil persepsi manusia (fisiologi dan aspek perilaku dari organ produksi suara) yang berarti pendekatan ini tidak dapat memproses frekuensi diatas 1 KHz. Pendekatan ini mempunyai dua tipe filter yang berjarak secara linear yang memiliki frekuensi rendah dibawah 1000 Hz serta jarak secara *logarithmic* diatas 1000 Hz[10]. Hasil dari MFCC adalah berupa koefisien yang nantinya akan digunakan untuk menentukan klasifikasi data suara, untuk alur dari MFCC dapat dilihat pada Gambar 2.2.[1]



Gambar 2.2 Alur MFCC

2.3.1 *Pre-emphasis*

Pre-emphasis adalah proses filter dengan maksud untuk mengambil bentuk spektral yang lebih halus dari frekuensi sinyal ucapan dan mengurangi kebisingan selama pengambilan suara. Filter *pre-emphasis* diperlukan setelah proses pengambilan sampel dalam proses sinyal ucapan. Penerapan filter pre-emphasis didasarkan pada proses pada input atau output di dalam time-domain, berikut rumus dari *pre-emphasis*[11], [12].

$$Y[n] = X[n] - \alpha \times X[n-1] \quad \text{.....(1)}$$

dimana:

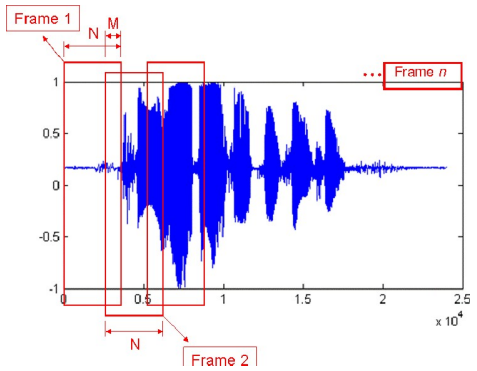
$Y[n]$ = sinyal output atau hasil *pre-emphasis*

$X[n]$ = sinyal input atau sebelum *pre-emphasis*

α = nilai koefisien $0 \leq \alpha \leq 1$, dalam penelitian ini menggunakan 0.97

2.3.2 *Framing*

Frame Blocking adalah proses membagi sampel suara menjadi beberapa *frame* atau slot. Setiap potongan dari sinyal suara disebut dengan frame. Sinyal suara yang dipotong berdurasi pendek dan saling tumpang tindih (*overlapping*) satu sama lain sebanyak N yang kemudian disimpan dalam matrix[13]. Blok *frame* sebanyak N sampel dengan spasi M (M < N)



Gambar 2.3 *Frame Blocking*

Gambar 2.3 menunjukkan satu frame terdiri dari beberapa sampel yang tergantung pada setiap detik suara yang akan di sampel dan berapa besar frekuensi sampling yang digunakan. Proses *frame blocking* menggunakan persamaan berikut:

$$J(f) = \left(\frac{(I-N)}{M} + I\right) \quad \text{.....(2)}$$

dimana:

$J(f)$ = jumlah frame

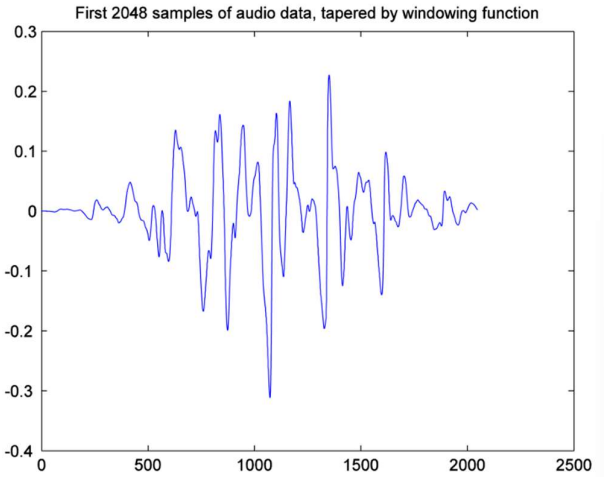
I = Panjang frame

N = ukuran frame

M = pergeseran antar frame

2.3.3 *Windowing*

Windowing adalah proses untuk mengurangi kesenjangan (diskontinuitas) sinyal pada awal dan akhir bingkai setelah proses *frame blocking*[14]. Sampel yang sudah di potong menjadi beberapa *frame* diperlukan konversi sinyal audio menjadi bersifat periodik dan kontinu ketika dibungkus kedalam sebuah *window* seperti pada gambar 2.4 hal ini dilakukan dengan membingkai *frame signal*. Dengan mengalikan setiap *frame* dengan *hamming window*, dengan begitu meningkatkan kontinuitas pada titik awal dan akhir sinyal.



Gambar 2.4 *Windowing*

Berilut rumus yang digunakan[15]:

$Y(n) = X(n) \times W(n), 0 \leq n \leq N - 1$ ……..….(3)

dimana:

$Y(n)$ = sinyal output atau singal hasil windowing ke-n

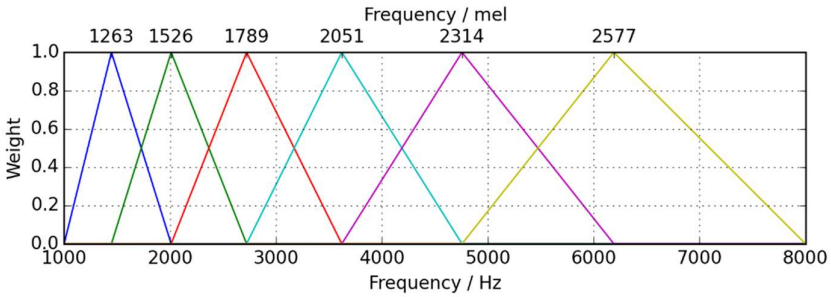
$W(n)$ = fungsi window

$X(n)$ = sinyal input atau signal ke-n hasil frame signal ke i

N = jumlah sampel dalam frame

2.3.4 *Mel Filterbank*

Mel-filterbank bertujuan untuk menentukan batas atas dan bawah dari filter. Skala *mel-frequency* adalah frekuensi linier berada dibawah 1000 Hz dan bentuk logaritmik berada diatas 1000 Hz[16]. Tujuan dari *filter bank* yaitu untuk menentukan ukuran energi dari suatu frekuensi-frekuensi tertentu, tetapi untuk tujuan MFCC, penerapan *filter bank* harus dilakukan pada *frequency domain*. upaya *filter bank* adalah dengan melakukan konvolusi representasi dalam melakukan filter terhadap sinyal. Konvolusi dapat terjadi dengan melakukan perkalian antara spektrum sinyal dengan koefisien *filter bank*.



