JURNAL SINTAKS LOGIKA Vol. 1 No. 2 , Mei - 2021,
https://jurnal.umpar.ac.id/index.php/sylog

E-ISSN : 2775-412X

SISTEM PENGENALAN UCAPAN BAHASA DAERAH MENGGUNAKAN METODE MEL FREQUENCY CEPSTRAL COEFFICIENT (MFCC) DAN ADAPTIVE NEURO-FUZZY INFERENCE SYSTEM (ANFIS)

Ade Hastuty[1] , Muh. Basri[2], Asfiad Amir[3]

[1,2,3]Program Studi Teknik Informatika, Universitas Muhammadiyah Parepare, Indonesia
asfiad.amir@gmail.com

Informasi Artikel

Riwayat Artikel:

Dikirim *Author* : 04-04-2021
Diterima Redaksi : 05-04-2021
Revisi *Reviewer*: 15-04-2021
Diterbitkan *online*: 05-05-2021

Keywords:

Bugis Language, dialects, classification, recognition, MFCC, ANFIS

ABSTRACT

Bugis language is one of the traditional languages in South Sulawesi which has various dialects. Different ways of speaking and tone of voice that cause voice of each person also vary. The design of the Bugis language speech recognition system is one way to digitize and indicate the Bugis regional language. This study uses Mel Frequency Cepstral Coefficient (MFCC) method as a feature extraction and Adaptive Neuro-Fuzzy Inference System (ANFIS) as an algorithm for classification and recognition. This research was conducted by collecting various voice data from various age groups and regions. Voice data is collected in wav audio format and then extracted into numerical results to use as database on system. Test was carried out using various database models and the amount of voice data. The test results show that databases with single data produce a higher level of accuracy with a percentage of truth above 50%. While the test results with a database of 37 speakers with a total data of 740 have a low level of accuracy and are below 50%. Testing in outdoors has a lower accuracy around 10%. Error value in the data training process in the range of 0.5 - 1.5.

Kata kunci:

Bahasa Bugis, dialek, pengelompokan, pengenalan, MFCC, ANFIS.

ABSTRAK

Bahasa bugis merupakan salah satu bahasa daerah yang ada di Provinsi Sulawesi Selatan yang memiliki beragam dialek. Perbedaan cara penuturan dan intonasi suara menyebabkan karakteristik suara dari setiap orang juga bervariasi. Perancangan sistem pengenalan ucapan bahasa daerah bugis merupakan salah satu cara untuk melakukan digitalisasi dan identifikasi terhadap bahasa daerah bugis. Penelitian ini menggunakan metode Mel Frequency Cepstral Coefficient (MFCC) sebagai fitur untuk melakukan ekstraksi ciri dan Adaptive Neuro-Fuzzy Inference System (ANFIS) sebagai algoritma untuk pengelompokan dan pengenalan suara. Penelitian ini dilakukan dengan mengumpulkan berbagai data suara dari berbagi kelompok umur dan daerah. Data suara dikumpulkan dalam format wav yang kemudian diekstraksi menjadi data numerik untuk digunakan sebagai database pada sistem. Pengujian dilakukan dengan menggunakan berbagai model database dan jumlah data suara. Pengujian dilakukan oleh beberapa orang pada kondisi di dalam ruangan dan di luar ruangan. Hasil pengujian menunjukkan bahwa penggunaan database dengan data tunggal menghasilkan tingkat akurasi yang lebih tinggi dengan persentase kebenaran di atas 50%. Sedangkan hasil pengujian dengan database 37 penutur dengan jumlah data 740 memiliki tingkat akurasi yang rendah dan berada di bawah 50%. Pengujian yang dilakukan di luar ruangan menghasilkan presentase yang lebih rendah yaitu sekitar 10%. Nilai error yang dihasilkan pada proses pelatihan database suara berada pada rentang 0,5 – 1,5.

Penulis Korespondensi:

Asfiad Amir
Program Studi Teknik Informatika
Universitas Muhammadiyah Parepare,
Jl.Jend.Ahmad Yani Km.6, Parepare
Email: asfiad.amir@gmail.com

This is an open access article under the CC BY-SA license.

