Makhraj Recognition of *Hijaiyah* Letter for Children Based on Mel-Frequency Cepstrum Coefficients (MFCC) and Support Vector Machines (SVM) Method

Lina Marlina[1], Cipto Wardoyo[2], W.S. Mada Sanjaya[3,4*], Dyah Anggraeni[3,4], Sinta Fatmala Dewi[3,4], Akhmad Roziqin[5], and Sri Maryanti[6]

[1]*dept. of Arabic Language Edu., Faculty of Education and Teaching*, *UIN Sunan Gunung Djati*, Bandung, Indonesia
[2]*dept. of English Language and Literature, Faculty of Adab and Humanities*, *UIN Sunan Gunung Djati*, Bandung, Indonesia
[3]*dept. of Physics, Faculty of Science and Technology*, *UIN Sunan Gunung Djati*, Bandung, Indonesia
[4]*Bolabot Techno Robotic Institute*, *CV Sanjaya Star Group*, Bandung, Indonesia
[5]*dept. of Madrasah Ibtidaiyah Teacher Edu., Faculty of Education and Teaching*,*UIN Sunan Gunung Djati*, Bandung, Indonesia
[6]*dept. of Biology Edu., Faculty of Education and Teaching*, *UIN Sunan Gunung Djati*, Bandung, Indonesia
*madasws@gmail.com

***Abstract*—*Makhraj* is the most important thing for Muslim to recite the Holy Quran properly besides of *Tajweed*. This paper describe the *Makhraj* recognition of *Hijaiyah* Letter for children education. To make the *Makhraj* recognition, the feature extraction is used Mel-Frequency Cepstrum Coefficients (MFCC) method and to classify the *Hijaiyah* letter use Support Vector Machines (SVM) method based on Python 2.7. The waveform analysis of each *Hijaiyah Makhraj* pronunciation shows the differences of each letter. The database of *Hijaiyah Makhraj* pronunciation using 12 feature extraction can be classified by SVM process.**

***Keywords*—*Makhraj* Recognition, *Hijaiyah* Letter, MFCC, SVM, Python.**

I. Introduction

Holy Quran is the living handbook for Muslims. Because of the importance to read the Holy Quran properly [1], every Muslim must pay attention the *Makhraj* to read the *Hijaiyah* letter (Arabic letter) [2]. *Makhraj* is the pronunciation to reciting the Holy Quran letter properly based on *Tajweed* which differentiated by the organ of speech to produce a letter like constant and vowel [3].

Speech recognition is a conversion of the speech audio data to the text [4]. The conversion process needs an audio signal to identified by the audio feature extraction and Machines learning with the result classifying the speech. The various methods of speech audio feature extraction, such as; Linear Predictive Coding (LPC) [5] [6] and Mel-Frequency Cepstrum Coefficient (MFCC) [7] [8] [9] [10]. The Machines learning method which used to classify the speech for example; Artificial Neural Networks (ANN) [7] [6] [11], Support Vector Machines (SVM) [7], Hidden Markov Model (HMM) [5], Principle Component Analysis (PCA), Adaptive Neuro-Fuzzy Inference System (ANFIS) [12], K-Nearest Neighbors (KNN) [13], Fuzzy Logic [14], and other. Speech recognition has been implemented in many field such as Robotic [15] [16] [17] [18], control/wireless comunication [19] [20] [21], criminal detection [22], *Makhraj* recognition [23] [3], language recognition [24] [25] [26], and other.

In this paper, the Speech Recognition method is used to identify the *Hijaiyah Makhraj* pronunciation. The audio processing is used Mel-Frequency Cepstrum Coefficients (MFCC) and Support Vector Machines (SVM) method to recognize the *Hijaiyah Makhraj* pronunciation based on Python 2.7. Then, each waveform of *Hijaiyah Makhraj* pronunciation is analyzed by 3 waveform analysis. Finally, the *Makhraj* recognition system is classified to distinguished the *Hijaiyah* letter and correcting the *Makhraj*.

The paper is organized as follows. In section 2 described the theoretical background of MFCC and SVM on details. In section 3 described the experimental design of method and system design. In section 4 described the Analysis and Result of the research. Finally, the concluding remarks are given in section 5.

