Dynamic Time Warping Features Extraction Design for Quranic Syllable-based Harakaat Assessment

Noraimi Shafie, Azizul Azizan, Mohamad Zulkefli Adam, Hafiza Abas, Yusnaidi Md Yusof, Nor Azurati Ahmad
Razak Faculty of Technology and Informatics
Universiti Teknologi Malaysia, Kuala Lumpur, Malaysia

***Abstract*—The use of technological speech recognition systems with a variety of approaches and techniques has grown exponentially in varieties of human-machine interaction applications. The assessment for Qur'anic recitation errors based on syllables utterance is used to meet the Tajweed rules which generally consist of *Harakaat* (prolonging). The digital transformation of Quranic voice signals with identification of Tajweed-based recitation errors of *Harakaat* is the main research work in this paper. The study focused on speech processing implemented using the representation of Quranic Recitation Speech Signals (QRSS) in the best digital format based on Al-Quran syllables and feature extraction design to reveal similarities or differences in recitation (based on Al-Quran syllables) between experts and student. The method of Dynamic Time Warping (DTW) is used as Short Time Frequency Transform (STFT) of QRSS syllable feature for *Harakaat* measurement. Findings from this paper include an approach based on human-guidance threshold classification that is used specifically to evaluate *Harakaat* based on the syllables of the Qur'an. The threshold classification performance obtained for *Harakaat* is above 80% in the training and testing stages. The results of the analysis at the end of the experiment have concluded that the threshold classification method for Minimum Path Cost (MPC) feature parameters can be used as an important feature to evaluate the rules of Tajwid Harakaat embedded in syllables.**

Keywords*—Speech processing; short time frequency transform; dynamic time warping; human-guided threshold classification

I. Introduction

The Tajweed Al-Quran guidelines [1], which outline the laws of Makhraj, Sifaat and Harakaat (MSH) scientifically have shown the comprehensiveness of rules and requirements for the proper pronunciation and articulation [2] of each syllable, forming precise recitation for each verse for Al-Quran recitation. This unique information is often embedded inconsistently in the Quranic recitation digital signal, due to different signal representation circumstances and recitation of the Quran by different readers [3]. The preparation of this recitation signal for further analysis, involving Digital Signal (Speech) Processing (DSP) up to the classification of Tajweed rules digitally, is not an easy task as it involves technical solutions, moreover religious opinions and sensitivity need to be factored in.

Difficulties in managing the complexity of recitation signals[3], especially in selecting a reasonable DSP method for the smooth implementation of every stage involved, are among the aspects of the study emphasized in this paper. The discovery of the error of Tajweed-based recitation [4] was completed by contrasting the comparability properties of the discourse signal between the learner and the expert recitation signal. Stochastic reading discourse is a very interesting issue for a person to show or evaluate reading errors based on technology. The specific model should be able to deal with even a relatively miniature form of differences.

The ability of DSP to reveal important features and classification approaches is very promising and becomes an important pillar in facing and solving the complexity of speech or voice signal analysis. Phoneme duration features and speech models are used to recognize some rules of reciting the Quran [5]. Therefore, an approach to reveal the characteristics and classification of time series to represent speech or voice using DSP techniques will be highlighted in this paper.

Harakaat, as one of the Tajweed rules in the recitation of the Qur'an [6], is the process of prolonging the pronunciation of syllables. In general, the law of Harakat Tajweed requires the prolonging to be performed on the particular syllable pronunciation guided by the character of ي (Ya), ا (Alif) and (وْ) (Wau). In addition, the prolongation of the syllable pronunciation [7] at the end of a sentence also involves 2, 4 or 6 Harakaat. The Harakaat rules, known as *Madd* [8] are generally divided into 12 types.