I. PENDAHULUAN

Indonesia merupakan negara kepulauan yang terdiri dari ribuan pulau yang terbentang dari Sabang sampai Merauke. Letak geografis ini juga berdampak pada keragaman suku yang dimiliki oleh Indonesia. Hal ini juga menyebabkan Indonesia memiliki ratusan bahasa daerah. Berdasarkan data Badan Pengembangan dan Pembinaan Bahasa Kementerian Pendidikan dan Kebudayan jumlah bahasa daerah yang telah diverifikasi adalah 652 tetapi ini belum termasuk dialek dan subdialek.

Bahasa merupakan salah satu unsur kebudayaan yang menjadi cara manusia untuk melakukan komunikasi, bersosialisasi dan mengekspresi-kan diri. Dalam perkembangan penggunaan bahasa daerah khususnya bahasa daerah bugis yang ada di Sulawesi, setiap daerah memiliki dialek yang berbeda. Berdasarkan peta bahasa, suku Bugis dibagi dalam beberapa dialek yaitu dialek Luwu, dialek Wajo, dialek Palakka, dialek Ennak (Sinjai), dialek Soppeng, dialek Sidenreng, dialek Parepare, dialek Sawitto (Pinrang), dialek Tallumpanua, dialek Ugi Riawa (Pasangkayu).

Dalam pengaplikasiannya, bahasa diwujudkan dalam bentuk suara. Data yang dikirim dalam proses komunikasi secara langsung antar manusia dalam sebuah suku dengan menggunakan bahasa daerah berbentuk suara. Namun, seiring dengan perkembangan teknologi, proses komunikasi tidak hanya terjadi antar manusia saja. Suara merupakan salah satu faktor penting dari interaksi antara manusia dengan komputer.

Sinyal suara manusia memiliki tingkat variabilitas yang sangat tinggi. Setiap orang dapat menghasilkan pola ucapan yang berbeda-beda. Hal ini juga terjadi dengan pengucapan bahasa daerah bugis dengan dialek yang berbeda. Atau pengucapan bahasa daerah bugis oleh orang yang tidak berasal dari suku bugis.

Speech recognition merupakan salah satu proses pengenalan suara manusia berupa kata atau kalimat agar bisa dikenali oleh perangkat computer. Pada penelitian ini digunakan fitur ekstraksi Mel Frequency Cepstral Coefficient (MFCC) dan klasifikasi menggunakan metode Adaptive Neuro-Fuzzy Inference System (ANFIS) untuk mendapatkan tingkat kesesuaian kata dengan dialek yang berbeda. Serta untuk mendapatkan tingkat kesesuaian kata dengan penutur yang bukan suku bugis. Pada penelitian sebelumnya, metode MFCC dan ANFIS telah digunakan untuk mengontrol nyala dan mati lampu AC secara otomatis [1].

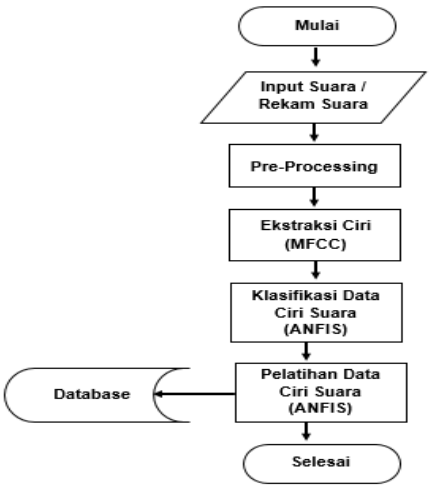
Berdasarkan pemaparan di atas, maka pada penelitian ini akan dirancang "Sistem Pengenalan Ucapan Bahasa Daerah Menggunakan Metode Mel Frequency Cepstal Coefficient (MFCC) dan Adaptive Neuro-Fuzzy Inference System (ANFIS)".