II. Theoretical Background

A. Feature Extraction using Mel Frequency Cepstrum Coefficient (MFCC) Method

Mel Frequency Cepstrum Coefficient (MFCC) is the extraction method for characterizing the audio signal. The extraction value can be used as the object or individual identity. The feature extraction is the coefficient of cepstral which used to consider the perception of the human hearing system. MFCC becomes the most used extraction method, because of considered quite good in representing the signal. Fig. 1 show the diagram process of MFCC. [19]

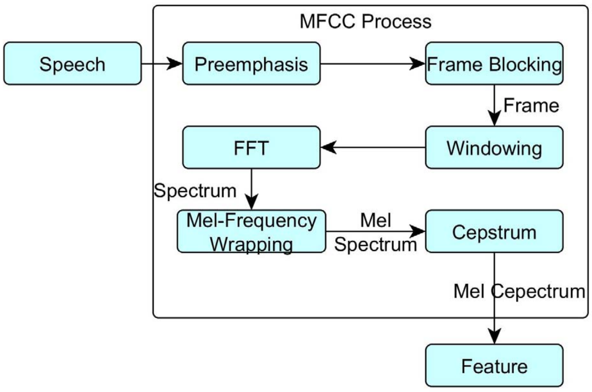


Fig. 1. MFCC process.

1) Preemphasis: Pre-emphasis is a filter process with the purpose to obtain a smoother spectral form of speech signal frequency and reduce a noise during sound capture. Pre-emphasis filter is required after the sampling process in the process of the speech signal. The pre-emphasis filter is based on the input/output relationship in the time domain on (1).

$$y(n) = x(n) - ax(n-1), \tag{1}$$

a is a pre-emphasis filter constant, and the value usually set as $0.9 < a < 1.0$.

2) Frame Blocking: Frame blocking is a segmentation of the audio signal into multiple overlapped frames. This process purpose to decreases the deletion of signals. This process continues until all signals have to get into one or more frames. By the short analysis, $x[n]$ is a long audio signal divided into some number of data frames. Each frame has N of the data sample of audio overlapping each other. The overlapping of N samples called as M which the value is not more than N or $N = 2xM$.

3) Windowing: Windowing is an analysis process of taking a sufficiently representative section from a long audio signal. This process removes the aliasing signal because of the discontinuity of the signal pieces by the Finite Impulse Response (FIR) digital filter approach. The discontinuities occur because of the frame blocking the process. The window define as $w(n), 0 \leq n \leq N-1$, N is the number of samples in each frame, the result of windowing is a signal present as (2).

$$y_1(n) = x_1(n)w(n), 0 \leq n \leq N-1, \tag{2}$$

$y(n)$ is the result signal of the convolution between the input signal and the window function, $x(n)$ represents the signal to be convolved by the window function. Where $w(n)$ is usually uses window Hamming which has the (3),

$$w(n) = 0.54 - 0.46cos(\frac{2\pi n}{N-1}), 0 \leq n \leq N-1. \tag{3}$$

4) Fast Fourier Transform (FFT): A function with a limited period can be expressed in Fourier series. Fourier transforms (FFT) are used to convert a time series of time-limited domain signals into a frequency spectrum. FFT is a fast algorithm of Discrete Fourier Transform (DFT) which is useful for converting each frame with N samples from the time domain into frequency domain and reduces the repeatable multiplication in the DFT.

$$X_n = \Sigma_{k=0}^{N-1} x_k e^{-2\pi jkn/N}, \tag{4}$$

(4) define that $j = \ sqrt{-1}$. $X[n]$ and $n = 0.1, 2, ..., N-1$ is the n-frequency of pattern generated from the Fourier transform, X_k is the signal of a frame. The result of this stage called Spectrum or Periodogram.