Gambar 2.5 *Mel-Filterbank*

Rumus *mel-filterbank* ialah

$$mel(f) = 2595 \times log_{10}\left(1 + \frac{f}{700}\right) \text{.....(4)}$$

2.3.5 *Fast Fourirer Transform*

Fast Fourier transformation (FFT) adalah algoritma bekerja untuk menguraikan sinyal diskrit menjadi gelombang sinusoidal[17]. Untuk menguraikan sebuah sinyal maka digunakanlah FFT, dengan begitu untuk memproses suara menjadi lebih sederhana dikarenakan menyesuaikan dengan indra pendengaran manusia. FFT sebuah teknik transformasi yang mengimplementasikan *discrete fourier transform* (DFT). DFT mentransformasikan sejumlah frame terhadap sampel N, Dari time domain ke frequency domain. berikut persamaan dari FFT

$$X(n) = \sum_{K=0}^{N-1} X_k e^{-2jkn} \text{ .....(5)}$$

dimana:

N = jumlah dari frame *segment*

X = nilai data hingga k

n = 0, 1, 2, 3, …, N-1

k = 0, 1, 2, 3, …, N-1

X_k = deretan aperiodik dengan nilai N

2.3.6 *Discrete Consine Transform*

Discrete Consine Transform yaitu proses merubah konversi *frequency domain* ke *time domain* dengan menggunakan DCT[18]. Tujuannya adalah untuk mendapatkan nilai koefisien dari hasil perkalian *mel-filterbank* yang sudah dikonversi ke dalam time *domain frequency*. Hasil perkalian log dari time domain menghasilkan *mel-frequency cepstral coefficients* (MFCC) dengan menggunakan persamaan berikut

$$Cn = \sum_{k=1}^{K} (logSk) \cos\left[n\left(k - \frac{1}{2}\right)\frac{\pi}{K}\right] \quad \text{.....(6)}$$

dimana:

Cn = koefisien MFCC

Sk = *power spectrum* mel frekuensi

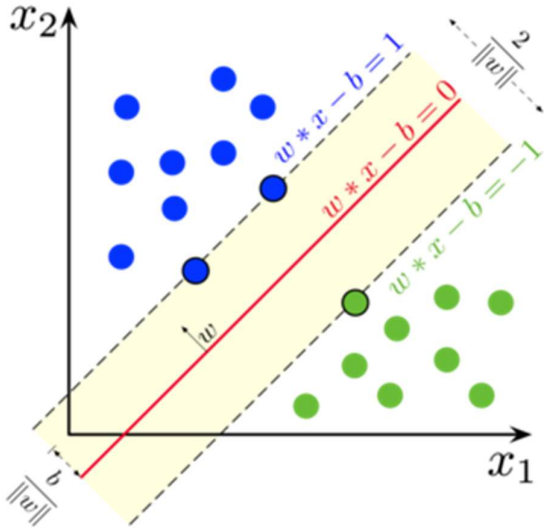
K = jumlah *filter bank*

2.4 *Support Vector Machine*

Machine learning merupakan bagian dari kecerdasan buatan yang memungkinkan sistem komputer untuk belajar dari data dan meningkatkan performanya secara otomatis tanpa deprogram secara eksplisit. Dalam pengenalan suara, *machine learning* memiliki peran penting dalam membangun model yang mampu mengenali pola suara berdasarkan ciri-ciri tertentu. Proses pembelajaran

dilakukan dengan memberikan data latih berisi fitur (input) dan label (target), sehingga algoritma dapat memahami hubungan antar data untuk melakukan klasifikasi[19].

Salah satu algoritma machine learning yang banyak digunakan untuk klasifikasi adalah *Support Vector Machine* (SVM). SVM merupakan suatu teknik klasifikasi data dengan proses pelatihan yang bersifat *supersived learning*. Prinsip dasar dari SVM yaitu dengan klasifikasi pada sebuah fungus pemisah atau yang umum disebut *hyperplane* yang berbentuk sebuah model linear dan akan membentuk *decision boundary* dengan cara memaksimalkan *margin*[20]. Margin merupakan jarak antara *hyperplane* tersebut dengan titik data yang paling dekat dari setiap kelas. Titik yang paling dekat dengna margin ini disebut dengan *support vector*.