II. METODOLOGI PENELITIAN

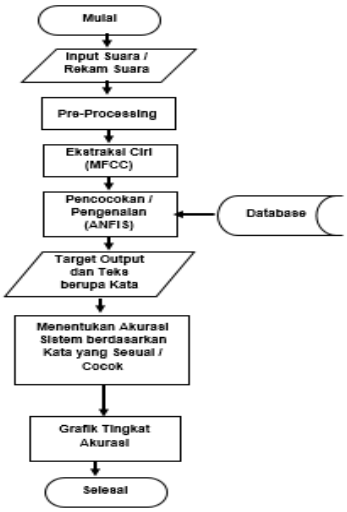
A. *Perancangan Sistem*

Perancangan sistem untuk sistem pengenalan suara bahasa daerah dengan metode MFCC dan ANFIS dilakukan dengan menggunakan software Matlab. Secara umum alur kerja dari sistem pengenalan suara yang dirancang pada penelitian ini dapat dibagi dalam dua tahapan, yaitu tahapan pelatihan suara dan tahapan pengujian.

Tahapan awal dari penelitian ini adalah pelatihan suara yaitu dengan melakukan penginputan suara. Peralatan yang digunakan dalam proses perekaman adalah mikrofon Laptop atau headset. Sinyal suara yang direkam berupa kata (1 kata) yang disimpan dalam file dengan format wav. Sinyal suara yang telah direkam akan diproses pada software MATLAB. Proses awal berupa pre-processing dengan melakukan normalisasi amplitudo / volume sinyal suara. Pada tahap ini juga dihilangkan gangguan-gangguan yang ada pada sinyal suara.



Gambar 1. Flowchart Pelatihan Suara



Gambar 2. Flowchart Pengujian Suara

Proses selanjutnya adalah ekstraksi ciri dengan menggunakan metode Mel Frequency Cepstral Coefficient (MFCC). Sinyal suara akan melalui berbagai proses pada MFCC dan menghasilkan output berupa vektor ciri dalam bentuk angka.Setelah itu, setiap suara akan diklasifikasi menggunakan metode Adaptive Neuro-Fuzzy Inference System (ANFIS). Setiap suara keluaran MFCC yang berbentuk vektor ciri akan diberikan target output berupa angka. Suara yang memiliki pengucapan yang sama akan diberikan target ouput yang sama.

Proses selanjutnya adalah pelatihan suara. Semua data vektor ciri dan target ouput akan diproses dengan ANFIS sehingga menghasilkan berbagai rule atau kemungkinan input dan ouput yang bisa terjadi. Rule ini akan menjadi pedoman pada proses pengujian/pengenalan suara. Data output ANFIS akan disimpan di database.

Tahapan kedua dari penelitian ini adalah pengujian. Proses perekaman, pre-processing dan ekstraksi ciri dengan menggunakan MFCC sama dengan metode yang digunakan pada tahap pelatihan. Objek penelitian yang digunakan adalah suara orang yang digunakan pada proses pelatihan dan suara orang yang tidak ada pada database pelatihan. Vektor ciri keluaran MFCC akan dikalkulasi dan dibandingkan dengan database dengan menggunakan ANFIS sehingga akan menghasilkan output berupa angka. Output angka yang dihasilkan akan merepresentasikan nama kata yang diucapkan sesuai dengan yang telah diatur pada tahap pelatihan. Nama kata yang dikenali akan keluar dalam bentuk teks.

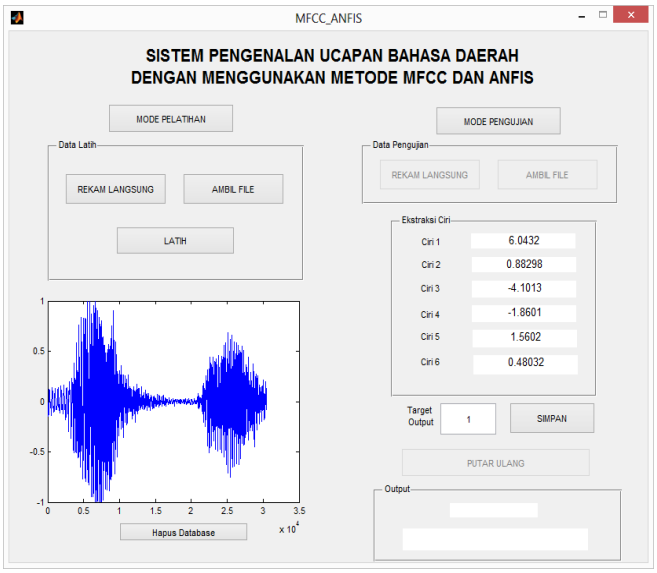
Proses selanjutnya adalah pencocokan antara kata yang diucapkan dengan teks yang keluar secara manual. Jika kata yang diucapkan sama dengan teks yang keluar maka dianggap sesuai / cocok. Tetapi jika kata yang diucapkan tidak sama dengan teks yang keluar maka dianggap tidak sesuai / cocok. Persentasi akurasi dihitung dengan membandingkan jumlah data yang sesuai / cocok dengan jumlah seluruh pengujian dikalikan dengan 100%. Hasilnya akan dibuat dalam bentuk grafik.

