

Model Identifikasi Kata Ucapan Tuna Wicara

Nuruddin Wiranda*¹, Agfianto Eko Putra²

¹Program Studi Pendidikan Ilmu Komputer, FKIP, ULM, Banjarmasin, Indonesia

²Departemen Ilmu Komputer dan Elektronika, FMIPA, UGM, Yogyakarta, Indonesia

e-mail: *¹nuruddin.wd@ulm.ac.id, ²agfi@ugm.ac.id

Abstrak

Tuna wicara adalah ketidak-mampuan seseorang berbicara, padahal kemampuan berbicara merupakan hal penting agar dapat berkomunikasi dengan orang lain. Menyiasati hal tersebut seseorang yang mengalami tuna wicara memiliki cara tersendiri dalam berkomunikasi, yaitu dengan menggunakan bahasa isyarat, akan tetapi tidak semua orang memahami bahasa isyarat tersebut. Metode MFCC dan JST Backpropagation diimplementasikan pada Single Board Computer (SBC) dirancang untuk mengatasi permasalahan komunikasi tuna wicara. Metode MFCC digunakan untuk mengambil fitur dari suara tuna wicara dan JST Backpropagation digunakan untuk pengenalan pola suara.

Sistem dilatih menggunakan 750 sampel suara yang terdiri dari 5 penutur, masing-masing mengucapkan sebanyak 30 kali pengulangan pengucapan kata (makan, kamar, kerja, harga dan lapar), kemudian diuji menggunakan 125 sampel suara yang terdiri dari 5 penutur, masing-masing mengucapkan 5 kali pengulangan kata. Pelatihan dan pengujian JST Backpropagation menggunakan masukan koefisien yang dihasilkan dari MFCC. Hasil penelitian menunjukkan bahwa Metode MFCC dan JST Backpropagation mampu mengidentifikasi ucapan kata tuna wicara dengan akurasi 60%, presisi 40% dan sensitivitas 40%.

Kata kunci— MFCC, JST Backpropagation, Single board computer, Tuna Wicara

Abstract

Speech impaired is the inability of someone to speak, even though speaking ability is important in order to communicate with other people. Dealing with this as someone who has speech impairments has their own way of communicating, namely by using sign language, but not everyone understands the sign language. The MFCC and Backpropagation ANN methods are implemented on a Single Board Computer (SBC) designed to overcome speech impaired communication problems. The MFCC method is used to retrieve the features of speech impairment and the Backpropagation ANN is used for sound pattern recognition.

The system was trained using 750 sound samples consisting of 5 speakers, each uttering as many as 30 repetitions of the pronunciation of words (makan, kamar, kerja, harga and lapar), then tested using 125 sound samples consisting of 5 speakers, each saying 5 repetitions of words. Training and testing of Backpropagation ANN using input coefficients generated from MFCC. The results showed that the MFCC and Backpropagation ANN methods were able to identify speech words with 60% accuracy, 40% precision and 40% sensitivity.

Keywords— MFCC, JST Backpropagation, Single board computer, Speech impaired

1. PENDAHULUAN

Tunawicara merupakan ketidak-mampuan seseorang dalam berbicara. Hal ini disebabkan oleh kurang atau tidak berfungsinya organ-organ untuk berbicara, seperti rongga mulut, langit langit, lidah dan pita suara. Selain itu juga adanya kekurangan pada indra pendengaran, keterlambatan perkembangan bahasa, kerusakan pada sistem syaraf dan struktur otot. Badan Pusat Statistik (BPS) menyajikan data statistik disabilitas dalam SUSENAS 2009 untuk penyandang tuna wicara berjumlah 151.427 dan tuna rungu wicara 73.587 jiwa di Indonesia[1].

Penelitian ini merancang-bangun sebuah sistem untuk membantu orang awam agar bisa mengerti apa yang diucapkan oleh tuna wicara menggunakan metode *Mel Frequency Cepstral Coefficient* (MFCC) dan Jaringan Saraf Tiruan (JST). Sistem ini memanfaatkan suara yang diucapkan oleh tunawicara tersebut, kemudian dilakukan proses pengenalan suara untuk mengetahui arti dari suara tersebut.