Gambar 2.6 SVM Data Terpisah Secara Linear

Metode SVM telah dikembangkan untuk memecahkan masalah data yang bersifat *non-linear* dengan memanfaatkan fungsi kernel ke dalam ruang kerja (*workspace*) dengan dimensi yang tinggi[16]. Perumpamaan terdapat *N* data training dengan setiap inputan x_i berdimensi *d* yang dinotasikan dengan $x_i \in R^d$, untuk label pada setiap kelas dinotasikan dengan$x_i \in \{-1, +1\}$ untuk $i = 1.2. \ldots, N$ yang mana *N* adalah jumlah data input yang dipergunakan. Apabila diasumsikan terdapat dua

buah kelas -1 dan +1 dapat dipisahkan oleh *hyperplane* yang memiliki dimensi *d*, dapat didefinisikan sebagai:

$$w.x + b = 0$$

Keterangan:

w = vektor bobot

x = data input

b = bias atau scalar

Untuk data input x_i, merupakan kelas +1 yang dapat dirumuskan dengan memenuhi pertidaksamaan sebagai berikut:

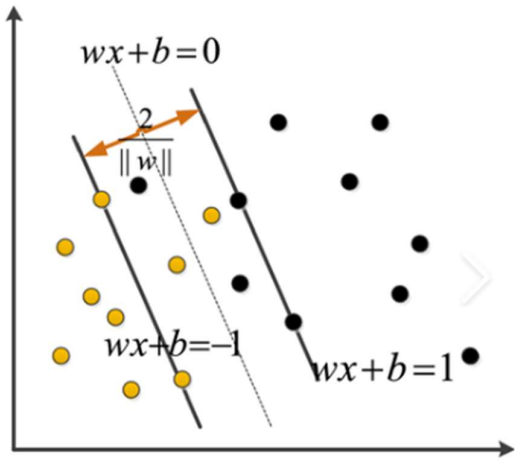
$$x_i, w + b \geq +1 \text{ untuk } y_i = 1$$

Dan untuk data input x_i, yang merupakan kelas -1 dapat digunakan rumus yang memenuhi pertidaksamaan sebagai berikut:

$$x_i, w + b \geq -1 \text{ untuk } y_i = -1$$

Jika digabungkan maka akan menjadi persamaan tersebut akan menjadi$y_i(w \times x_i + b) \geq 1$

Persamaan itu berlaku untuk semua titik data x_i dimana y_i adalah label kelas yang bernilai +1 atau -1.



Gambar 2.7 SVM Data Non-Linier

SVM *Non-Linear* digunakan ketika data tidak dapat dipisahkan secara linier. Dalam hal ini, fungsi kernel digunakan untuk memetakan data ke ruang fitur berdimensi lebih tinggi. Fungsi kernel memungkinkan kita untuk melakukan pemisahan tanpa perlu menghitung koordinat baru secara eksplisit.

Persamaan untuk SVM Non-Linear dengan kernel dapat dituliskan sebagai:

$$f(x) = \sum_{i=1}^{n} \alpha_i y_i K(x_i x) + b$$

Keterangan:

$K(x_i x)$: fungsi kernel yang mengukur kesamaan antara titik data x_i dan x

α_i : koefisien yang dioptimalkan selama pelatihan

y_i : label kelas dari titik data

SVM mempunyai fungsi kernel yang berfungsi untuk membantu permasalahan data yang tidak memungkinkan untuk dipisahkan secara linear dengan mengubah ruang data menjadi ruang baru yang memiliki dimensi yang lebih tinggi seperti.

Adapun fungsi *kernel* yang dimaksud adalah sebagai berikut:

1. Polynomial Kernel menggunakan polinomial untuk memetakan data ke dimensi yang lebih tinggi, Adapun persamaannya adalah sebagai berikut:

   $$K(x_i, x_j) = \gamma(x_i, x_j) + r^d$$

   Keterangan:

   $K(x_i, x_j)$

   x_i, x_j = vektor input

   γ = parameter skala

   r = konstanta bias

   d = derajat polinomial

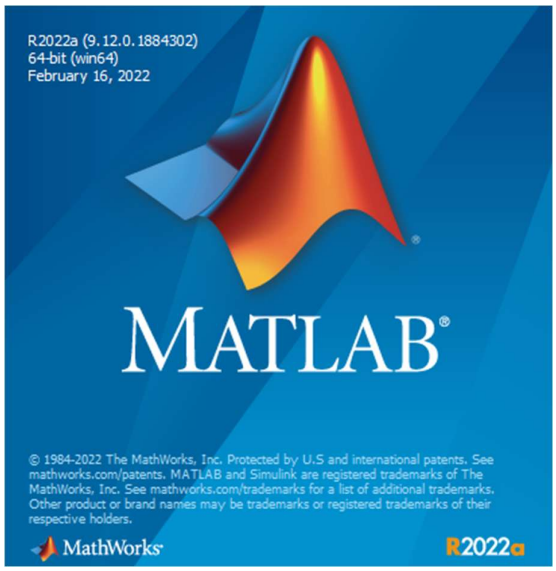
2. Radial Basis Function (RBF) Kernel, aplikasi non-linear ini mengukur jarak dari titik pusat dan menghasilkan fungsi Gaussian. Adapun persamaannya adalah sebagai beriku:

   $$K(x_i, x_j) = exp\left(-\gamma \left\| x_i - x_j \right\|^2\right)$$

$\|x_i - x_j\|^2$ = jarak *Euclidean* kuadrat antara dua vektor

Dalam penilitian ini kernel yang digunakan adalah kernel *radial basis function* (RBF).

2.5 *MATLAB*



Gambar 2.8 MATLAB

MATLAB (*Matrix Laboratory*) merupakan sebuah perangkat lunak komputasi numerik yang banyak digunakan dalam bidang teknik dan sains untuk melakukan analisis data, pemodelan sistem, serta visualisasi hasil secara interaktif. Dalam penelitian ini, MATLAB digunakan sebagai platform utama untuk mengembangkan sistem pengenalan suara, karena kemampuannya dalam menangani pengolahan sinyal digital, termasuk sinyal suara. MATLAB menyediakan berbagai fungsi bawaan yang mendukung proses seperti pembacaan file audio (*audioread*), perekaman suara (*audiorecorder*), pemrosesan sinyal, serta analisis frekuensi melalui *Fast Fourier Transform* (FFT).

Selain itu, MATLAB juga digunakan dalam proses ekstraksi fitur suara menggunakan metode *Mel-Frequency Cepstral Coefficients* (MFCC), serta pelatihan model klasifikasi menggunakan algoritma *Support Vector Machine* (SVM). Kelebihan MATLAB terletak pada konsistensi hasil komputasinya, kemudahan dalam visualisasi data fitur, serta integrasi dengan berbagai toolbox yang relevan untuk keperluan machine learning dan digital signal processing.