5) Mel-Frequency Wrapping: The mel scale is the unit on the frequency axis reflecting the perception of human speech. The lower the frequency, the narrower the interval, the higher the frequency, the interval will be wider. Apparently, humans can understand well the difference in sound heights at low frequencies, but increasingly higher frequencies are less likely to know the difference in pitch (high low-pitch in a sound). Equation 2.5 denotes the relation of the mel scale to the frequency in Hz is shown in (5).

$$F_{mel} = \{^{2595*[log]_{10}(1+\frac{F_{HZ}}{700}), F_{HZ}>1000}_{F_{HZ}, F_{HZ}<1000)}, \tag{5}$$

where F is the frequency in Hz and F_{mel} is the mel scale. Filter Bank is an approach the frequency spectrum in the mel scale with the working function as the human ear filter. FFT signal result is grouped into triangular filter file in Mel-frequency wrapping. The wrapping process to the signal in the frequency domain is performed by (6).

$$X_i = log_{10}(\Sigma_{k=0}^{N-1}|X(k)|H_i(k)). \tag{6}$$

From (6) define that $i = 1, 2, 3, ..., M$ (M is the number of triangle filters) and $H_i(k)$ is the value of the $i-$ triangle filter for the acoustic frequency of k.

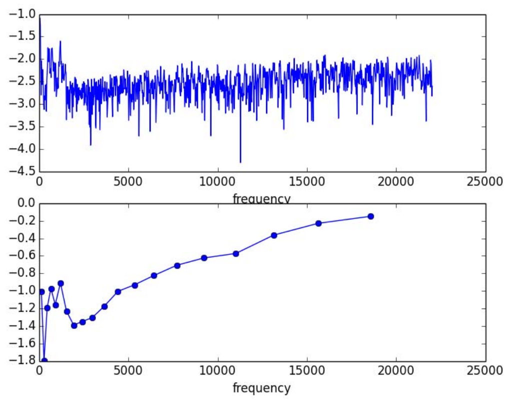


Fig. 2. The original amplitude spectrum and the Mel Bank filter.

6) Cepstrum: Humans can listen to the voice information based on time domain signals. At this section, the mel-spectrum be converted to the time domain using Discrete Cosine Transform (DCT), will get the result called mel-frequency cepstrum coefficient (MFCC). The cosine transformations shown on (7).

$$C_j = \Sigma_{j=1}^{K} X_j cos(j(i-1)/2\frac{\pi}{K}), \tag{7}$$

where C_j is the MFCC coefficient, X_j is the power spectrum of mel frequency, $j = 1, 2, 3, ..., K$, K is the number of desired coefficients, and M is the number of filters.

B. Machines Learning using Support Vector Machines (SVM) Method

Support Vector Machines (SVM) is a kernel based discriminative classification algorithm which proposed by Boser et al in 1992 [27]. The SVM concept can be explained simply as a search for the best hyperplane that serves as a separator of two classes in the input space. SVM is a binary classification algorithm. It is comprised of sums of the kernel function $k(xi; xj)$. [28]

$$f(x) = \Sigma_{i=1}^{N} \alpha_i t_i K(x_i, x_j + d). \tag{8}$$

From (8), $\Sigma_{i=1}^{N} \alpha_i t_i = 0$, $\alpha_i > 0$, and t_i represent the ideal outputs either $+1$ or -1 depends of the class which have a sample data. To decides the output class of certain test sample, $f(x)$ compare with the threshold. An one-vs-all approach usually adapted to achieve classification for multi-class data problem. The SVM train by the Gaussian RBF kernel have the data point x_i and x_j get from (9).

$$K(x_i, x_j) = exp(\gamma \parallel x_i - x_j \parallel)^2). \tag{9}$$

After multiple iterations on the train and test data, the optimal hyper-parameters γ and regularization constant C select for the SVM.

The advantages of SVM are effectiveness, low of memory, versatile, and common kernels are provided. The disadvantages of SVM are avoided the over-fitting in choosing Kernel functions and regularization term is crucial the number of features is much greater than the number of samples, and SVM do not directly provide probability estimates. SVM can be used as classifier such as; language recognition, speech recognition, hand-written character recognition, speaker recognition, object recognition, and other. [29]

III. Experimental Method

A. Method and System Design

The main hardware which used in this research is Personal Computer, Microphone, connections, and others. Fig. 3 is the illustration of *Makhraj* recognition of this research describe that; when the system ready to record, and human recites the *Hijaiyah* letter, the system will process the recognition and analyze the *Hijaiyah Makhraj* pronunciation result.