Beberapa penelitian dalam [2], [3], [4] dan [5] menyatakan bahwa ekstraksi fitur menggunakan MFCC lebih unggul daripada menggunakan metode linear lain seperti LPCC dan LPC. Penelitian [6] melakukan penelitian menggunakan JST *Backpropagation* untuk melatih fitur yang di dapat dari MFCC dengan hasil akurasi mencapai 99.8%. Sedangkan [7] dalam penelitiannya menggunakan ekstraksi fitur MFCC dan klasifikasi JST untuk mengidentifikasi Qori Qur'an memperoleh akurasi sebesar 91.2%. Penelitian [8] menyimpulkan bahwa JST merupakan metode yang baik untuk pengenalan ucapan, sedangkan [9] dan [10] menggunakan JST dan MFCC untuk pengenalan suara dengan hasil akurasi sangat baik.

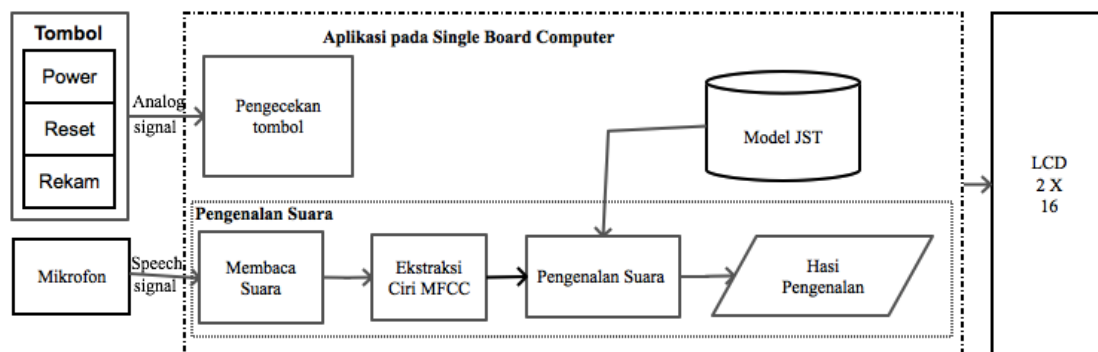
Tujuan dari penelitian yang dilakukan adalah merancang-bangun sistem yang mampu mengidentifikasi kata ucapan seorang tunawicara menggunakan metode MFCC dan JST *Backpropagation*. Hasil penelitian ini diharapkan dapat memberikan manfaat bagi orang awam agar bisa mengerti kata yang diucapkan oleh tunawicara.

2. METODE PENELITIAN

2.1 Gambaran Keseluruhan Sistem

Masukan sistem alat bantu tunawicara adalah suara yang ditangkap dari mikrofon dan 3 tombol yang berfungsi untuk mematikan, mereset sistem dan untuk merekam suara. Masukan dari mikrofon atau tombol diproses oleh aplikasi yang telah ditanamkan pada SBC, hasilnya ditampilkan pada LCD 2 x 16, sebagaimana diagram blok-nya ditunjukkan pada Gambar 1.

Terdapat 2 tahapan proses utama dalam sistem identifikasi kata ucapan kata tunawicara yang dirancang-bangun, yaitu tahap pelatihan dan pengujian. Tahap pelatihan merupakan proses untuk melatih atau membimbing sistem agar dapat mengenali pola dari kata ucapan tunawicara. Sedangkan tahap pengujian merupakan proses untuk mengetahui kemampuan sistem dalam pengenalan pola berdasarkan tahapan pelatihan yang telah dilakukan.



Gambar 1 Bagan alir sistem identifikasi kata ucapan tuna wicara

2.2 Data Penelitian

Metode JST *backpropagation* membutuhkan data masukan, dalam hal ini menggunakan fitur suara yang telah diekstraksi menggunakan metode MFCC. Adapun data yang digunakan pada penelitian ini ditunjukkan pada Tabel 1 dan 2.

