

Deteksi Klasifikasi Ruangan Berdasarkan Reverberation Time Dengan Metode *Mel-Frequency Cepstral Coefficients* (Mfcc) Dan *Learning Vector Quantization* (Lvq)

1st Ariza Rizky Pratama
Fakultas Teknik Elektro
Universitas Telkom
Bandung, Indonesia

arizarizky@student.telkomuniversity.ac.id

2nd Jangkung Raharjo
Fakultas Teknik Elektro
Universitas Telkom
Bandung, Indonesia

jangkungraharjo@telkomuniversity.ac.id

3rd Bambang Hidayat
Fakultas Teknik Elektro
Universitas Telkom
Bandung, Indonesia

bhidayat@telkomuniversity.ac.id

Abstrak—Rekaman suara berupa file percakapan menjadi salah satu alat bantu dalam berbagai keperluan salah satunya bagi pihak kepolisian. Mengidentifikasi suatu kelas ruangan menjadi salah satu petunjuk dalam mengolah tempat kejadian perkara. Membuat sistem deteksi kelas ruangan adalah salah satu cara pihak kepolisian untuk mengidentifikasi kelas ruangan. Untuk menentukan suatu kelas ruangan dapat diukur melalui reverberation time menggunakan algoritma MFCC untuk mengekstraksi ciri data latih rekaman suara. Setelah didapat ekstraksi cirinya, ciri data akan di masukan sebagai dataset yang selanjutnya akan dilakukan proses pengujian. Kemudian data uji yang belum diketahui kelas ruangnya akan dilakukan proses pengujian. Data uji yang di proses akan di klasifikasikan menggunakan LVQ berdasarkan ciri data yang sudah ada di dataset. Hasil akhir pada sistem ini yaitu untuk memunculkan nilai akurasi dan waktu komputasi. Penelitian ini akan memproses data rekaman suara menggunakan aplikasi MATLAB, lalu diekstraksi ciri menggunakan MFCC dan mencari klasifikasi menggunakan LVQ. Menggunakan 48 data latih dan 18 data uji sebagai proses perhitungan dan simulasi. Hasil nilai akurasi yang didapat dari sistem deteksi kelas berdasarkan reverberation time dengan metode MFCC dan LVQ bernilai 94,44 % dan waktu komputasi selama 32,969474 detik dengan parameter Koef MFCC 40, Frame Size 0,05, Node Layer 5 dan Epoch 10.

Kata kunci— Reverberation Time, Mel Frequency Coefficients Cepstral, Learning Vector Quantization.

I. PENDAHULUAN

Perkembangan ekonomi Indonesia yang terjadi pada masa sekarang masih belum merata. Upaya pemerintah dalam mengembangkan ekonomi di Indonesia masih terus dilakukan. Perkembangan terhambat disebabkan pemerataan distribusi pendapatan masyarakat yang belum merata [1]. Salah satu faktor yang menunjang pemerataan adalah kurangnya lapangan kerja bagi masyarakat. Akibatnya, beberapa oknum masyarakat menghalalkan segala cara untuk mendapat keuntungan bagi dirinya. Salah satu target oknum tersebut adalah garasi pada suatu rumah yang berisikan kendaraan pribadi untuk dicuri. Masyarakat pun resah akan tindak yang dilakukan oknum tersebut, dan juga menjadi kerugian yang dialami bagi masyarakat yang dirugikan. Pihak aparat pun harus melakukan inovasi dalam menindak

kejahatan tersebut agar meminimalisir tindakan kejahatan pada masa saat ini. Namun dengan adanya *Reverberation Time* sebagai salah satu cara untuk mengolah data rekaman suara agar pihak aparat memudahkan dalam mengidentifikasi bukti khususnya data dalam bentuk rekaman suara.

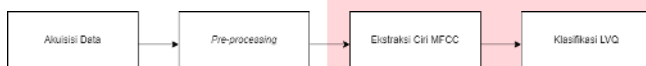
Reverberation adalah sinyal suara akustik yang dihasilkan dari pantulan berulang dari suatu ruang yang tertutup akibat dari superposisi dan juga gelombang yang dihasilkan dari dinding ruangan tersebut dan juga benda yang ada di dalamnya [2]. Hasil sinyal suara yang telah tereverberasi harus dilakukan proses pengolahan data suara dengan beberapa metode agar bisa mendapat hasil akurasi yang sesuai. Dengan konsep tersebut, penelitian ini bertujuan agar memudahkan pihak aparat untuk mengolah bukti data berupa rekaman suara dalam mengidentifikasi lokasi atau suatu kelas ruangan tempat dimana pelaku kejahatan melakukan aksinya.

Pada Tugas Akhir ini akan dilakukan dua kali pengujian dalam proses pengolahan data suara berdasarkan *Reverberation Time* menggunakan metode MFCC sebagai ekstraksi ciri dan mengklasifikasikan data suara tersebut menggunakan metode LVQ dengan ukuran ruangan yang berbeda-beda. Pengujian pertama ini menggunakan data latih yang di-*input*-kan selanjutnya akan di ekstraksi pada MFCC. Ciri data suara akan disimpan di file database dan akan digunakan di proses selanjutnya. Pengujian kedua ini menggunakan data uji, setelah selesai di-*input*-kan pada proses ekstraksi ciri selanjutnya data suara akan diklasifikasikan LVQ menggunakan ciri data yang sudah disimpan. Penulisan Tugas Akhir ini diharap dapat membantu pihak aparat dalam mengidentifikasi kelas ruangan.