Fig. 4 is the general scheme of *Makhraj* recognition system which describes that after the system start to record *Hijaiyah*

Fig. 3. General scheme system of *Makhraj* recognition.

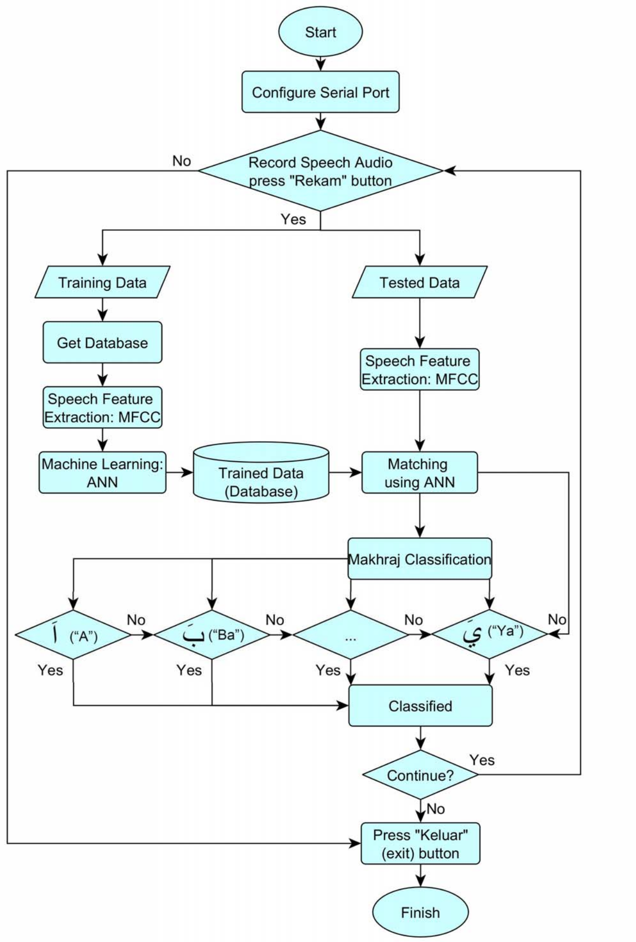


Fig. 4. Genaral scheme of *Makhraj* recognition system.

Makhraj pronunciation, the process divided by 2 processes: The first process makes a database using MFCC for features extraction of audio and SVM method to classifying the *Hijaiyah Makhraj* pronunciation. After that, the database called trained data. The second, the testing process with recording new audio of the *Hijaiyah Makhraj* pronunciation data will get the new feature extraction. Then, the new data matched with the Trained Data, classifying and analysis by using SVM method. The *Makhraj* recognition process based on Python 2.7.

B. Interface Design

The Graphical User Interface (GUI) of *Makhraj* recognition system based on Python 2.7 shown on Fig. 5. The interface consists by menu "Record" and "Exit", the shell windows of Python to monitoring the result and graphical interface to display the audio visualization of the *Makhraj* recognition shown on Fig. 5.

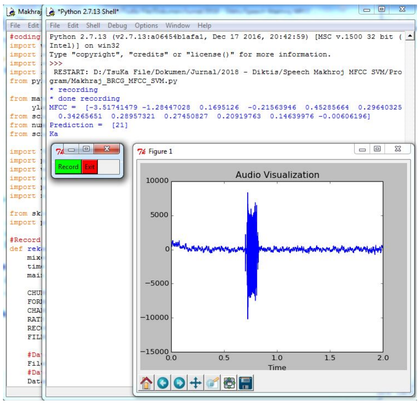


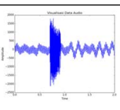
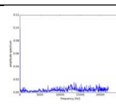
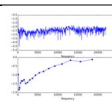
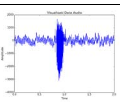
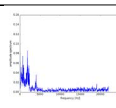
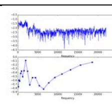
Fig. 5. The interface of *Makhraj* recognition system.