Tabel 1 Data untuk pelatihan

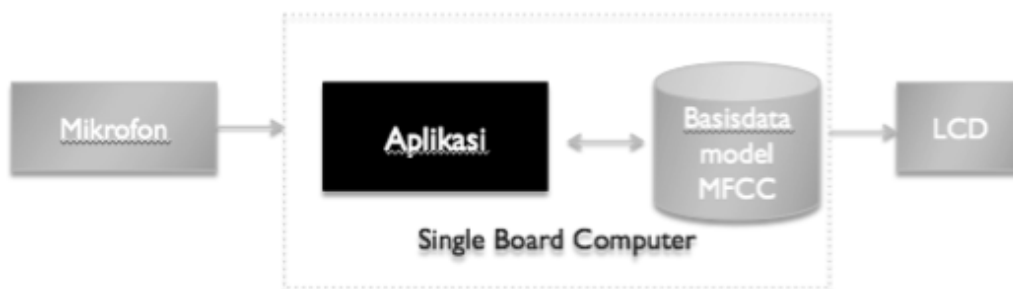
No.	Kata	Jumlah Suara
1	Harga	150
2	Kerja	150
3	Makan	150
4	Lapar	150
5	Kamar	150

Tabel 2 Data untuk pengujian

No.	Kata	Jumlah Suara
1	Harga	25
2	Kerja	25
3	Makan	25
4	Lapar	25
5	Kamar	25

2.3 Perancangan Sistem dan Alat Identifikasi Kata

Sebagaimana ditunjukkan pada Gambar 1, yang berfungsi sebagai masukan adalah sensor suara atau mikrofon, *Single Board Computer* (SBC) sebagai pemroses dan LCD sebagai luaran.



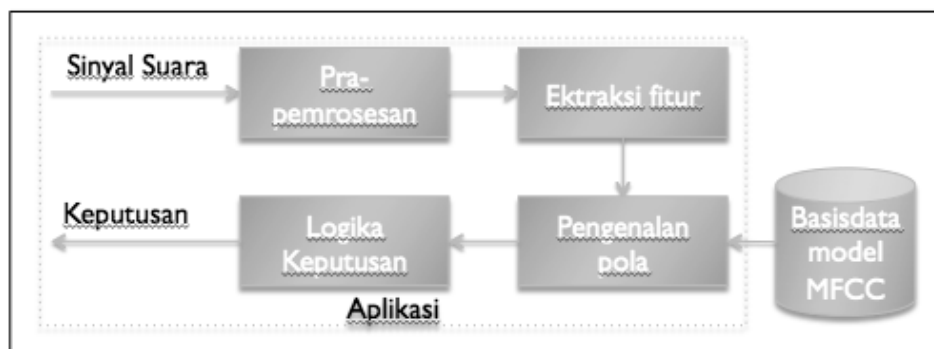
Gambar 2 Diagram blok sistem dan alat identifikasi kata

2.3.1 Mikrofon

Mikrofon akan menangkap suara tunawicara sebagai masukan untuk sistem, lalu suara tersebut akan diproses oleh sistem yang sudah tertanam pada SBC.

2.3.2 Single Board Computer (SBC)

SBC digunakan untuk memproses sinyal suara, termasuk proses pengenalan pola, sehingga akan diketahui arti dari sinyal suara tersebut, hasil tersebut akan ditampilkan pada LCD. Aplikasi yang tertanam pada SBC bertugas untuk membaca masukan dari mikrofon, melakukan pra-pemrosesan, ekstraksi fitur, pengenalan pola (perbandingan fitur masukan ke basisdata model MFCC) dan menampilkan arti suara melalui LCD. Diagram blok aplikasi yang dijalankan SBC ditunjukkan pada Gambar 3.



Gambar 3 Diagram blok aplikasi

Alur proses yang dijalankan pada SBC meliputi:

a. Pra-pemrosesan

Sinyal yang ditangkap oleh mikrofon akan ditapis menggunakan penapis *lowpass* untuk menghilangkan atau mereduksi derau atau *noise*, meningkatkan energi sinyal dan melakukan normalisasi dari sinyal mentah ke dalam sinyal berskala sama.

b. Ekstraksi fitur

Setelah dilakukan pra-pemrosesan sinyal, proses selanjutnya adalah pencarian fitur menggunakan algoritma MFCC.

c. Pengenalan pola

Melakukan perbandingan pola dari pola masukan mikrofon dengan pola-pola yang tersimpan pada basisdata model MFCC.

d. Logika keputusan

Setelah pola dikenali, maka akan diperoleh arti kata dari suara yang ditangkap oleh mikrofon, arti kata akan ditampilkan di LCD.