II. KAJIAN TEORI

A. Sinyal Audio

Audio dapat diartikan sebagai bunyi atau suara. Bunyi adalah suatu gelombang yang merambat melalui suatu medium lalu dirasakan oleh panca indera manusia [8]. Gelombang suara adalah gelombang yang dihasilkan karena adanya getaran pada suatu benda. Sinyal audio bergetar melalui perambatan pada udara, air, dan juga material lainnya. Satu-satu penyebab sinyal audio tidak bisa merambat yaitu pada ruangan hampa udara. Gelombang suara yang dihasilkan memiliki satu lembah dan satu bukit, atau bisa disebut dengan satu siklus atau periode. Siklus ini pun berlangsung bergantian dan membentuk pada konsep frekuensi. Rentang frekuensi pendengaran manusia dengan fungsi pendengaran yang normal yaitu berkisar antara 20 – 20.000 Hz. Berikut adalah gambar proses sinyal audio.

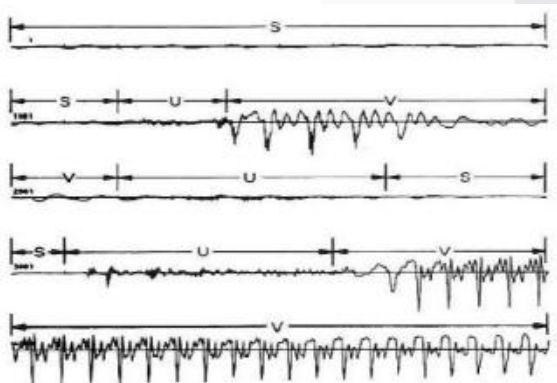


GAMBAR 2.1
Proses Sinyal Audio

Pada Gambar 2.1 data suara yang diambil dari sumber suara akan diekstraksi cirinya melalui software matlab menggunakan metode MFCC. Setelah mendapatkan ekstraksi cirinya, proses selanjutnya yaitu pemrosesan klasifikasi dengan metode LVQ.

1. Sinyal Suara Manusia

Proses suara manusia, suara yang di hasilkan dari dorongan udara di paru-paru melalui suara vokal, dan keluar melalui mulut untuk menjadi wicara. Pada proses tersebut dijelaskan bahwa paru-paru adalah sumber dari suara yang akan dihasilkan, dan saluran vokal adalah filter untuk menghasilkan berbagai suara yang akan keluar dari mulut seseorang ketika berbicara. Fonem adalah perangkat terbatas suara individu yang dihasilkan. Ada dua kategori pada fonem yaitu, *voiced sounds* dan *unvoiced sounds*. Kategori tersebut sangat penting dalam proses produksi suara [9]. Ada beberapa klasifikasi dari sinyal suara manusia yaitu *silence*, *unvoiced*, dan *voiced*.



GAMBAR 2.2
Representasi Sinyal Suara Manusia

Pada Gambar 2.2 merupakan representasi sinyal suara yang dihasilkan selama 100 ms. Bagian S merupakan daerah *silence*, U merupakan daerah *unvoiced*, dan v merupakan daerah *voiced*. Bisa dilihat pada sinyal S getarannya hanya

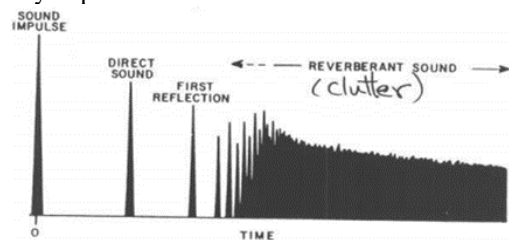
datar menandakan tidak ada suara yang dihasilkan. Pada daerah U getaran sinyal hanya sedikit yang menandakan suara yang dihasilkan sedikit. Pada daerah V getaran sinyal bergerak sangat besar yang menandakan keluaran suara yang besar.

B. Reverberation Time

Reverberation time (RT) atau waktu dengung dihasilkan ketika pantulan bunyi atau suara mengenai suatu permukaan benda. Ketika di suatu ruangan suara akan terdengar lebih lama setelah sumber suara berhenti pada rentang waktu tertentu. Setelah sumber suara dihentikan, RT akan mengukur suara yang sudah dihasilkan berdasarkan gema pada suatu ruangan tersebut [4]. Pada saat menghitung RT ada beberapa faktor yang mempengaruhi RT yaitu, ukuran volume di suatu ruangan tersebut, bentuk ruangan yang akan di ukur, dan bahan yang digunakan pada ruangan tersebut seperti dinding, lantai, dan plafon [5]. Untuk menghitung nilai dari RT pada suatu ruangan, digunakan rumus seperti berikut.

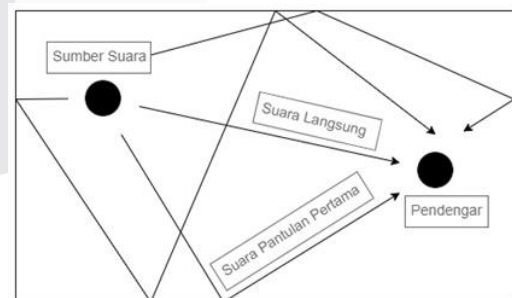
$$T = \frac{0,161V}{A} \tag{2.1}$$

Keterangan:
T = Reverberation Time.
V = Volume Ruangan.
A = Penyerapan Total.