IV. Results and Discussion

A. Waveform Analysis

In this section, the *Hijaiyah Makhraj* pronunciation are analysis and compare each other. The data audio of *Hijaiyah Makhraj* pronunciation is compared by 3 analysis waveform, they are; the initial (audio visualization), FFT, and Mel by using MFCC feature extraction algorithm based on Python 2.7. For the first comparison, compare the similar pronunciation between letter أ ("A") and ء (" 'a") in the TABLE I. The audio visualization shows the waveform of letter ء (" 'a") is thin than letter أ ("A"). For FFT and Mel waveform analysis show the differences each other.

TABLE I
Comparison of letter between letter أ ("A") and ء (" 'a").

Hijaiyah Letter	Audio Visualization	FFT Waveform	Mel Waveform
أ ("A")			
ء (" 'a")			

Next, compare the other similar pronunciation between letter ح ("ha") and ه ("Ha") on the TABLE II. The audio visualization shows the waveform of letter ح ("ha") is thin than ه ("Ha"). FFT and Mel waveform analysis has a differences form.

And then, compare the similarity of some *Hijaiyah Makhraj* pronunciation of letter ج ("ja"), ذ ("dza"), and ز ("za") on TABLE III. From Audio Visualization on TABLE III is just a little differences appear on the waveform. But, on the other analysis like FFT and Mel show the difference of each waveform.

TABLE II
Comparison of letter between letter ح ("ha") and ه ("Ha").

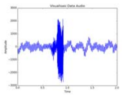
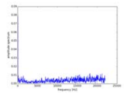
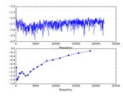
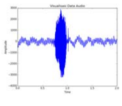
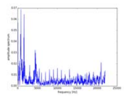
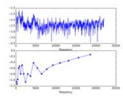
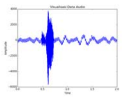
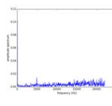
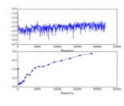
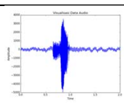
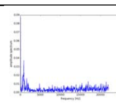
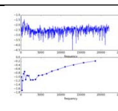
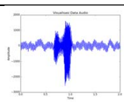
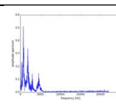
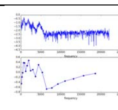
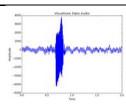
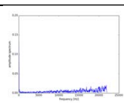
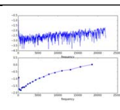
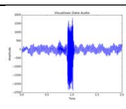
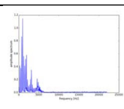
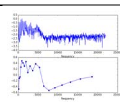
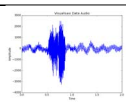
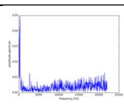
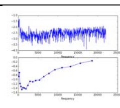
Hijaiyah Letter	Audio Visualization	FFT Waveform	Mel Waveform
ح ("ha")			
ه ("Ha")			

TABLE III
Comparison of letter; ج ("ja"), ذ ("dza"), and ز ("za")

Hijaiyah Letter	Audio Visualization	FFT Waveform	Mel Waveform
ج ("ja")			
ذ ("dza")			
ز ("za")			

Finally, compare the similarity *Hijaiyah Makhraj* pronunciation of letter ث ("tsa"), س ("sa") and ش ("sya"). From Audio Visualization on TABLE IV is a little differences of the waveform too. But, on the other analysis like FFT and Mel, the waveform is very different.

TABLE IV
Comparison of letter; ث ("tsa"), س ("sa") and ش ("sya").

Hijaiyah Letter	Audio Visualization	FFT Waveform	Mel Waveform
ث ("tsa")			
س ("sa")			
ش ("sya")			

With the waveform analysis using the initial, FFT, and Mel method, we can see the differences of each *Hijaiyah Makhraj* pronunciation waveform. Although the letter has a similar waveform in Audio Visualization, in other analysis (FFT and Mel), each *Hijaiyah Makhraj* pronunciation can be distinguished each other. Therefore, the *Makhraj* can be gone to the classification.

B. Building a Database

Hijaiyah letter divided by 28 letter from أ ("A") to ي ("ya"). To building a system that can recognize a *Makhraj*, it takes a collection of *Hijaiyah Makhraj* pronunciation audio to make a database. While recording *Hijaiyah Makhraj* pronunciation of each Hijaiyah letter, there is a different waveform with each other. Therefore, each *Hijaiyah Makhraj* pronunciation audio has its own characteristics.