2.3.3 Basisdata model MFCC

Basisdata ini menyimpan model fitur dari suara yang telah melewati berbagai proses, yaitu:

1. Ekstraksi fitur

Ekstraksi fitur speech akan menggunakan MFCC yang nantinya akan menghasilkan vektor fitur MFCC. Fitur MFCC dihitung menggunakan parameter-parameter berikut:

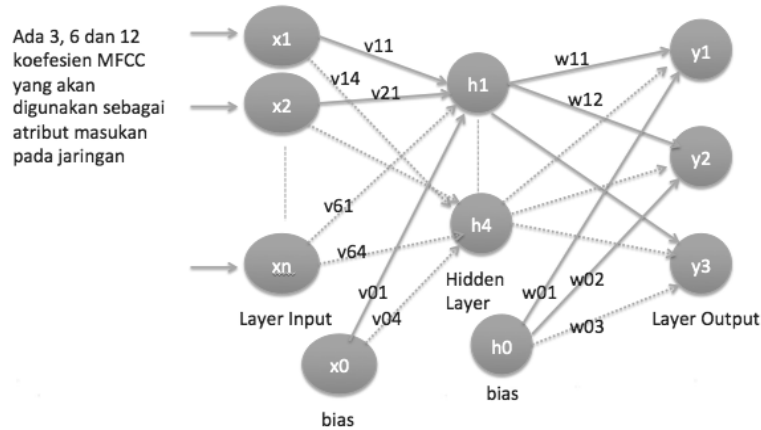
- *Pre-emphasis coefficients* = -0,95
- *Frame size* = 256 samples (16ms)
- *Frame overlap* = 85 samples (5,3ms)
- *Number of Triangular bandpass filters* = 20
- *Number of MFCCs* = 6 dan 12

2. Klasifikasi fitur

Klasifikasi fitur akan menggunakan JST *backpropagation* (seperti ditunjukkan pada Gambar 4) dengan *multi-layer perceptron* yang memiliki fase maju (*forward*) dan fase balik (*back*) dan fitur yang akan digunakan adalah MFCC. Adapun hasil pelatihan menggunakan JST yang terbaik dalam kasus ini adalah yang memiliki nilai *epoch* dan *error* terkecil. Fitur MFCC berupa vektor (bernilai bilangan asli) dari fitur sinyal yang dihasilkan dari proses *frame blocking & windowing*, FFT, *Mel Frequency wrapping*, Log dan *Discrete Cosine Transform* (DCT). Konfigurasi yang digunakan JST sebagai berikut :

- *Input layers* = 6 dan 12 nodes
- *Hidden layers* = 4 nodes

- *Output layers = 3 nodes*
- *Hidden layer transfer function = Hyperbolic tangent*
- *Output layer transfer function = Linear*
- *Epoch = 0, MSE = 1.*
- *Maksimum Epoch = 100.000,*
- *Target Error = 1e-010*
- *Learning Rate = 0,2*

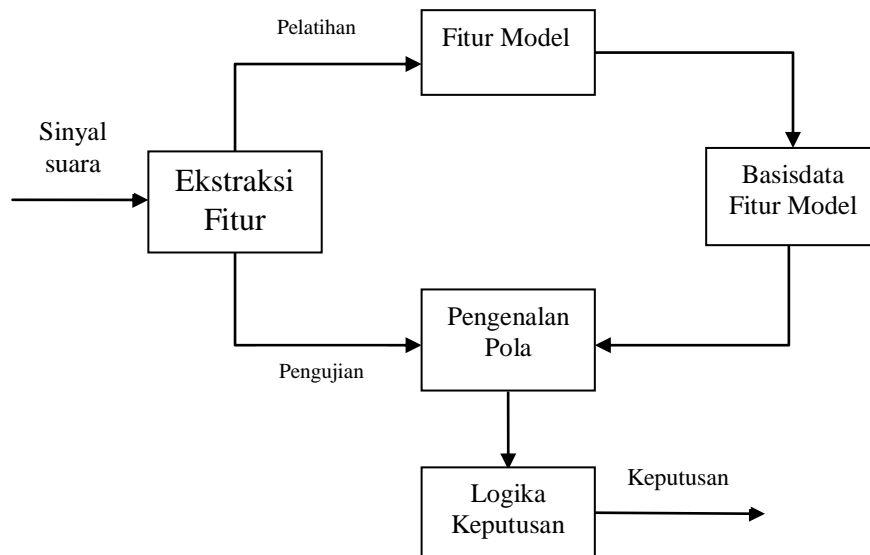


Gambar 4 Arsitektur pelatihan JST

Jumlah node masukan sama dengan jumlah koefisien yang diperoleh dari MFCC, jadi menggunakan 6 dan 12 node masukan.