Dengan dukungan tersebut, MATLAB menjadi lingkungan yang efisien dan andal untuk pengembangan serta pengujian sistem pengenalan suara berbasis fitur dan klasifikasi[21].

2.6 *Principal Component Analysis* (PCA)

Principal Component Analysis (PCA) merupakan salah satu metode reduksi dimensi yang digunakan untuk menyederhanakan representasi data tanpa menghilangkan informasi penting di dalamnya. PCA bekerja dengan cara mengubah sekumpulan variabel yang saling berkorelasi menjadi kumpulan variabel baru yang tidak saling berkorelasi, yang disebut dengan *principal components*. Setiap *principal component* dibentuk sedemikian rupa sehingga mampu menjelaskan variasi data terbesar dari komponen sebelumnya. Melalui pendekatan ini, data berdimensi tinggi dapat direpresentasikan ke dalam dimensi yang lebih rendah, namun tetap mempertahankan sebagian besar karakteristik penting dari data awal.[22]

Proses PCA dimulai dengan menstandarkan data, kemudian menghitung matriks kovarians untuk mengetahui hubungan antarfitur. Selanjutnya, dilakukan perhitungan nilai eigen dan vektor eigen dari matriks kovarians tersebut. Nilai eigen menunjukkan besarnya kontribusi setiap vektor eigen terhadap total variansi data. Vektor eigen dengan nilai eigen terbesar dipilih sebagai *principal component*, karena komponen tersebut merepresentasikan arah dengan variasi data terbesar. Dengan menyeleksi beberapa *principal components* teratas, data baru dapat dibentuk dengan dimensi yang lebih sedikit namun tetap informatif.

Dalam penelitian ini, PCA digunakan pada tahap *pre-processing* setelah proses ekstraksi fitur suara menggunakan metode MFCC. Hasil ekstraksi MFCC menghasilkan sejumlah besar koefisien yang merepresentasikan karakteristik suara setiap pembicara. Jumlah fitur yang tinggi tersebut dapat meningkatkan beban komputasi dan berpotensi menyebabkan terjadinya *overfitting* pada proses klasifikasi menggunakan Support Vector Machine (SVM). Oleh karena itu, PCA diterapkan untuk mereduksi dimensi fitur MFCC dengan mempertahankan variasi data yang paling signifikan. Dengan cara ini, data suara dapat direpresentasikan

secara lebih efisien, proses pelatihan SVM menjadi lebih cepat, serta sistem pengenalan pembicara dapat mencapai tingkat akurasi yang lebih optimal tanpa kehilangan informasi penting dari sinyal suara asli.

2.7 *Mic Fifine K669B*

Gambar 2.9 *Microphone Fifine* K669B

Mikrofon Fifine K669B merupakan mikrofon kondensor dengan konektivitas USB yang dirancang untuk keperluan perekaman suara dengan kualitas tinggi. Mikrofon ini memiliki karakteristik *cardioid polar pattern*, yang artinya hanya menangkap suara dari arah depan dan meminimalkan gangguan dari samping atau belakang, sehingga sangat efektif untuk digunakan dalam sistem pengenalan suara berbasis komputer. Dalam penelitian ini, Fifine K669B digunakan untuk merekam suara pengguna secara langsung sebagai data input sistem pengenalan suara, karena mampu menghasilkan rekaman dalam format WAV 16-bit/44.1 kHz, yang sesuai dengan standar untuk ekstraksi ciri menggunakan metode MFCC. Fifine K669B juga mendukung *plug-and-pla***y** tanpa perlu sound card tambahan, sehingga memudahkan integrasi langsung dengan MATLAB. Berdasarkan hasil pengujian dan ulasan teknis, mikrofon ini menawarkan rasio *signal-to-noise* (SNR) yang cukup tinggi, sehingga menghasilkan suara yang jernih dan *minim noise* faktor yang sangat penting dalam sistem *speaker recognition*.

BAB III
METODE PENELITIAN

3.1 Waktu dan Tempat Penelitian

Adapun pelaksanaan penelitian ini dilakukan di Laboratorium Elektronika Jurusan Teknik Elektro Universitas Lampung. Adapun pelaksanaanya yaitu bulan Maret 2025 sampai dengan bulan Juli 2025.

3.2 Alat dan Bahan Penelitian

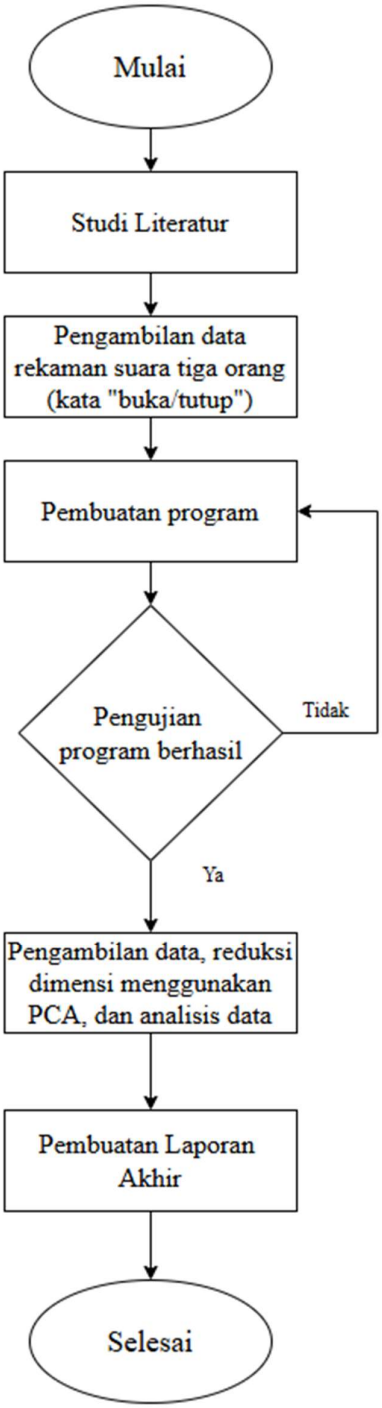
Adapun alat dan bahan yang digunakan dalam penelitian ini dapat dilihat pada tabel 3.1