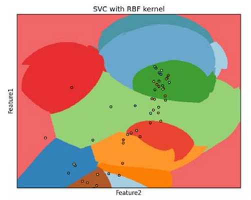
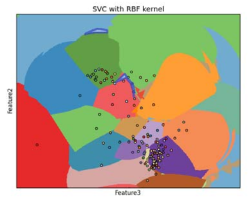
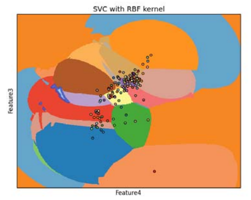
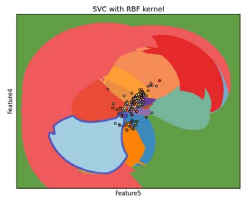
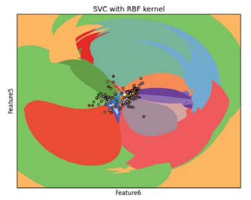
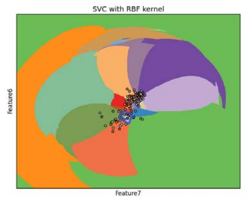
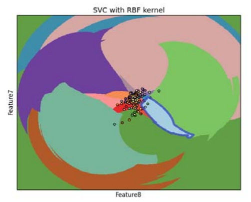
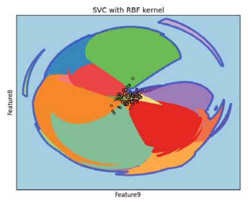
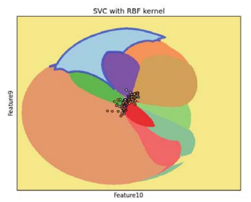
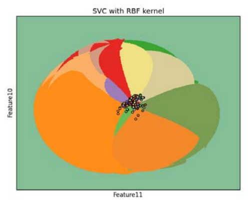
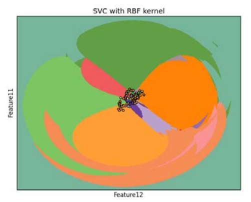
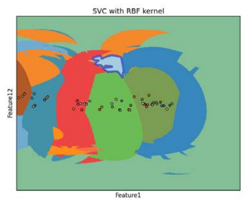
To develop the *Makhraj* recognition is need to collect the database of *Hijaiyah Makhraj* pronunciation data. To get the data characteristic of sound data is used MFCC method for the feature extraction. In this research, the data which will get the feature extraction is from the *Hijaiyah Makhraj* pronunciation audio data.

The database of *Makhraj* recognition is made from 12 feature extraction (from the coefficient of MFCC feature extraction) and 28 targets of *Hijaiyah Makhraj* pronunciation with 5 iterations for each letter. To distinguish data of each letter is use a target in the form of a value, "0" for letter أ ("A"), ..., and "27" for letter ي ("ya"). The database is collected on the ".txt" file, then the database is called as the Trained Data. The Trained Data will be used to classify and analyze the *Hijaiyah Makhraj* pronunciation using SVM machine learning.

C. Makhraj Classification

The database which made on the previous section is classified by SVM method with RBF kernel. TABLE V show the comparison of each features extraction of *Hijaiyah Makhraj* pronunciation for classification. From the classification shows that comparison on TABLE V (c) to TABLE V (k) cannot be classified because the distance of each feature extraction target is close together. But, on the TABLE V (a) to TABLE V (c), and TABLE V (l), the distance of each feature extraction target can be separate, thus each *Hijaiyah Makhraj* pronunciation can be classified.