3. Pelatihan dan Pengujian Fitur

Kegiatan pelatihan tergolong berat karena banyaknya proses yang dilakukan, oleh karena itu dilakukan menggunakan komputer dan perangkat lunak Matlab. Diagram blok pelatihan dapat ditunjukkan pada Gambar 5.



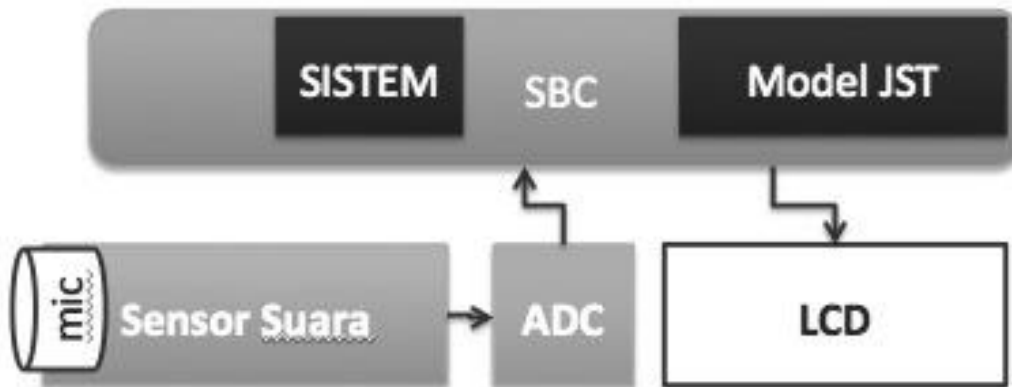
Gambar 5 Blok diagram klasifikasi fitur

2.3.4 LCD

LCD sebagai sarana luaran, menampilkan arti dari kata yang sudah ditentukan oleh sistem.

2.3 Perancangan Sistem

Mikrofon digunakan sebagai masukan untuk sistem, sebagaimana ditunjukkan pada Gambar 6. Mikrofon, sebagai sensor suara, menghasilkan sinyal analog yang kemudian dilewatkan ADC (*Analog to Digital Converter*) agar diperoleh sinyal digital terkait. Sinyal digital tersebut menjadi masukan untuk SBC yang sudah dilengkapi dengan modul JST.



Gambar 6 Blok diagram alat

Setelah alat dirakit (sebagaimana ditunjukkan pada Gambar 10) dan ditanamkan program terkait kedalamnya, selanjutnya dilakukan pengujian. Gambar 7 menunjukkan saat alat dihidupkan untuk pertama kali, yaitu melakukan proses inisialisasi, kemudian beberapa saat alat akan siap untuk menerima data suara yang akan dikenali (ditunjukkan pada Gambar 8). Kata yang diucapkan, akan ditangkap oleh sensor, lalu akan di proses oleh sistem yang sudah ditanam di SBC, lalu akan menampilkan hasil, sebagaimana ditunjukkan pada Gambar 9.



Gambar 7 Proses inisialisasi



Gambar 8 Alat siap untuk menerima masukan kata



Gambar 9 Hasil diteksi suara pada alat



Gambar 10 Tampilan alat yang sudah dirakit

3. HASIL DAN PEMBAHASAN

3.1 Hasil Penelitian

Hasil pelatihan dan pengujian JST *Backpropagation* menggunakan beberapa kombinasi *learning rate*, maksimal *epoch* dan nilai toleransi *error* ditunjukkan pada Tabel 3.