GAMBAR 2.3
Grafik Energi Reverberation Terhadap Waktu

Pada Gambar 2.3 dijelaskan bahwa suara yang sudah tereverberasi. Impuls pertama menunjukkan dari sumber suara yang di hasilkan, lalu ada impuls kedua menunjukkan suara langsung yang terdengar oleh pendengar, terakhir ada implus yang berdempetan dihasilkan dari suara pantulan dari ruangan tersebut atau disebut *reverberation time*.

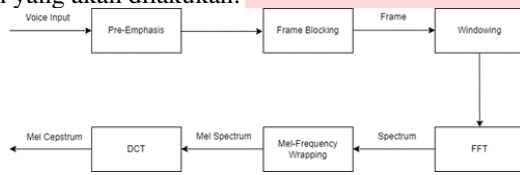


GAMBAR 2.4
Pantulan Sinyal Suara Pada Ruangan

Ada beberapa parameter yang mempengaruhi proses *reverberation* yaitu, waktu tunda (*decay*) dan intensitas refleksi awal, waktu tunda dan intensitas refleksi akhir, dan waktu peluruhan. Setiap parameter tersebut saling mempengaruhi dan berhubungan dalam sebuah efek gema. [10]. Ada beberapa parameter yaitu, *decay*, *early reflections*, *early intensity*, *early delay*, *late intensity*, dan *late delay*.

C. Mel Frequency Coefficients Cepstral (MFCC)

Mel Frequency Coefficients Cepstral (MFCC) merupakan salah satu ekstraksi ciri yang digunakan setelah pengambilan data suara untuk mencari ciri akustik pada data suara. Metode ekstraksi ciri MFCC digunakan untuk mendapatkan *cepstral coefficient* dan *frame*, sehingga dalam memproses pengenalan suara memiliki ketepatan yang akurat [6]. Pada umumnya metode ini didasarkan pada bandwidth dan frekuensi yang dapat diterima manusia [7]. Ada beberapa langkah untuk memproses sinyal suara yang sudah di inputkan, yaitu *Pre-emphasis*, *frame blocking*, *windowing*, *fast fourier transform*, *mel frequency wrapping* dan *discrete cosine transform* (DCT). MFCC akan mengekstrak ciri akustik dari sinyal yang di-input-kan untuk memperoleh nilai koefisien, lalu menyimpannya di dataset. Dalam proses pengkodean sinyal analog suara manusia, berikut ilustrasi langkah yang akan dilakukan.



GAMBAR 2.5 Diagram Blok MFCC Dalam Proses Audio

1. *Pre-Emphasis* mengacu pada memfokuskan proses penyaringan pada frekuensi tinggi dan bertujuan untuk menyetarakan seluruh frekuensi suara. *Pre-emphasis* dapat dihitung dengan persamaan:

$$y(n) = S(n) - \alpha * S(n - 1) \tag{2.2}$$

2. *Frame Blocking* pada proses ini dilakukan pemotongan pada sinyal menjadi blok. *Frame* diproses agar memperoleh frekuensi yang stabil.

$$f(n) = y(Ml + n) \tag{2.3}$$

3. *Windowing* dilakukan pada setiap *frame* untuk meminimalisir adanya kontinuitas sinyal pada awal dan akhir dari setiap *frame* yang terjadi karena dilakukannya *frame blocking*.

$$w(n) = 0.54 - 0.64\cos\left(\frac{2\pi.n}{M-1}\right) \tag{2.4}$$

Berikut adalah persamaan digunakan pada proses *hamming windowing*.

$$x(n) = f(n)*w(n); 0 \leq n \leq M - 1. \tag{2.5}$$

4. *Fast Fourier Transform* (FFT) adalah mengkonversi tiap pada sampel N dari bentuk domain waktu ke dalam bentuk domain frekuensi.

$$T[k] = \sum_{n=0}^{n-1} x(n)\cos\left(\frac{2\pi kn}{N}\right) - \sum_{n=0}^{n-1} x(n)\sin\left(\frac{2\pi kn}{N}\right) \tag{2.6}$$

5. *Mel-Frequency Wrapping* pada proses ini untuk pitch subjektif diukur dengan satuan ‘mel’. Skala frekuensi ‘mel’ memiliki sifat linier untuk frekuensi dibawah 1 kHz dan bersifat logaritmik untuk frekuensi diatas 1kHz.