Tabel 3.1 Alat dan Bahan

Alat dan Bahan	Keterangan Penggunaan
Laptop MSI G10	Sebagai perangkat keras untuk tempat program dan pengujian
Software Matlab	Sebagai perangkat lunak membuat perangcangan sistem.
Mic *Fifine K669B*	Sebagai perangkat keras untuk input suara

3.3 Diagram Alir Penelitian

Adapun tahapan yang dilakukan pada penelitian ini digambarkan melalui diagram alir pada Gambar 3.1.



Gambar 3.1 Diagram Alir Penelitian

Diagram alir pada Gambar 3.1 menggambarkan tahapan sistematis dalam pelaksanaan penelitian ini, yang bertujuan untuk mengembangkan sistem pengenalan suara berbasis metode MFCC dan algoritma klasifikasi SVM dengan kernel RBF. Proses penelitian dimulai dari tahap studi literatur, yang dilakukan untuk mengidentifikasi teori-teori dasar terkait pengolahan sinyal suara, ekstraksi fitur menggunakan *Mel Frequency Cepstral Coefficients* (MFCC), serta metode klasifikasi dengan *Support Vector Machine* (SVM), termasuk pemanfaatan kernel non-linear seperti *Radial Basis Function* (RBF). Studi ini juga mencakup penelaahan berbagai pendekatan yang relevan dalam sistem autentikasi berbasis biometrik suara.

Tahapan berikutnya adalah pengambilan data rekaman suara, yang merupakan langkah krusial dalam penyusunan dataset. Pada tahap ini, dilakukan perekaman suara dari tiga orang pembicara yang telah ditentukan, yaitu Yandi, Gibran, dan Faris. Masing-masing pembicara mengucapkan dua kata perintah, yaitu “buka” dan “tutup”, dalam kondisi lingkungan yang dikontrol agar konsisten. Proses perekaman menggunakan mikrofon eksternal dengan kualitas memadai untuk memastikan akurasi ekstraksi fitur. Data rekaman ini selanjutnya digunakan sebagai input untuk pelatihan dan pengujian sistem klasifikasi suara.

Setelah data terkumpul, tahap selanjutnya adalah pembuatan program yang mencakup proses ekstraksi fitur suara menggunakan MFCC, pelatihan model klasifikasi menggunakan SVM dengan kernel RBF. Program diuji untuk memastikan bahwa setiap proses berjalan sesuai dengan tujuan sistem. Jika pada tahap pengujian program ditemukan bahwa program belum berjalan dengan tidak benar maka dilakukan perbaikan dan penyempurnaan program hingga pengujian dapat berjalan dengan baik.

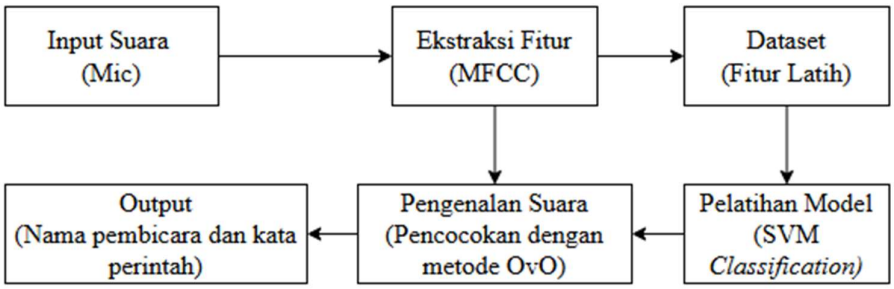
Apabila program telah berhasil diuji, penelitian dilanjutkan pada tahap pengambilan data hasil pengujian dan analisis data. Analisis dilakukan terhadap hasil klasifikasi suara baik dari sisi akurasi sistem secara keseluruhan maupun dari masing-masing pembicara. Selain itu, analisis visual dari hasil PCA juga digunakan untuk memahami distribusi data serta mendeteksi potensi tumpang tindih antar kelas suara, yang dapat memengaruhi performa klasifikasi. Termasuk dalam

analisis ini adalah pembuatan *confusion matrix* untuk setiap pembicara guna mengevaluasi ketepatan sistem dalam mengenali kata perintah.

Tahap akhir dari proses penelitian adalah pembuatan laporan akhir, yang merangkum seluruh kegiatan mulai dari perancangan, implementasi, pengujian, hingga evaluasi sistem. Laporan ini disusun secara sistematis sebagai bentuk pertanggungjawaban ilmiah dari hasil penelitian. Setelah seluruh proses selesai, penelitian ini diakhiri dengan simpulan dan saran pengembangan untuk penelitian selanjutnya.

3.4 Diagram Blok Perancangan Sistem

Adapun diagram blok pada penelitian ini ialah sebagai berikut.