Tabel 3 Hasil akurasi pelatihan dari masing-masing penutur

a. *Learning rate* = 0,2, maksimal *epoch* = 100.000, toleransi *error* = 0,05

Data Penutur	Koefisien	
	6	12
1	40%	50%
2	60%	60%
3	60%	50%
4	60%	50%
5	50%	50%
Rata-rata	54%	52%

b. *Learning rate* = 0,03, maksimal *epoch* = 100.000, toleransi *error* = 0,00001

Data Penutur	Koefisien	
	6	12
1	50%	50%
2	50%	60%
3	60%	50%
4	60%	50%
5	60%	50%
Rata-rata	56%	52%

c. *Learning rate* = 0,03, , maksimal *epoch* = 1.000.000, toleransi *error* = 0,00001

Data Penutur	Koefisien	
	6	12
1	60%	50%
2	50%	60%
3	70%	50%
4	60%	50%
5	50%	50%
Rata-rata	58%	52%

3.2 Pelatihan dan Pengujian Sistem

Tabel 3a, Tabel 3b, dan Tabel 3c menunjukkan bahwa masukan JST *Backpropagation* menggunakan 6 koefisien memiliki rerata akurasi lebih baik, yaitu 54%, 56% dan 58%. Berdasarkan rerata akurasi koefisien 6 dan 12 (yang diperoleh dari rata-rata akurasi semua penutur) dapat disimpulkan bahwa hasil pada Tabel 3c mendapat hasil yang paling tinggi, menggunakan masukan 6 koefisien MFCC dan dilatih menggunakan JST dengan parameter learning rate = 0.03, maksimal epoch = 1.000.000 dan toleransi error = 0.00001.

Tabel 4 Hasil akurasi pengujian dari masing-masing penutur

Data penutur	Koefisien	
	6	12
1	54%	50%
2	50%	50%
3	60%	50%
4	50%	34%
5	54%	50%
Akurasi Pengujian (%)		

Proses pengujian mendapatkan hasil yang berbeda-beda, sebagaimana ditunjukkan pada Tabel 4. Rata-rata akurasi tertinggi untuk semua penutur sebesar 54% ketika menggunakan MFCC dengan 6 koefisien.

3.3 Pengujian Alat Identifikasi Kata Ucapan Tunawicara

Selain melakukan pengujian fungsional, dilakukan juga pengujian secara *realtime* menggunakan alat. Hasil akurasi, presisi dan sensitivitas dari masing-masing penutur ditunjukkan pada Tabel 5. Pengujian menggunakan gabungan dari semua data penutur mendapatkan akurasi sebesar 40% menggunakan 6 koefisien MFCC. Tabel 5 juga menunjukkan akurasi tertinggi ketika menggunakan data penutur ke 3 yaitu 60%, pengujian menggunakan data penutur 1, 2, 4 dan 5 mendapatkan hasil yang sama yaitu 50%. Sedangkan untuk presisi dan sensitivity dari pengujian per penutur, memiliki persentase yang sama, yaitu 40% untuk penutur 1, 2, 3 dan 5 serta 50% untuk penutur 3.

Tabel 5 Hasil akurasi alat dari masing-masing penutur

Penutur	Akurasi	Presisi	Sensitivity
1	50%	40%	40%
2	50%	40%	40%
3	60%	50%	50%
4	50%	40%	40%
5	50%	40%	40%

4. KESIMPULAN

Telah dibuat sebuah sistem (alat) yang mampu mengidentifikasi kata ucapan seorang tunawicara menggunakan metode MFCC dan JST *Backpropagation*. Akurasi alat mencapai 50% pada pengujian menggunakan data penutur 1, 2, 4 dan 5 serta 60% ketika menggunakan data penutur 3. Presisi dan sensitivitas alat identifikasi mendapatkan persentase yang sama,

yaitu 40% untuk penutur 1, 2, 4 dan 5 serta 50% untuk penutur 3. Alat memiliki sensitivitas 40%, presisi 40%, akurasi 50% untuk pengujian menggunakan semua data penutur.

5. SARAN

Perlu dilakukan penambahan fitur yang digunakan sebagai masukan JST dan memilih sensor suara yang lebih baik lagi, yang mampu bertahan terhadap noise.