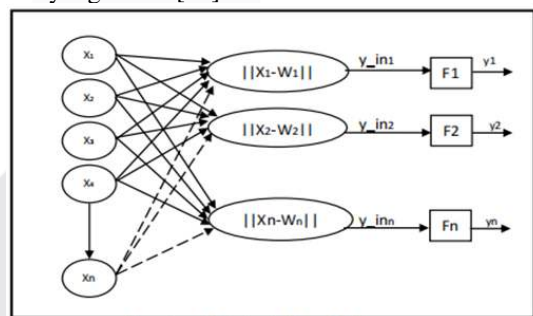
$$Y[k] = \sum_{j=1}^G T[k]H[j] \tag{2.7}$$

6. *Discrete Cosine Transform* (DCT) pada proses ini spectrum yang dihasilkan frekuensi mel dikonversikan kembali ke dalam bentuk domain waktu yang hasilnya disebut MFCC.

$$C_m = \sum_{k=1}^K (\log_{10} Y[k])\cos\left[m\left(k - \frac{1}{2}\right)\frac{\pi}{K}\right]; m = 1,2, \dots, K \tag{2.8}$$

D. Learning Vector Quantization (LVQ)

Learning Vector Quantization (LVQ) merupakan metode jaringan saraf tiruan dengan tipe arsitektur jaringan lapis tunggal. LVQ bertujuan mendapat pembelajaran pada lapisan kompetitif yang terawasi berdasarkan jarak antara vektor masukan dan bobot menggunakan persamaan *euclidian distance* [17]. LVQ salah satu metode klasifikasi yang sangat populer dengan algoritma pembelajaran berbasis prototipe intuitif dengan aplikasi yang sukses di bidang telekomunikasi. [11]. Klasifikasi ini dapat menentukan sebuah kelompok dari output yang dihasilkan. LVQ mempelajari terhadap lapisan kompetitif yang nantinya akan menargetkan kepada klasifikasi yang didefinisikan oleh pengguna. Lapisan kompetitif akan mempelajari vektor-vektor yang di-input-kan secara otomatis. Lapisan kompetitif akan menghasilkan kelas-kelas sesuai jarak antara vektor-vektor yang di input. Jika vektor yang di input-kan mendekati sama, maka kedua vektor tersebut akan dimasukkan kedalam klasifikasi yang sama. [12]



GAMBAR 2.6 Arsitektur Klasifikasi LVQ [12]

Keterangan:

Keterangan pada arsitektur:

- a. X1 sampai Xn : Nilai input
- b. $\|X1-W1\|$ sampai $\|Xn-Wn\|$: Jarak bobot
- X1-W1 : Jarak bobot pertama
- Xn-Wn : Jarak bobot akhir
- c. F1-Fn : Lapisan output
- d. y1-yn : Nilai output
- e. n : Jumlah kelas

Algoritma dalam melakukan LVQ adalah sebagai berikut [18]:

1. Tentukan maksimum *epoch*, untuk memaksimalkan banyaknya pembelajaran berulang dan menentukan nilai *alpha* (α).

2. Hasil dari ekstraksi fitur merupakan pola yang akan digunakan sebagai data awal. Data awal ini akan diisi sebagai nilai bobot awal (w).
3. Nilai epoch akan dimulai dari nilai iterasi 0.
4. Selama $epoch < max\ epoch$, maka lanjut ke Langkah selanjutnya
5. $Epoch = Epoch + 1$.
6. Untuk setiap ekstraksi fitur, lakukan hal berikut:
 - a. Set x = hasil ekstraksi fitur pada pola.
 - b. Set T = nomor urut dari setiap kelas.
 - c. Hitung jarak dengan rumus *euclidean distance*.

$$J = \sqrt{(x_{11} - w_{11})^2 + \dots + (x_{1m} - w_{1m})^2} \quad (2.9)$$

- d. Jika kelas bobot memiliki jarak terkecil, maka.

$$W_j(\text{baru}) = W_j(\text{lama}) + \alpha(x - W_j(\text{lama})) \quad (2.10)$$

- e. Bila tidak, maka.

$$W_j(\text{baru}) = W_j(\text{lama}) - \alpha(x - W_j(\text{lama})) \quad (2.11)$$

7. Kurangi nilai α .

$$\alpha(\text{baru}) = 0.1 \times \alpha(\text{lama}) \quad (2.12)$$

III. METODE

Rancangan sistem ini akan dijelaskan seperti Gambar 3.1.



GAMBAR 3.1
Desain Sistem

A. Akuisisi Data

Akuisisi data dilakukan pada tahap awal. Data suara yang direkam menggunakan suara 2 orang laki-laki dan 1 orang perempuan. Suara manusia yang direkam berupa dialog dengan menggunakan beberapa kata, mulai dari 1 kata, 3 kata, dan 5 kata. Beberapa kata pada dialog yang digunakan bertujuan agar data suara didapatkan bervariasi dan juga durasi yang dihasilkan berbeda-beda. Ruang pada perekaman suara dibedakan menjadi 3 kelas yaitu, ruangan besar (RB), ruangan sedang (RS), dan ruangan kecil (RK), dengan menghasilkan 66 data suara. Data suara yang dihasilkan akan dibagi menjadi 2, yaitu data training dan data testing. Data training menggunakan 48 data, dengan data ruangan besar berjumlah 15 data, ruangan sedang berjumlah 14 data, dan ruangan kecil berjumlah 19 data. Sedangkan pada data testing menggunakan 18 data, dengan data ruangan besar berjumlah 8 data, ruangan sedang berjumlah 8 data, dan ruangan kecil berjumlah 8 data. Data diambil menggunakan spesifikasi microphone mono condenser dengan polarisasi unidirectional. Software yang digunakan untuk proses pengambilan data yaitu Fruity Loops Studio 12 dengan format data “,wav” dengan frekuensi sampling 44.100 Hz.