Gambar 3.2 Diagram Blok Sistem

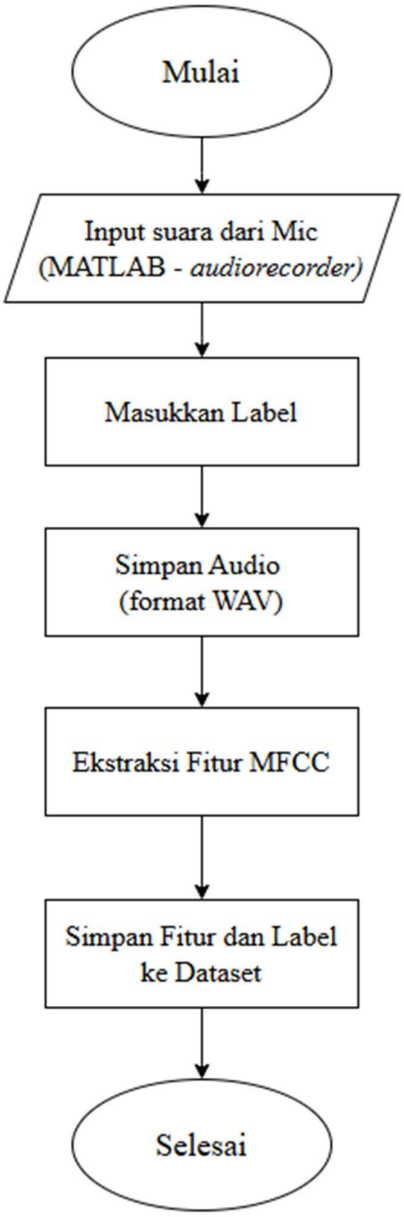
Diagram blok di atas menggambarkan alur kerja sistem pengenalan pembicara berbasis MFCC dan SVM yang dibangun dalam penelitian ini. Proses dimulai dari input suara yang direkam secara langsung menggunakan mikrofon.

Kemudian data suara memasuki tahap ekstraksi fitur menggunakan *Mel-Frequency Cepstral Coefficients* (MFCC). Metode MFCC digunakan untuk mengambil ciri-ciri penting dari suara yang mencerminkan karakteristik unik masing-masing pembicara dan kata perintah yang diucapkan. Hasil ekstraksi fitur ini kemudian disimpan ke dalam dataset yang berisi fitur-fitur suara beserta label identitas pembicara dan kata perintah yang diucapkan. Dataset ini digunakan untuk melatih model SVM agar sistem dapat mempelajari pola-pola dari data suara yang ada.

Setelah proses pelatihan selesai, sistem masuk ke tahap pengenalan suara, di mana fitur dari suara baru yang diinputkan akan dicocokkan dengan model yang telah dilatih sebelumnya. Apabila pola fitur suara yang diinputkan sesuai dengan model yang ada, sistem akan menampilkan output berupa nama pembicara dan kata perintah yang dikenali. Sebaliknya, jika suara tidak dikenali, maka sistem akan memberikan output "suara tidak dikenal".

3.5 Diagram Alir Sistem Data *Train*

Adapun tahapan sistem data *train* pada penelitian ini ialah sebagai berikut.



Gambar 3.3 Diagram Alir Sistem Data *Train*

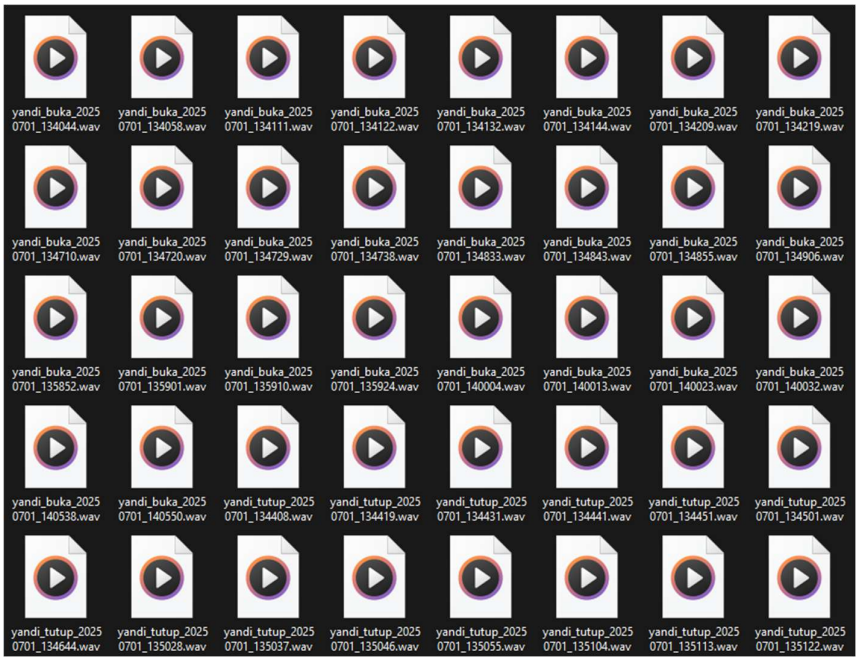
Proses dimulai dengan tahapan pengambilan input suara secara langsung dari mikrofon menggunakan fungsi *audiorecorder* di MATLAB. Fitur ini memungkinkan sistem untuk merekam suara pengguna secara real-time dengan durasi dan pengaturan tertentu.

Setelah perekaman, pengguna diminta untuk memasukkan label yang berfungsi sebagai identifikasi dari data suara yang direkam. Label ini mengandung informasi penting seperti nama pembicara, kata perintah, dan *timestamp*, misalnya: yandi_buka_20250701_134122.

Langkah berikutnya adalah penyimpanan suara dalam format .wav, karena format ini umum digunakan dalam pengolahan sinyal suara, dan didukung penuh oleh MATLAB.

Setelah file audio tersimpan, sistem melakukan proses ekstraksi fitur MFCC (*Mel-Frequency Cepstral Coefficients*). Tahapan ini merupakan inti dari pengolahan sinyal suara, di mana ciri-ciri unik dari suara seperti intonasi, timbre, dan artikulasi diubah menjadi representasi numerik berbentuk vektor.

Setelah ekstraksi selesai, fitur MFCC dan label yang telah ditetapkan disimpan ke dalam dataset. Dataset ini akan digunakan sebagai data latih untuk algoritma klasifikasi (seperti SVM).