DAFTAR PUSTAKA

- [1] Hikmat, R. H., "Booklet Kementerian Sosial, Badan Pendidikan dan Penelitian Kesejahteraan Sosial Pusat Data dan Informasi Kesejahteraan Sosial Jakarta," halaman 54. 2012 [Online]. Available: http://perpustakaan.bappenas.go.id/lontar/file?file=digital/143165-%5B_Konten_%5D-Konten%20D72.pdf [Accessed: 13-Juli- 2019].
- [2] Chandra, E., Manikandan, K., dan Sivasankar, M., "A Proportional Study On Feature Extraction Method In Automatic Speech Recognition System," International Journal Of Innovative Research In Electrical, Electronics, Instrumentation And Control Engineering Vol. 2, Issue 1 , January 2014, ISSN(Online) : 2321 – 2004, ISSN(Print) : 2321 – 5526. [Online]. Available: https://ijireeice.com/wp-content/uploads/2013/03/IJIREEICE5C_s_siva_A_proportional.pdf [Accessed: 13-Juli-2019].
- [3] Srinivasan, V., Ramalingam, V., dan Arulmozhi, P., "Artificial Neural Network Based Pathological Voice Classification Using MFCC Features," International Journal Of Science, Environment And Technology, Vol. 3, No 1, 2014, 291 – 302, ISSN 2278-3687. [Online]. Available: <https://pdfs.semanticscholar.org/241b/313fd5758095d74abe8da7b8aa22e2348075.pdf> [Accessed: 13-Juli- 2019].
- [4] Chauhan, H.B. and Tanawala, B.A., "Comparative study of MFCC and LPC algorithms for Gujrati isolated word recognition," International Journal of Innovative Research in Computer and Communication Engineering, 3(2), pp.822-826. 2015 [Online]. Available: <https://pdfs.semanticscholar.org/7c26/a9a579e14a1034eee909a2f34f12f5e44ae3.pdf> [Accessed: 13-Juli- 2019].
- [5] Desai N. , Dhameliya Kinnal., dan Desai Vijayendra., "Feature Extraction and Classification Techniques for Speech Recognition: A Review," International Journal of Emerging Technology and Advanced Engineering, Volume 3, Issue 12, December 2013, ISSN 2250-2459, ISO 9001:2008 Certified Journal. 2013 [Online]. Available: https://ijireeice.com/wp-content/uploads/2013/03/IJIREEICE5C_s_siva_A_proportional.pdf [Accessed: 13-Juli-2019].
- [6] Wang, Y. and Lawlor, B., "Speaker recognition based on MFCC and BP neural networks," In 2017 28th Irish Signals and Systems Conference (ISSC) (pp. 1-4). IEEE. 2017 [Online]. Available: <https://ieeexplore.ieee.org/abstract/document/7983644> [Accessed: 13-Juli- 2019].
- [7] Asda, T.M.H., Gunawan, T.S., Kartiwi, M. and Mansor, H., "Development of Quran reciter identification system using MFCC and neural network," *Indonesian Journal of Electrical Engineering and Computer Science*, 1(1), pp.168-175. 2016 [Online]. Available:

- https://www.researchgate.net/profile/Teddy_Gunawan/publication/299433726_Development_of_Quran_Reciter_Identifier_Using_MFCC_and_Neural_Network/links/5919b70b0f7e9b1db6527de1/Development-of-Quran-Reciter-Identification-System-Using-MFCC-and-Neural-Network.pdf [Accessed: 13-Juli- 2019].
- [8] Srivastava, N., “Speech recognition using artificial neural network,” *International Journal of Engineering Science and Innovative Technology (IJESIT)*, 3(3), pp.406-408. 2014 [Online]. Available:<https://pdfs.semanticscholar.org/5f0c/0fa960bb59294c8351a315824d3582735ab5.pdf> [Accessed: 13-Juli- 2019].
- [9] Sivan, S. dan Gopakumar, C., “An MFCC Based Speaker Recognition using ANN with Improved Recognition Rate,” *International Journal of Emerging Technologies in Computational and Applied Sciences (IJETCAS)*. 2014 [Online]. Available:<http://iasir.net/IJETCASpapers/IJETCAS14-426.pdf> [Accessed: 13-Juli- 2019].
- [10] Sharma S., Shukla A., Mishra Pankaj, “Speech and Language Recognition using MFCC and DELTA-MFCC,” *International Journal of Engineering Trends and Technology (IJETT)* – Volume 12 Number 9 - Jun 2014. [Online]. Available:<http://ijettjournal.org/archive/ijett-v12p286>[Accessed: 13-Juli- 2019].