The structure of this paper begins with Section I which explains about the introduction to the complexity in speech or the nature of the voice in reciting the Qur'an. Section II explains about the type of reading rate of the Al-Quran that specifically involves the law of Tajwid Harakaat. Next, the work related to Harakaat analysis is explained in Part III. This is followed by a discussion about the Dynamic Time Warping (DTW) method for feature vectors in Part IV. Descriptions of the extraction design work and a description of the relevant experiments that have been carried out are included in Section V and VI respectively. Part VII of this paper explains the experimental results and the discussion of issues related to the results of the analysis. Finally, the conclusions from the research results in this paper are summarized in Section VIII.

II. Al-Quran Recitation Speed Rate

Islam requires its followers to read the Qur'an in an orderly and prudent manner. Furthermore, reading it with melodious and serious appreciation while following Tajweed rules [5] is very much demanded. There are four types of recitation speed rate which relate to Harakaat Tajweed [9].

A. At-Tahqiq (التحقيق)

This is a common reading type for beginners who are just learning Tajweed. The way it is read is similar to at-Tartil, but at-Tahqiq is identified as slower and calmer. The recitation of at-Tahqiq is also as practiced and considered as an initial stage for a novice reciter before the recitation improves to the next level.

B. At-Tartil (الترتيل)

In terms of terminology, at-Tartil carries the meaning of slow reading. This is in line with the tafsir stated in Tafsir Ibn Katsir, where tartil means reading the Qur'an in accordance with the law of Tajweed.

This means each of these readings is given the rights and properties of the letters that should be. Basically, reading with the al-Tartil approach allows the reader to understand and reflect on the verses of the Qur'an.

C. At-Tadwir (التدوير)

At-Tadwir is a type of reading in which the reading rate is intermediate between slow reading and fast reading. Usually, at-Tadwir recitation is practiced in congregational (Jama'ah) prayers, where the reference can be observed easily based on the recitation of Harakaat, known as Mad Munfasil which is recited not more than 6 Harakat.

D. Al-Hadar (الحدر)

Al-Hadar is defined as a way of reading at a relatively fast pace within the laws of Tajweed. In fact, Al-Hadar is the fastest level of reading and usually practiced by memorizers of the Qur'an who requires repetition in the process of memorization. A simple reference can be observed when the reading involved in the law of Tajweed namely Mad Munfasil, the reading is read with only two Harakat only. However, even though the whole recitation is read quickly, the reader adheres to the relevant Tajweed laws in every verse read.

In this paper, the collection of data is implemented for the at-Tadweer category. This important Harakaat attributes are evaluated in the experimental works taking into account the differences in recitation speed rate in every category, producing different Harakaat duration requirements.

III. Related Works

One of Tajweed rules is the prolongation, which is related to rhythm of recitation. There are various types of prolongation or Harakaat. The approach of Harakaat analysis concerns on sequences of voiced or unvoiced sound because of recitation are related to the prosody. The production of voice or speech involves the movement of air from the lung to vocal tract towards the lips. The combination of voice production mechanism produces a variety of vibration and spectral-temporal composition that produce different speech sound and its prosody. Apparently, the Arabic phonetic sound was produced from the specific articulation places or regions in the vocal tract.

MFCC features are widely used to determine the type or error for Al-Quran word pronunciation or Tajweed rules [10]. In addition, there are further discussion on the uses of MFCC as features in Harakaat recognition or identification [11], [12] and the prosody characteristics of speech signals associated with Harakaat [13]. All this emphasizes that it has a high potential to be used to obtain important characteristics. Among the successes reported, one of them is encouraging results in the assessment of English pronunciation [14]. Likewise with the successful use of prosody features in identifying the correct recitation of the Qur'an [15].