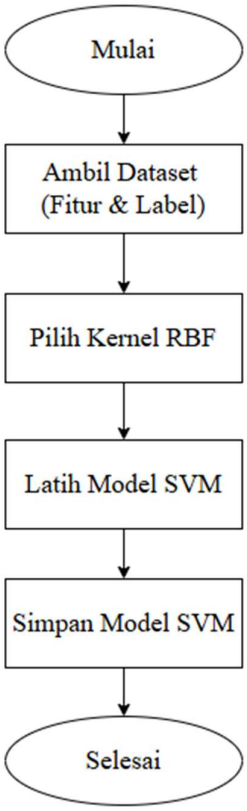
Gambar 3.4 Data Train

Pada gambar 3.4 merupakan sekumpulan file suara untuk data train. Sebagai ilustrasi, data latih yang digunakan berasal dari tiga orang pembicara: Yandi, Gibran, dan Faris. Pada gambar yang ditampilkan, terlihat kumpulan file audio latih milik Yandi yang tersimpan dalam folder SuaraYandi. Meskipun nama folder mencerminkan salah satu pembicara, sistem secara keseluruhan mencakup beragam data latih dari seluruh pengguna agar proses pelatihan model dapat mengenali variasi suara dan perintah secara umum. Setiap file suara memiliki format penamaan yang konsisten untuk memudahkan proses ekstraksi label otomatis saat membangun dataset.

Dengan tahapan ini, sistem memiliki fondasi data yang kuat untuk dilatih dan diuji secara berulang guna membangun akurasi yang tinggi dalam mengenali siapa pembicara dan apa perintah yang diucapkan.

3.6 Diagram Alir Pemodelan SVM

Adapun diagram alir pemodelan SVM pada penelitian ini ialah sebagai berikut.



Gambar 3.5 Diagram Alir Pemodelan SVM

Diagram alir di atas menjelaskan tahapan-tahapan dalam proses pemodelan *Support Vector Machine* (SVM) yang digunakan dalam sistem pengenalan suara berbasis klasifikasi pembicara dan perintah. Proses diawali dengan pengambilan dataset yang terdiri dari fitur hasil ekstraksi *Mel Frequency Cepstral Coefficients* (MFCC) serta label yang telah diberikan secara terstruktur. Fitur-fitur tersebut merepresentasikan karakteristik akustik dari sinyal suara, sedangkan label digunakan sebagai acuan dalam proses pembelajaran. Dataset yang digunakan mencakup data suara dari tiga pembicara berbeda, yakni Yandi, Gibran, dan Faris. Masing-masing pembicara mengucapkan dua jenis perintah, yaitu “buka” dan “tutup”, sehingga total terdapat enam kelas kategori dalam data, yaitu: yandi_buka, yandi_tutup, gibran_buka, gibran_tutup, faris_buka, dan faris_tutup.

Setelah dataset tersedia, proses dilanjutkan dengan pemilihan kernel yang akan digunakan dalam pelatihan model SVM. Pada sistem ini, digunakan kernel *Radial Basis Function* (RBF) karena sifatnya yang sangat cocok untuk menangani data non-linear. Kernel RBF mampu memetakan data input ke dalam ruang berdimensi lebih tinggi, sehingga batas pemisah antar kelas dapat ditemukan secara optimal. Kemampuan kernel ini dalam menangani pola kompleks sangat relevan untuk data suara yang memiliki banyak variasi frekuensi dan intensitas.

Langkah selanjutnya adalah melatih model SVM menggunakan dataset yang telah disiapkan. Dalam tahap ini, SVM akan mempelajari hubungan antara fitur MFCC dan label enam kelas tersebut, sehingga terbentuk model klasifikasi yang mampu mengenali dan membedakan suara berdasarkan pembicara serta kata perintahnya. Proses pelatihan bertujuan untuk menentukan hyperplane non-linear terbaik yang dapat memisahkan masing-masing kelas dengan akurasi setinggi mungkin.

Setelah proses pelatihan selesai, model yang dihasilkan akan disimpan ke dalam file agar dapat digunakan kembali pada tahap pengujian atau implementasi sistem secara real-time. Penyimpanan model ini penting untuk efisiensi, karena memungkinkan proses klasifikasi langsung tanpa perlu melatih ulang model setiap kali digunakan. Tahap terakhir menunjukkan bahwa keseluruhan proses pemodelan telah selesai, dan sistem siap untuk melakukan pengenalan suara terhadap input baru dengan memanfaatkan pola-pola yang telah dipelajari sebelumnya

BAB V
KESIMPULAN DAN SARAN

5.1 Kesimpulan

Adapun Kesimpulan yang diperloeh berdasarkan penelitian ini adalah sebagai berikut:

1. Sistem speaker recognition berhasil dikembangkan menggunakan metode MFCC untuk ekstraksi ciri suara dan SVM sebagai algoritma klasifikasi multi-kelas, yang mampu mengenali perintah serta identitas pembicara.
2. Pengujian sistem dilakukan dalam dua kondisi, yaitu tanpa *noise* dan dengan *noise*. Hasil menunjukkan akurasi rata-rata sebesar 63,3% pada kondisi ideal dan turun menjadi 41,6% pada kondisi berisik, menandakan penurunan performa yang signifikan akibat gangguan lingkungan.
3. Efektivitas sistem diukur dari kemampuannya mempertahankan akurasi antar kondisi. Dengan penurunan akurasi sebesar 21,7%, dapat disimpulkan bahwa kombinasi MFCC dan SVM belum efektif dalam menghadapi variasi kondisi lingkungan jika model hanya dilatih dengan data dari kondisi ideal.