TABLE V
THE CLASSIFICATION EACH FEATURE EXTRACTION OF HIJAIYAH MAKHRAJ PRONUNCIATION.

 a. Feature1 Vs Feature2	 b. Feature2 Vs Feature3
 c. Feature3 Vs Feature4	 d. Feature4 Vs Feature5
 e. Feature5 Vs Feature6	 f. Feature6 Vs Feature7
 g. Feature7 Vs Feature8	 h. Feature8 Vs Feature9
 i. Feature9 Vs Feature10	 j. Feature10 Vs Feature11
 k. Feature11 Vs Feature12	 l. Feature12 Vs Feature1

V. Conclusions

This study has been presented the development of the *Makhraj* recognition of *Hijaiyah* Letter for children education. This research is used MFCC and SVM method based on Python 2.7 to make the *Makhraj* recognition system. The waveform analysis shows that each *Hijaiyah Makhraj* pronunciation can be distinguished from each other. The database of *Makhraj* recognition used 12 feature extraction can be classified by SVM method. The future works of this research will be enhancing the classification of *Hijaiyah Makhraj* pronunciation by using Artificial Neural Networks (ANN) method or other deep learning.

References

[1] H. M. A. Tabbaa and B. Soudan, "Computer-Aided Training for Quranic Recitation," *Procedia - Social and Behavioral Sciences*, vol. 192, pp. 778–787, 2015. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S1877042815035636

[2] S. S. B. Hassan and M. A. B. Zailaini, "Analysis of Tajweed Errors in Quranic Recitation," *Procedia - Social and Behavioral Sciences*, vol. 103, no. Tq 1000, pp. 136–145, 2013. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S1877042813037634

[3] N. Arshad, S. A. Aziz, R. Hamid, R. A. Karim, F. Naim, and N. F. Zakaria, "Speech processing for makhraj recognition," *International Conference on Electrical, Control and Computer Engineering 2011 (InECCE)*, pp. 323–327, 2011.

[4] N. W. Arshad, S. N. Abdul Aziz, R. Hamid, R. Abdul Karim, F. Naim, and N. F. Zakaria, "Speech processing for makhraj recognition: The design of adaptive filter for noise removal," *InECCE 2011 - International Conference on Electrical, Control and Computer Engineering*, pp. 323–327, 2011.

[5] Thiang and Wanto, "Speech Recognition Using LPC and HMM Applied for Controlling Movement of Mobile Robot," *Seminar Nasional Teknologi Informasi*, 2010.

[6] Thiang and S. Wijoyo, "Speech Recognition Using Linear Predictive Coding and Artificial Neural Network for Controlling Movement of Mobile Robot," in *International Conference on Information and Electronics Engineering*, 2011.

[7] P. A. Sawakare, R. R. Deshmukh, and P. P. Shrishrimal, "Speech Recognition Techniques: A Review," *International Journal of Scientific & Engineering Research*, vol. 6, no. 8, pp. 1693–1698, 2015.

[8] A. Setiawan, A. Hidayatno, and R. R. Isnanto, "Aplikasi Pengenalan Ucapan dengan Ekstraksi Mel-Frequency Cepstrum Coefficients (MFCC) Melalui Jaringan Syaraf Tiruan (JST) Learning Vector Quantization (LVQ) untuk Mengoperasikan Kursor Komputer," Tech. Rep. 3, 2011.

[9] I. B. Fredj and K. Ouni, "Optimization of Features Parameters for HMM Phoneme Recognition of TIMIT Corpus," in *International Conference on Control, Engineering & Information Technology*, vol. 2. IPCO, 2013, pp. 90–94.

[10] E. S. Wahyuni, "Arabic Speech Recognition Using MFCC Feature Extraction and ANN Classification," *ICITISEE*, vol. 2, pp. 22–25, 2017.

[11] B. P. Das and R. Parekh, "Recognition of Isolated Words using Features based on LPC , MFCC , ZCR and STE , with Neural Network Classifiers," *International Journal of Modern Engineering Research*, vol. 2, no. 3, pp. 854–858, 2012.

[12] W. S. M. Sanjaya and D. Anggraeni, "Sistem Kontrol Robot Arm 5 DOF Berbasis Pengenalan Pola Suara Menggunakan Mel-Frequency Cepstrum Coefficients (MFCC) dan Adaptive Neuro-Fuzzy Inference System (ANFIS)," *Wahana Fisika*, vol. 1, no. 2, pp. 152–165, 2016.