B. Pre-processing

Pre-processing, merupakan langkah setelah proses sebelumnya yang bertujuan untuk memperbaiki data audio yang sudah diambil sebelum mencari ekstraksi cirinya. Ada 3 operasi pada tahap pre-processing ini [19]:

1. Signal pre-emphasis

Proses penyangiran data audio sehingga frekuensi tinggi akan disetarakan.

2. Noise Elimination

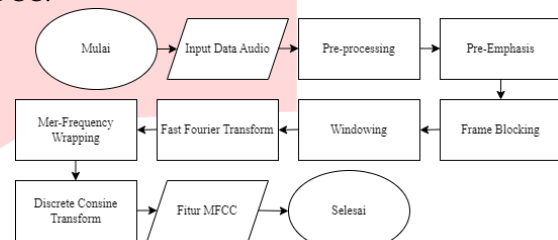
Proses ini *noise* atau sinyal yang ada pada data audio akan dihilangkan agar suara lebih jernih.

3. Signal Enhancement

Tahap terakhir proses ini bertujuan untuk membuat peak pada format sinyal audio menjadi lebih memuncak.

C. Ekstraksi Ciri MFCC

Pada tahap ekstraksi ciri menggunakan algoritma MFCC untuk mencari ciri data pada data yang sudah di-input-kan. MFCC akan mengekstraksi data untuk mendapatkan *cepstral coefficient* dan *frame*, agar sehingga pada sehingga dalam memproses pengenalan suara memiliki ketepatan yang akurat. Gambar 3.2 merupakan diagram blok pada proses MFCC.

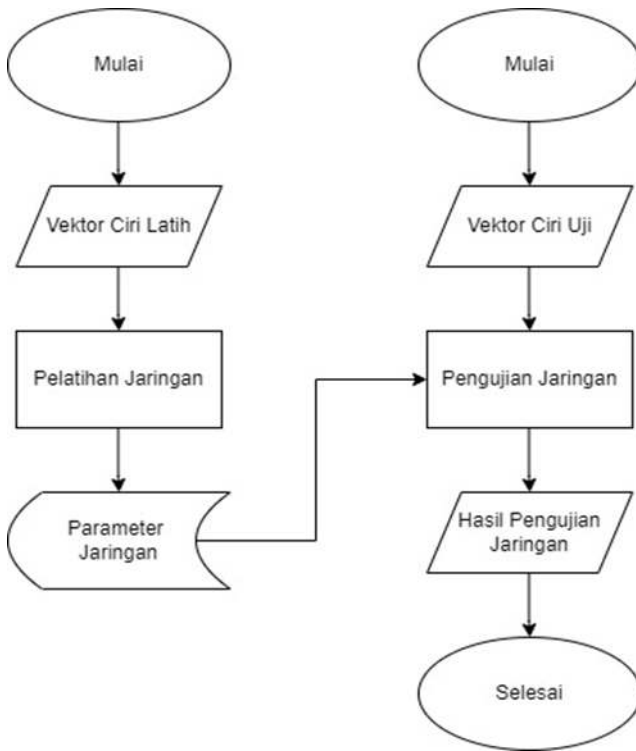


GAMBAR 3.2
Diagram Blok MFCC

Tahap pertama pada proses ekstraksi ciri MFCC yaitu sinyal audio akan dibaca pada sistem setelah di run. Pada tahap *pre-processing* berfungsi untuk menghilangkan *noise* pada data yang di-input-kan agar kualitas suara menjadi lebih baik. Selanjutnya, tahap *pre-emphasis* yaitu penyangiran frekuensi tinggi untuk menyetarakan seluruh frekuensi pada data suara. *Frame blocking* sinyal akan di potong per-*frame* agar frekuensi dihasilkan lebih stabil. Kemudian tahap *windowing* yaitu mengurangi kontinuitas pada sinyal awal dan akhir pada setiap *frame*. *Fast fourier transform* berfungsi untuk mengkonversi *frame* yang sudah di filter dari bentuk domain waktu ke dalam bentuk domain frekuensi. Selanjutnya tahap *mel frequency wrapping* berfungsi untuk mendapatkan skala nilai mel pada sinyal suara. Tahap terkahir yaitu *discrete cosine transform*, proses ini *spectrum* yang dihasilkan frekuensi mel dikonversikan kembali ke dalam bentuk domain waktu. Fitur yang dihasilkan pada ekstraksi ciri ini yaitu berupa *output* nilai matriks koefisien, delta, dan delta-delta.

D. Klasifikasi LVQ

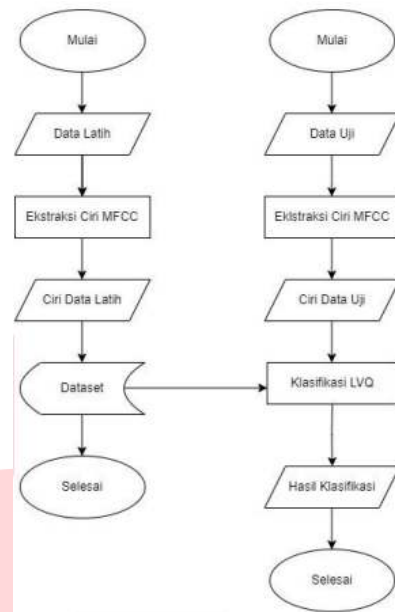
Klasifikasi pada sistem ini berfungsi untuk mengklasifikasikan data menurut kelasnya. Klasifikasi pada sistem ini menggunakan LVQ. Tiga kelas dibuat pada sistem ini yaitu, ruang kecil, ruang sedang, dan ruang besar. Dengan memasukan data latih dan data uji yang sudah di ekstraksi menggunakan MFCC, kemudian akan digunakan *epoch* dan *node layer* yang sudah ditentukan. *Epoch* dan *node layer* adalah parameter pada klasifikasi LVQ. Pelatihan data akan di ambil dari folder yang sudah didefinisikan sebagai data latih yang akan di proses. Proses terakhir pada klasifikasi ini melakukan pengujian terhadap data uji untuk membandingkan antara data uji dan data latih. Berikut adalah ilustrasi pada klasifikasi LVQ pada Gambar 3.3.