B. *Pelatihan Sistem*

Tahapan awal dari pelatihan data adalah pengumpulan data dengan melakukan perekaman suara. Proses perekaman suara dapat dilakukan secara langsung pada aplikasi dengan menggunakan tombol REKAM LANGSUNG di Mode Pelatihan. Untuk menggunakan fungsi AMBIL FILE, proses perekaman dapat menggunakan coding Matlab yang dilakukan terpisah dari sistem utama yang telah dirancang. File yang telah direkam dengan format WAV akan diinputkan ke dalam sistem dengan tombol AMBIL FILE.

Selanjutnya data yang diinputkan dalam bentuk perekaman langsung atau pengambilan file dari komputer akan secara otomatis dilakukan ekstraksi data ciri oleh sistem. Data ciri akan ditampilkan dalam kolom Ekstraksi Ciri seperti yang ditunjukkan pada Gambar 3.

Jumlah data ciri yang ditampilkan pada sistem ini adalah 6 ciri, yaitu Ciri 1 – Ciri 6. Setiap data suara akan menampilkan data ciri yang berbeda. Setiap data ciri akan diberikan Target Output secara manual. Nilai target output diberikan untuk memudahkan dalam proses klasifikasi suara. Setiap suara dengan pengucapan yang sama diberikan nilai Target Output yang sama.



Gambar 3. Proses Ekstraksi Ciri

Pada penelitian ini digunakan 10 kata bahasa bugis yaitu Jokka, Lao, Leko, Leppang, Lisu, Mabbaca, Mabbicara, Mammekko, Menre dan Soro. Setiap kata diberikan target ouput dari 1 – 10 seperti yang ditampilkan pada Tabel 1.

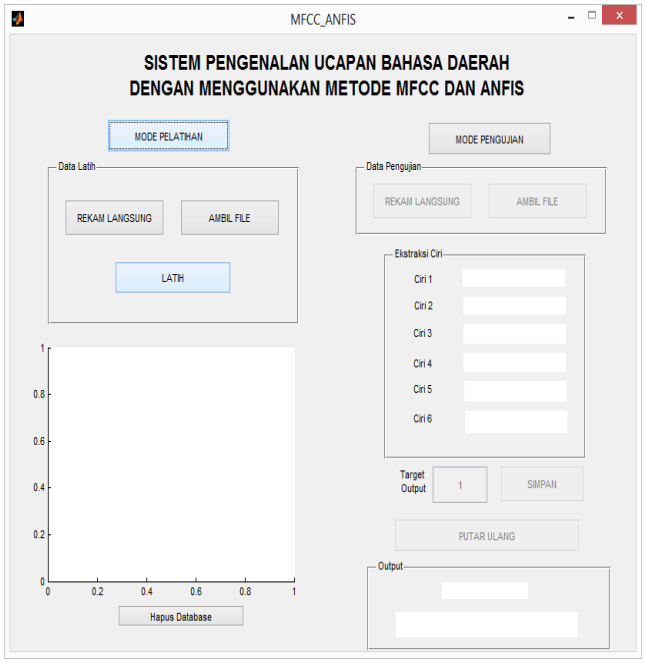
Tabel 1. Target Output

Kata	Target Output
Jokka	1
Lao	2
Leko	3
Leppang	4
Lisu	5
Mabbaca	6
Mabbicara	7
Mammekko	8
Menre	9
Soro	10

Semua file suara yang diinputkan ke dalam grup dalam bentuk REKAM LANGSUNG atau AMBIL FILE

akan disimpan sebagai database untuk data pelatihan. Database ciri suara disimpan dalam format file .dat, pada penelitian ini database ciri suara disimpan dalam nama file simpansuara.dat.