5.2 Saran

Adapun saran yang dapat diajukan dari penelitian ini sebagai berikut:

1. Sistem sebaiknya dilatih menggunakan data suara dari berbagai kondisi lingkungan, tidak hanya dari kondisi ideal. Pelatihan dengan data yang mengandung noise alami seperti suara lingkungan sekitar akan membantu SVM membentuk batas keputusan yang lebih adaptif terhadap gangguan suara

2. Teknik augmentasi seperti penambahan noise sintetis (*white noise*, *background noise*) pada data latih yang sudah ada dapat memperluas cakupan fitur tanpa perlu merekam ulang. Ini akan membantu meningkatkan kemampuan generalisasi model saat menghadapi kondisi dengan *noise*.

DAFTAR PUSTAKA

[1] K. K. Rekayasa, F. Dharma Adhinata, D. Putra Rakhmadani, A. Jala, dan T. Segara, "Terbit online pada laman web jurnal: http://journal.ittelkom-pwt.ac.id/index.php/dinda JURNAL DINDA Pengenalan Jenis Kelamin Manusia Berbasis Suara Menggunakan MFCC dan GMM," 2021.

[2] I. Sidik Permana, Y. I. Nurhasanah, dan A. Zulkarnain, "Implementasi Metode Mfcc Dan Dtw Untuk Pengenalan Jenis Suara Pria Dan Wanita," *MIND Journal | ISSN*, vol. 3, no. 1, hlm. 49–63, 2018.

[3] P. Prasetyawan, "Perbandingan Identifikasi Pembicara Menggunakan Mfcc Dan Sbc Dengan Ciri Pencocokan LBG-VQ," 17 Februari 2018.

[4] F. Fauziah, I. I. Tritoasmoro, dan S. Rizal, "Sistem Keamanan Berbasis Pengenalan Suara Sebagai Pengakses Pintu Menggunakan Metode Mel Frequency Cepstral Coefficient (Mfcc) Door Entry Using Voice Recognition Security System While Untilizing Mel Frequency Cepstral Coefficient (Mfcc)." e-Proceeding of Engineering: Vol.8, No.6 Desember 2021

[5] G. Ajinurseto, L. O. Bakrim, dan N. Islamuddin, "Penerapan Metode Mel Frequency Cepstral Coefficients pada Sistem Pengenalan Suara Berbasis Desktop," *Infomatek*, vol. 25, no. 1, hlm. 11–20, Jun 2023.

[6] A. A. Sundawa, A. G. Putrada, dan N. A. Suwastika, "Implementasi dan Analisis Simulasi Deteksi Emosi Melalui Pengenalan Suara Menggunakan Mel-Frequency Cepstrum Coefficient dan Hidden Markov Model Berbasis IOT." e-Proceeding of Engineering: Vol.6, No.1 April 2019

[7] N. Florencia, M. Tatiana, dan D. Alamsyah, "Speaker Recognition Menggunakan MFCC dan Algoritma DTW," *July 2020*, vol. x, No.x, hlm. 1–5.

[8] Shambhavi. S. S dan D. N. Nitnaware, "Emotion Speech Recognition using MFCC and SVM." International Journal of Engineering Research and Technology (IJERT) Vol.4 Issue 6, June 2015

[9] I. Suara Pada Sistem Presensi, Y. Fitri Widodo, dan A. Fadlil, "Identifikasi Suara Pada Sistem Presensi Karyawan Dengan Metode Ekstraksi MFCC," 2019.

[10] S. M. Kamruzzaman, S. C. College, M. S. Islam, M. E. Haque, S. M. Kamruzzaman, dan A. N. M. R. Karim, "Speaker Identification using MFCC-Domain Support Vector Machine," 2010.

[11] T. Bäckström, "Signals and Communication Technology Speech Coding with Code-Excited Linear Prediction." ISSN 1860-4862, ISSN 1860-4870 (electronic), ISBN 978-3-319-50202-1, ISBN 978-3-319-50204-5 (eBook), 2017

[12] Jerome Sueur, “preemphasis: Pre-emphasis speech filter,” 19 October 2023

[13] S. Santoso, R. Hartayu, C. Anam, dan Dimas Abdul Aziz, “Simulasi Simulasi Ekstraksi Fitur Suara menggunakan Mel-Frequency Cepstrum Coefficient,” *Jurnal Sains dan Informatika*, vol. 8, no. 1, hlm. 80–87, Jul 2022.

[14] M. Ahied, “Teknik Pengenalan Sinyal Wicara Dengan Metoda Hidden Markov Model-Neural Network (Syaraf Tiruan),” 2016.

[15] K. M. M Prabhu, “Window Functions and Their Rpplications in Signal Processing.” CRC Press, Taylor & Francis Group, Boca Raton, FL 33487-2742 --dec 2023

[16] S. Helmiyah, A. Fadlil, dan Anton Yudhana, dan U. Ahmad Dahlan Yogyakarta, “Ekstraksi Ciri Emosi Manusia Berdasarkan Ucapan Menggunakan Mel-Frequency Cepstral Coefficients (Mfcc).” Prosiding SNST 2018

[17] R. A. Fauzi, “Simulasi Control Smart Home Berbasis Mel Frequency Cepstral Coefficients Menggunakan Metode Support Vector Machine (SVM).,” Jul 2020.

[18] R. Fahmi, N. Imanudin, I. Kustiawan, dan S. Elvyanti, “Steganografi Citra Digital Menggunakan Pendekatan Least Significant Bit dan Discrete Cosine Transform.” 17 November 2021

[19] Hemant A. Patil and S. Elgar, “Voice-based Human-Machine Interaction using Machine Learning: Review and Future Directions,” IEEE, 2016, hlm. 1–6.

[20] A. Mansour dan Z. Lachiri, “SVM based Emotional Speaker Recognition using MFCC-SDC Features,” 2017.

[21] E. S. Gopi, *Digital Speech Processing Using MATLAB*. Springer, 2013.

[22] Hediyati et al. n. d. "Penerapan Principal Component Analysis (PCA) Untuk Reduksi Dimensi Pada Proses Clustering Data Produksi Pertanian Di Kabupaten Bojonegoro" JIEET, Vol 05 No. 02, 2021.