GAMBAR 3.3
Diagram Blok LVQ

E. Desain Perancangan Sistem

Pada rancangan sistem di penelitian ini, menggunakan perangkat lunak Matlab pada saat memproses data yang telah direkam. Sistem ini terbagi mejadi 2 langkah, yaitu tahap pelatihan dan tahap pengujian. Tahap pelatihan akan di-input-kan data latih sebanyak 48 sampel yang akan digunakan sebagai pelatihan audio. Proses selanjutnya data akan di proses untuk mendapatkan ekstraksi cirinya dengan menggunakan *Mel Frequency Coefficients Cepstral* (MFCC). Ekstraksi ciri ini juga akan memproses data latih untuk mencari ciri akustik pada data yang di-input-kan. Hasil dari proses dari ekstraksi ciri akan disimpan di dataset dengan berbagai kelas ukuran ruangan dalam bentuk file excel. Pada proses selanjutnya data yang di-input-kan tahap pengujian sebanyak 12 sampel. Tahap pengujian akan melalui proses ekstraksi ciri. Selanjutnya sistem ini akan mendeteksi ruangan berdasarkan pada dataset yang telah dibuat sebelumnya dengan klasifikasi *Learning Vector Quantization* (LVQ). Setelah klasifikasi ini menentukan kelas data audio berdasarkan dataset maka akan keluar output berupa nilai akurasi dan waktu komputasi yang sesuai dengan data yang diujikan dan proses dari nilai akurasi. Diagram blok sistem tugas akhir dapat dilihat pada Gambar 3.4.



GAMBAR 3.4
Diagram Blok Rancangan Sistem Data Latih dan Data Uji

IV. HASIL DAN PEMBAHASAN

Pada bab ini membahas tentang pengujian sistem yang telah dibuat menggunakan data set yang telah diambil dan diproses. Adapun tujuan dari pengujian sistem ini yaitu untuk menganalisis pengaruh nilai parameter yang digunakan. Beberapa parameter dirancang pada simulasi sistem ini yaitu, *Koef MFCC*, *Frame size*, *Node Layer*, dan *Epoch*. Sehingga dari hasil analisis pengujian sistem tersebut akan didapatkan nilai akurasi dan waktu komputasi yang terbaik dari beberapa skema parameter yang telah dibuat.

Skenario pada pengujian sistem dibagi menjadi 2 yaitu, pertama dengan menggunakan parameter *Koef MFCC*, *Frame Size 0,025 dan 0,05*, *Node Layer (3, 5, 7, dan 10)*, dan *Epoch (5, 10, 15, dan 20)*. Kedua dengan menggunakan parameter *Koef MFCC 40*, *Frame Size 0,025 dan 0,05*, *Node Layer (3, 5, 7, dan 10)*, dan *Epoch (5, 10, 15, dan 20)*. Tetapi yang ditampilkan hanya skenario dengan akurasi tertinggi yaitu, *Koef MFCC 40*, *Frame Size 0,05*, *Node Layer 5* dan *Epoch 10*.

TABEL 4.1
Nilai Akurasi Data Latih

Nilai Akurasi Data Latih				
Frame Size 0,05 dan Koef MFCC 40				
	Epoch			
	5	10	15	20
Node Layer = 3	68,75%	68,75%	68,75%	68,75%
Node Layer = 5	100%	100%	100%	100%
Node Layer = 7	97,91%	100%	100%	100%
Node Layer = 10	100%	97,91%	97,91%	100%

Pada Tabel 4.1 menjelaskan hasil nilai akurasi data latih. Hasil nilai akurasi tertinggi pada pengujian sistem didapatkan nilai sebesar 100%.

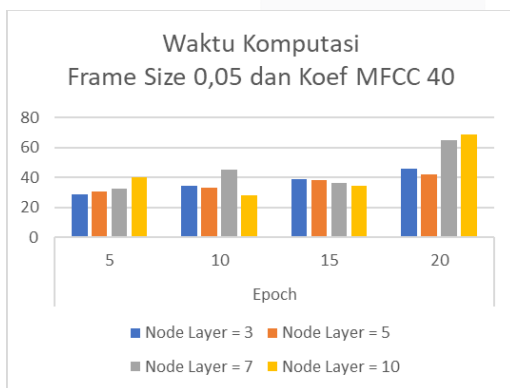
Tabel 4. 2
Nilai Akurasi Data Uji

Nilai Akurasi Data Uji				
Frame Size 0,05 dan Koef MFCC 40				
	epoch			
	5	10	15	20
Node Layer = 3	61,11%	61,11%	61,11%	61,11%
Node Layer = 5	88,88%	94,44%	94,44%	88,88%
Node Layer = 7	94,44%	88,88%	88,88%	94,44%
Node Layer = 10	94,44%	94,44%	94,44%	94,44%