Tahapan selanjutnya adalah pelatihan data suara. Proses pelatihan data adalah menggunakan aplikasi GUI yang telah dirancang. File database ciri suara disimpan pada direktori yang sama dengan file matlab dengan nama file simpansuara.dat. Sistem akan mengenali secara langsung sesuai yang telah diatur pada coding setelah menekan tombol LATIH.



Gambar 4. Mode Pelatihan Suara

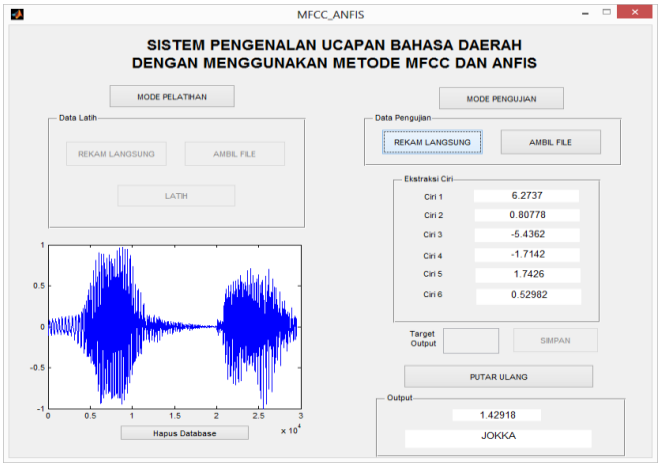
Membership function yang digunakan sudah dikonfigurasi pada code program dengan MF Numbers sebesar 3 dan MF Type adalah gauss2mf. Jumlah iterasi yang digunakan pada sistem ini adalah 50. Data yang telah dilatih akan disimpan secara otomatis pada direktori yang sama dengan nama suarasukses.fis.

Model pelatihan yang digunakan pada penelitian ini dilakukan secara manual seperti cara pertama. Pada cara pertama proses pengubahan iterasi dapat dilakukan secara mudah ketika nilai error yang didapatkan belum mencapai yang telah ditetapkan. Grafik perbandingan antara data output FIS dan data input bisa dibandingkan.

C. *Metode Pengujian Sistem*

Pengujian pada sistem pengenalan suara dilakukan dengan mengaktifkan Mode Pengujian pada Aplikasi. Metode pengujian dilakukan dengan dua cara yaitu merekam secara langsung (REKAM LANGSUNG) atau melakukan pengambilan file dari data suara yang telah direkam dalam bentuk wav (AMBIL FILE).

Salah satu cara pengujian yang dilakukan ditampilkan pada Gambar 7. Dengan menekan tombol REKAM LANGSUNG, penguji mengucapkan salah satu kata. Hasil kalkulasi data ciri dari suara yang diucapkan saat pengujian akan dibandingkan dengan data yang ada pada database. Sistem akan memberikan target output sesuai dengan data yang didapatkan pada database.



Gambar 5. Pengujian Sistem

Data output yang dihasilkan akan disandingkan dengan output kata yang telah diatur pada coding program. Pada Tabel 2 ditampilkan rentang nilai target output dan output kata.