[13] R. P. Gadhe, R. R. Deshmukh, and V. B. Waghmare, "KNN based emotion recognition system for isolated Marathi speech," *International Journal of Computer Science Engineering (IJCSE)*, vol. 4, no. 04, pp. 173–177, 2015.

[14] I. B. Fredj and K. Ouni, "A novel phonemes classification method using fuzzy logic," *Science Journal of Circuits, Systems and Signal Processing*, vol. 2, no. 1, pp. 1–5, 2013.

[15] I. N. K. Wardana and I. G. Harsemadi, "Identifikasi Biometrik Intonasi Suara untuk Sistem Keamanan Berbasis Mikrokomputer," *Jurnal Sistem dan Informatika*, vol. 9, no. 1, pp. 29–39, 2014.

[16] W. S. M. Sanjaya, D. Anggraeni, and I. P. Santika, "Speech Recognition using Linear Predictive Coding (LPC) and Adaptive Neuro-Fuzzy (ANFIS) to Control 5 DoF Arm Robot," in *ICCSE*. Bandung: IOP Conference, 2017.

[17] Z. H. Abdullahi, N. A. Muhammad, J. S. Kazaure, and F. A. Amuda, "Mobile Robot Voice Recognition in Control Movements," *International Journal of Computer Science and Electronics Engineering*, vol. 3, no. 1, pp. 11–16, 2015.

[18] D. Anggraeni, W. S. M. Sanjaya, M. Y. Solih, and M. Munawwaroh, "The Implementation of Speech Recognition using Mel-Frequency Cepstrum Coefficients (MFCC) and Support Vector Machine (SVM) method based Python to Control Robot Arm," *Annual Applied Science and Engineering Conference*, vol. 2, pp. 1–9, 2018.

[19] W. S. M. Sanjaya and Z. Salleh, "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, 2014.

[20] A. Kumar, P. Singh, A. Kumar, and S. K. Pawar, "Speech Recognition Based Wheelchair Using Device Switching," *International Journal of Emerging Technology and Advanced Engineering*, vol. 4, no. 2, pp. 391–393, 2014.

[21] K. P. Tiwari and K. K. Dewangan, "Voice Controlled Autonomous Wheelchair," *International Journal of Science and Research*, no. April, pp. 10–11, 2015.

[22] N. Zheng and X. Li, "A Robust Keyword Detection System for Criminal Scene Analysis Nengheng," *Proceedings of the 2010 5th IEEE Conference on Industrial Electronics and Applications, ICIEA 2010*, pp. 2127–2131, 2010.

[23] M. Subali, M. Andriansyah, and C. Sinambela, "Analysis of Fundamental Frequency and Formant Frequency for Speaker Makhraj' Pronunciation with DTW Method," in *Springer*. Springer, 2016, pp. 373–381.

[24] A. A. Almisreb, A. F. Abidin, and N. M. Tahir, "Arabic letters corpus based Malay speaker-independent," *Proceedings - 2013 IEEE 3rd International Conference on System Engineering and Technology, ICSET 2013*, pp. 232–236, 2013.

[25] M. S. Abdullah, M. M. Rahman, A. S. K. Pathan, and I. F. A. Shaikhli, "A practical and interactive web-based software for online Qur'anic Arabic learning," *Proceedings - 6th International Conference on Information and Communication Technology for the Muslim World, ICT4M 2016*, pp. 76–81, 2017.

[26] Z. A. Othman, Z. Razak, N. A. Abdullah, M. Yakub, and Z. B. Zulkifli, "Jawi character speech-to-text engine using linear predictive and neural network for effective reading," *Proceedings - 2009 3rd Asia International Conference on Modelling and Simulation, AMS 2009*, pp. 348–352, 2009.

[27] B. E. Boser, I. M. Guyon, and V. N. Vapnik, "A Training Algorithm for Optimal Margin Classiiers," *Proceedings of the fifth annual workshop on Computational learning theory*, pp. 144–152, 1992.

[28] H. Ali, A. Jianwei, and K. Iqbal, "Automatic Speech Recognition of Urdu Digits with Optimal Classification Approach," *International Journal of Computer Applications*, vol. 118, no. 9, pp. 1–5, 2015.

[29] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.