Tabel 4. 3
Nilai Waktu Komputasi

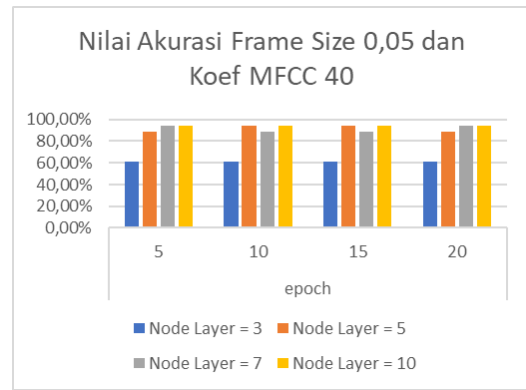
Nilai Waktu Komputasi				
Frame Size 0,05 dan Koef MFCC 40				
	Epoch			
	5	10	15	20
Node Layer = 3	28,894468	34,15934	39,050348	45,960829
Node Layer = 5	30,572193	32,969474	38,20962	42,118431
Node Layer = 7	32,570986	44,935421	36,442943	65,082167
Node Layer = 10	40,169598	27,821387	34,222178	68,769598

Untuk memudahkan visualisasi dalam menganalisis hasil nilai akurasi pada Tabel 4.2, dilakukan pemindahan data dari tabel kedalam grafik seperti pada Gambar 4.1 sebagai berikut.



GAMBAR 4.1
Grafik Akurasi

Untuk memudahkan visualisasi dalam menganalisis hasil nilai waktu komputasi pada Tabel 4.3, dilakukan pemindahan data dari tabel kedalam grafik seperti pada Gambar 4.2 berikut.



GAMBAR 4.2
Waktu Komputasi

Pada Tabel 4.2 dan Gambar 4.1 grafik menampilkan hasil skenario pengujian dengan menggunakan parameter *Frame Size* bernilai 0,05. Hasil pengujian ini menunjukkan nilai akurasi tertinggi pada nilai parameter *Node Layer* 5 dengan *Epoch* yang bernilai 10. Nilai akurasi yang di dapatkan pada pengujian ini bernilai 94,44 % dengan hasil waktu komputasi yang bernilai 32,969474 detik. Adapun nilai akurasi terendah pada pengujian ini, pada nilai parameter *Node Layer* 3 dan nilai *Epoch* 20, dengan hasil nilai akurasi 61,11% dan nilai waktu komputasi 45,960829 detik. Namun, pada pengujian ini ada beberapa nilai akurasi yang sama dengan 94,44 %. Pada nilai parameter *Node Layer* 7 dengan nilai *Epoch* 5 dan *Node Layer* 10 dengan nilai *Epoch* 10 dan 15. Tetapi nilai tersebut tidak termasuk nilai akurasi tertinggi pada pengujian ini, karena nilai akurasi data latih yang tidak mencapai 100 %. Menjadikan nilai akurasi tersebut menjadi tidak tervalidasi. Trend grafik hasil nilai akurasi yang ditampilkan, menunjukkan perkembangan nilai parameter yang digunakan terhadap hasil nilai akurasi. Pada nilai *Node Layer* 3 dan 10 dengan nilai *Epoch* 5 hingga 20 yang ditampilkan pada grafik menunjukkan bahwa nilai akurasi yang konsisten dengan nilai 61,11 % dan 94,44 %. Pada *Node Layer* 5 menghasilkan nilai akurasi yang cenderung konsisten. Karena pada grafik yang ditampilkan mengalami penurunan dan kenaikan. Pada nilai *Epoch* 5 hingga 10 grafik mengalami kenaikan hingga nilai akurasi 94,44 %, lalu mengalami penurunan pada nilai *Epoch* 15 hingga 20 hingga nilai akurasi menjadi 88,88 %. Pada *Node Layer* 7 juga mengalami nilai akurasi yang cenderung konsisten. Pada nilai *Epoch* 5 hingga 10 grafik menunjukkan penurunan dan pada nilai *Epoch* 15 hingga 20 grafik kembali mengalami kenaikan. Namun pada nilai *Node Layer* 5 dan 7 kenaikan dan penurunan hasil nilai akurasi tidak terlalu signifikan. Trend pada nilai waktu komputasi juga dapat dilihat pada grafik yang ditampilkan. Pada *Node Layer* 3 dan 5 hasil waktu komputasi menunjukkan grafik yang konsisten menaik, karena bertambahnya nilai *Epoch* yang digunakan menjadikan melakukan pembelajaran beberapa kali. Pada *Node Layer* 7 dan 10 mengalami kenaikan dan penurunan ada grafik. Hal ini disebabkan karena hasil dari waktu komputasi yang cenderung tidak konsisten.