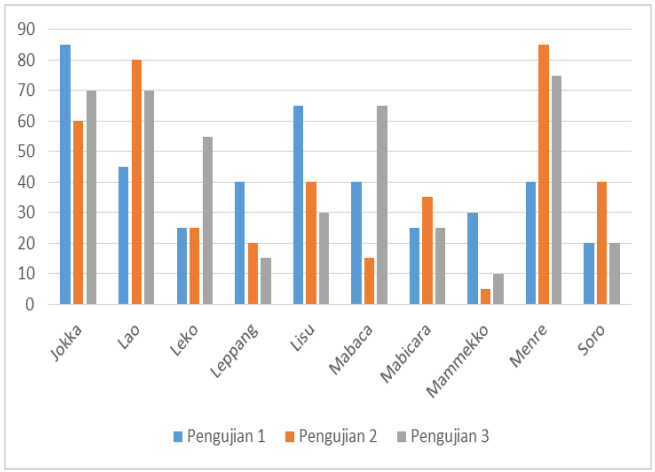
Tabel 2. Target Output pada Program

Kata	Target Output
Jokka	0,5 ≤ 1 < 1,5
Lao	1,5 ≤ 2 < 2,5
Leko	2,5 ≤ 3 < 3,5
Leppang	3,5 ≤ 4 < 4,5
Lisu	4,5 ≤ 5 < 5,5
Mabbaca	5,5 ≤ 6 < 6,5
Mabbicara	6,5 ≤ 7 < 7,5
Mammekko	7,5 ≤ 8 < 8,5
Menre	8,5 ≤ 9 < 9,5
Soro	9,5 ≤ 10 < 10,5

III. HASIL DAN PEMBAHASAN

A. *Pengujian dengan Database 1 Kata*

Pengujian ini dilakukan dengan menggunakan database 1 kata yang terdiri dari 1 orang penutur dengan jumlah data 30. Pengujian dilakukan sebanyak 20 kali dengan rincian 15 rekam langsung dan 5 dengan menggunakan file suara yang telah direkam. Pengujian pertama dilakukan dengan orang yang suaranya digunakan pada database sedangkan pengujian kedua dan ketiga dilakukan dengan orang yang berbeda.

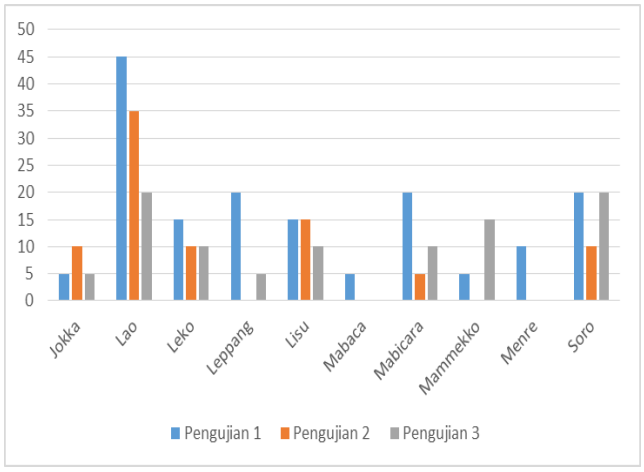


Gambar 6. Hasil Pengujian Database 1 Kata

Berdasarkan grafik diatas, pengujian kata Jokka memiliki presentasi kebenaran yang paling tinggi diantara semua kata dengan nilai sekitar 60 – 85 %. Sedangkan kata Mammekko memiliki hasil pengujian dengan rata-rata persentasi kebenaran yang lebih rendah dibandingkan dengan yang lainnya

B. *Pengujian dengan Database 5 Kata*

Pengujian ini dilakukan dengan menggunakan database 5 kata yang terdiri dari 1 orang penutur dengan jumlah data 150. Pengujian dilakukan sebanyak 20 kali dengan rincian 15 rekam langsung dan 5 dengan menggunakan file suara yang telah direkam. Pengujian pertama dilakukan dengan orang yang suaranya digunakan pada database sedangkan pengujian kedua dan ketiga dilakukan dengan orang yang berbeda.

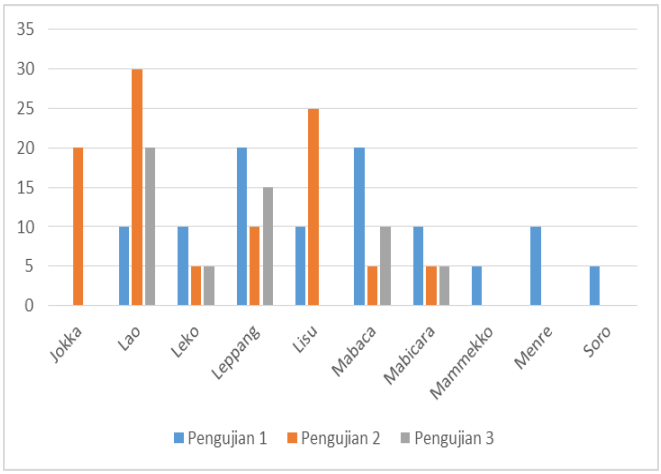


Gambar 7. Hasil Pengujian Database 5 Kata

Berdasarkan grafik diatas, pengujian kata Lao memiliki presentasi kebenaran yang paling tinggi diantara semua kata dengan nilai mencapai 45% pada pengujian pertama. Sedangkan kata Mabbaca memiliki hasil pengujian dengan rata-rata persentasi kebenaran yang lebih rendah dibandingkan dengan yang lainnya.

C. *Pengujian dengan Database 10 Kata*

Pengujian ini dilakukan dengan menggunakan database 10 kata yang terdiri dari 1 orang penutur dengan jumlah data 300. Pengujian dilakukan sebanyak 20 kali dengan rincian 15 rekam langsung dan 5 dengan menggunakan file suara yang telah direkam. Pengujian pertama dilakukan dengan orang yang suaranya digunakan pada database sedangkan pengujian kedua dan ketiga dilakukan dengan orang yang berbeda.

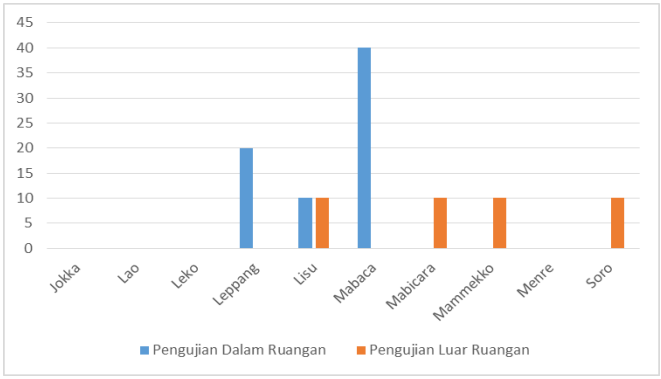


Gambar 8. Hasil Pengujian Database 10 Kata

Berdasarkan grafik diatas, pengujian kata Lao memiliki presentasi kebenaran yang paling tinggi diantara semua kata dengan nilai mencapai 30% pada salah satu pengujian. Sedangkan kata Mammekko dan Soro memiliki hasil pengujian dengan rata-rata persentasi kebenaran yang lebih rendah dibandingkan dengan yang lainnya.

D. *Pengujian Sistem*

Pengujian ini dilakukan dengan menggunakan database 37 orang yang berasal dari Pinrang dan Parepare, setiap orang mengucapkan 10 buah kata yang diulang sebanyak 2 kali. Jumlah data suara pada database ini adalah 740. Pengujian dilakukan sebanyak 10 kali, setiap penguji melakukan pengujian setiap kata sebanyak 2 kali..



Gambar 9. Hasil Pengujian Sistem

Berdasarkan grafik diatas, pengujian dari semua kata menghasilkan presentase kebenaran yang sangat rendah. Kata Mabbaca memiliki presentasi kebenaran yang lebih bagus dibandingkan dengan yang lainnya dengan nilai 40%. Sedangkan kata Jokka, Lao, Leko, Mabbicara, Mammekko, Menre dan Soro tidak memiliki kata yang benar dalam pengujian. Untuk pengujian di luar ruangan, hasil pengujian sangat rendah dengan persentase di bawah 10%

E. *Analisis Hasil Pengujian*

Analisis tingkat akurasi sistem dilakukan berdasarkan data pengujian yang diperoleh yang ditampilkan dalam bentuk grafik. Tingkat akurasi untuk setiap pengujian didapatkan dengan membandingkan jumlah pengucapan kata yang benar dengan jumlah pengujian yang dilakukan.

Pengujian dengan menggunakan data tunggal dengan variasi 1 kata dalam database memiliki presentase kebenaran yang lebih baik dibandingkan dengan yang lainnya. Pada pengujian data tunggal, terdapat beberapa pengujian dengan persentase kebenaran 50% ke atas. Sedangkan pengujian dengan menggunakan database dari beragam penutur menghasilkan persentase kebenaran pada pengujian yang lebih rendah.