V. KESIMPULAN

Berdasarkan hasil dari analisis kedua pengujian sistem deteksi klasifikasi ruangan menggunakan metode ekstraksi ciri MFCC dan klasifikasi LVQ, dapat disimpulkan bahwa, akurasi terbaik pada sistem yang digunakan dalam

mendeteksi ukuran ruangan mencapai nilai sebesar 94,44 %.. Hasil dari nilai waktu komputasi yang dibutuhkan dalam memproses pengujian untuk mendapatkan nilai akurasi tersebut memerlukan 32,969474 detik. Beberapa nilai dan jenis parameter yang digunakan mempengaruhi pengujian sistem anatra lain pada nilai *Koef MFCC* 40, *Frame Size* 0,05, *Node Layer* 5, dan *Epoch* 10. Proses *Preprocessing audio* sangat mempengaruhi kualitas data yang akan diujikan dalam mendeteksi ukuran ruangan dan akan mempengaruhi nilai akurasi.

REFERENSI

- [1] "Pertumbuhan Ekonomi yang Kurang Optimal Sebabkan Pendapatan Tidak Merata," Kontan Nasional, diakses pada tanggal 27 Januari 2023, <https://nasional.kontan.co.id/news/pertumbuhan-ekonomi-yang-kurang-optimal-sebabkan-pendapatan-tidak-merata>.
- [2] A. B. Cendana, "Teknik Peningkatan Kualitas Rekaman Sinyal Musik dengan Sistem Blind Dereverberation menggunakan Pemodelan Autoregresif untuk Mengurangi Dampak Reverberasi," *Jurnal Teknik Elektro*, vol. 7, no. 1, hlm. 1-8, 2016.
- [3] D. E. Fajar, "Identifikasi Emosi Melalui Suara Menggunakan Support Vector Machine dan Covolutional Neural Network," *Jurnal Sains dan Teknologi*, vol. 12, no. 2, hlm. 25-30, 2018.
- [4] K. W. Rintel, *Essentials of Audiology*, 2nd ed., New York: Thieme, 2015.
- [5] Yahya M.N., dkk, "Investigation the capability of neural network in predicting reverberation time on classroom," *Journal of Acoustical Engineering and Technology*, vol. 7, no. 2, hlm. 105-112, 2016
- [6] A. B. Cendana, "Ekstraksi Ciri Mel Frequency Cepstral Coefficient (MFCC) dan Rerata Coefficient untuk Pengecekan Bacaan Al-Qur'an," *Jurnal Ilmu Komputer dan Informasi*, vol. 6, no. 2, hlm. 45-50, 2017.
- [7] D. E. Fajar, "Klasifikasi Jenis Vokal Manusia menggunakan MFCC dan Convolutional Neural Network," *Jurnal Teknologi dan Sistem Komputer*, vol. 15, no. 3, hlm. 125-130, 2019.
- [8] K. W. Rintel, *Digital Signal Processing in Audio and Acoustical Engineering*, 2nd ed., New York: Springer, 2018.
- [9] A. B. Cendana, "Klasifikasi Emosi Berdasarkan Sinyal Suara Manusia Menggunakan Metode K-Nearest Neighbor (K-NN)," *Jurnal Teknologi dan Sistem Komputer*, vol. 14, no. 4, hlm. 150-155, 2018.
- [10] D. E. Fajar, "Simulasi Efek Suara Gema Dimensi Ruang (Reverb) dengan Menggunakan Aplikasi Android," *Jurnal Teknologi dan Sistem Komputer*, vol. 16, no. 1, hlm. 40-45, 2020.
- [11] T. U. Vianto, *Self organizing maps 3rd ed*, Jakarta: Penerbit Universitas, 2019
- [12] A. B. Cendana, "Identifikasi Citra Pola Rugae Palatina pada populasi Mahasiswa S1 Teknik Telekomunikasi 2015 Universitas Telkom menggunakan Metode Image Registration dan Binary Large Object (BLOB) dengan Teknik Klasifikasi Learning Vector Quantization (LVQ)," *Jurnal Ilmiah Teknik Telekomunikasi*, vol. 10, no. 2, pp. 1-10, 2015.
- [13] A. Author, "Klasifikasi Emosi Berdasarkan Sinyal Suara Manusia Menggunakan Metode K-Nearest Neighbor (K-NN)," *Journal of Speech and Audio Processing*, vol. 12, no. 3, pp. 123-145, 2020.
- [14] D. Smith, "Microphone Pickup or Polar Patterns," *Disc Makers Blog*, July 2012, [Online]. Available: <https://blog.discmakers.com/2012/07/microphone-pickup-or-polar-patterns/>. [Accessed: 23-Feb-2021].
- [15] S. Errede, "Department of Physics," University of Illinois, [Online]. Available: <https://www.physics.illinois.edu/people/faculty/prof-steven-errede>. [Accessed: 23-Feb-2021].
- [16] R. Hasan and A. N. Putra, "Verifikasi Suara Mahasiswa Sebagai Alternatif Presensi Kehadiran Menggunakan Ekstraksi Fitur MFCC dan Klasifikasi LVQ," *Journal of Speech and Audio Processing*, vol. 14, no. 1, pp. 56-78, 2020.
- [17] K. Williams and J. Smith, "Implementation of MFCC and LVQ Methods for Learning English Pronunciation," *Journal of Language Learning*, vol. 16, no. 2, pp. 89-104, 2020.
- [18] L. Fausett, *Fundamentals of Neural Networks: Architectures, Algorithms, and Applications*. Prentice-Hall, 1994.
- [19] B.-H. J. Lawrence Rabiner, *Fundamental of Speech Recognition*, Englewood Cliffs: PTR Prentice-Hall, Inc., 1993.