Kompleksitas dalam proses kalkulasi vektor ciri dan output menjadi salah satu faktor penting yang menyebabkan tingkat akurasi yang rendah pada database dengan jumlah penutur yang lebih banyak. Keterbatasan kemampuan perangkat komputer yang digunakan dalam melakukan iterasi dalam proses pelatihan juga menjadi faktor yang menyebabkan tingkat akurasi rendah. Sehingga nilai error yang dihasilkan dalam proses pelatihan masih berada pada rentang 0,5 – 1,5.

Proses perekaman suara pada beberapa penutur yang memiliki banyak gangguan / noise juga menjadi salah satu faktor yang menyebabkan proses pelatihan dan pengujian suara belum maksimal. Perbedaan pengujian di dalam ruangan dan di luar ruangan tidak terlalu berbeda secara signifikan.

IV. KESIMPULAN

Perancangan sistem pengenalan ucapan bahasa daerah dengan MFCC sebagai algoritma untuk ekstraksi ciri dan pengelompokan dengan menggunakan ANFIS. Kata yang digunakan dalam penelitian ini adalah kata berbahasa bugis yaitu Jokka, Lao, Leko, Leppang, Lisu, Mabbaca, Mabbicara, Mammekko, Soro dan Menre. Hasil pengujian menunjukkan bahwa tingkat akurasi dengan menggunakan database data tunggal lebih baik daripada dengan menggunakan berbagai penutur pada database. Terdapat beberapa pengujian menghasilkan persentase kebenaran mencapai atau lebih dari 50%. Tingkat akurasi yang lebih rendah didapatkan pada pengujian dengan database dengan 37 penutur disebabkan oleh keterbatasan perangkat komputer dalam melakukan proses iterasi pada pelatihan suara sehingga nilai error yang dihasilkan masih berada pada rentang 0,5 – 1,5.

REFERENSI

[1] W.S.M. Sanjaya dan Z. Salleh. 2014. "Implementasi Pengenalan Pola Suara Menggunakan Mel-Frequency Cepstrum Coefficients (MFCC) Dan Adaptive Neuro-Fuzzy Inferense System (ANFIS) Sebagai Kontrol Lampu Otomatis". *Al-HAZEN Jurnal of Physics,* vol 1, no.1, hlm. 43-54, Desember 2014

[2] T. Nasution. "Metoda Mel Frequency Cepstrum Coefficients (MFCC) untuk Mengenali Ucapan pada Bahasa Indonesia". *Jurnal Sains dan Teknologi Informasi,* vol. 1, no. 1, hlm. 22-31, Juni 2012

[3] N.R. Pradipta, Tasripan, H. dan Kusuma. "Perancangan Perangkat Antarmuka Berbasis Pengenalan Suara pada Purwarupa Mesin Cetak Huruf Braille ITS untuk Operator Tunanetra". *Jurnal Teknik ITS,* vol. 7, no. 2, hlm. A401-A407, 2019.

[4] M.B. Gunawan. "Konversi Suara ke Teks Menggunakan Metode Hidden Markov Model". Skripsi, Universitas Sanata Dharma, Yogyakarta, Agustus 2010.

[5] David "Penerapan Algoritma Levenberg-Marquadt dan Backpropagation Neural Network Untuk Klasifikasi Suara Manusia". *Jurnal Buana Informatika,* vol. 4, no. 1, Januari 2013

[6] M.N. Rabbani, A. Rizal, dan F.Y. Suratman. "Implementasi Kunci Berbasis Suara Menggunakan Metode Mel Frequency Cepstral Coefficient (MFCC)". *e-Proceeding of Engineering,* vol. 3, no. 3, hlm. 3998-4007, Desember 2016.

[7] R. Umar, I. Riadi dan A. Hanif. "Analisis Bentuk Pola Suara Menggunakan Ekstraksi Ciri Mel-Frequency Cepstral Coefficients (MFCC)". *Cogito Smart Journal,* vol. 4, no.2, hlm. 294-304, 2019.

[8] U. Sutisna, R. Hidayat, dan Litasari. "Pengenalan Tutur Terisolasi Dalambahasa Indonesia Menggunakan MFCC, FCM, Dan ANFIS". *Seminar Nasional Teknologi Informasi dan Multimedia 2013,* hlm. 22-7 – 22-12.

.