IV. Dynamic Time Warping

Speech is a time varying process [16] in which the duration of a word and its sub-words varies randomly. Time alignment may be applied in recitation which is required to find the best alignment and ratio between utterance sequences of expert and learner [17]. The recitation feature vectors represent the syllable similarly length on time occurrence. Dynamic Time Warping (DTW) is used to warp two feature vector sequences in time. DTW is well known in speech recognition to cope with different speaking speeds [18].DTW is used as feature matching with the deployment of the technique of minimum Euclidean distance or another distance-based approach [19]. DTW measures the similarity between the two temporal sequences which may vary in time [17]. This technique is also been used for finding the optimal alignment between the two-times series if one time series may be "warped" non-linearly by stretching or shrinking it along its time axis [20][21]. The recitation is also re-mapped as if all utterances were produced from the same vocal channel. The recitation speech signal produced by the learner is warped with respect to the Quranic speech signal that belonged to the expert by using the DTW algorithm. As a result, the sequence of Quranic speech signal of expert reference, X, with the length of |X|, and the learner, Y, with the length of |Y| can be expressed as in equation 1.

$$X = x_1, x_2, ..., x_i, ..., x_{|X|}$$

and

$$Y = y_1, y_2, ..., y_j, ..., y_{|Y|} \quad (1)$$

The warp path, W, is then can be constructed as,

$$W = w_1, w_2, ..., w_K$$

where

$$\max(|X|, |Y|) \le K < |X| + |Y| \quad (2)$$

where K is the length.

And, kth element of the warp path is $w_k = (i, j)$

where i = index of time series of X, and j = index of time series of Y.

The optimal warp path is presented as 'local match' scores matrix by getting the lowest-cost distance warp path and the first frame optimal warp start at $w_1 = (1,1)$ where the distance of a warp path is given by,

$$Dist(W) = \sum_{k=1}^{k=K} Dist(w_{ki}, w_{kj}) \quad (3)$$

$Dist(W)$ is the cosine distance of warp W, and $Dist(w_{ij}, w_{kj})$ is the distance between the two data frame indexes of X and Y in the kth element of the warp path.

The lowest-cost path for the first frame is $D(1,1) = 0$ and can be calculated by,

$$D(i, j) = Dist(i, j) + \min[D(i-1, j), D(i, j-1), D(i-1, j-1)] \quad (4)$$

The minimum-cost alignment is determined from the optimal warp path that ends at $D(|X|, |Y|)$ by looking at the lowest-cost warp path. The lowest-cost warp path is indicated by the ratio of STFT features vector between expert and reciter recitation signals. The search grid is used as a one-to-one monotonic transformation of the time axis in which all the movements have equal weight.

V. Features Extraction Design for Harakaat

To represent the parameters of measurement for the duration of syllable pronunciation, DTW of minimum path cost parameters coefficient is used. DTW calculates the local stretch on the time axis for two-series objects to optimally map one (query) to the other (reference) by calculating the distance between unequal sequences of length [19]. Minimum path cost alignment calculation uses the cosine distance [17] to determine the optimal warp path from the start point of syllable utterance until the end of an utterance. For each of syllable's utterance the search grid is used with one-to-one monotonic transformation along the time axis. The lowest-cost path is calculated from the optimal warp and known as minimum path cost alignment parameters co-efficient (refer to equation 4). The comparison between Quranic syllable-based Recitation Speech Signal (QRSS) syllables is done by the recitation ratio calculation between the expert and the learner recitation based on Short Time Frequency Transform (STFT) features. The STFT is given by

$$X(k, m) = \sum_{n=0}^{N-1} x(n+m)\omega(n)W_N^{nk}$$

Where k, m = 0. 1. …, N-1

ω(n) = time shifted window function.

m = the amount of shift. (5)

The value of the sequence distance represents as the Minimum Path Cost (MPC) for evaluating the syllable characteristics in each period referring to the Harakaat Tajweed rule feature. The minimum path cost (MPC) is used to obtain the optimal alignment by computing all the possible cost path and determine the lowest overall cost from the path to achieve high similarity between the two temporal recitation speech signals. Fig. 1 shows the minimum path cost (MPC) parameters with optimal path. The steps to calculate the minimum path cost parameters are given as follows [22]:

1) Load the QRSS from the expert as a reference template and followed by learner waveform of the same syllable.

2) Calculate the STFT features for both QRSSs (with 25% window overlaps).

3) Construct the 'local match' score matrix as the cosine distance between the STFT magnitudes.

4) Use dynamic programming to find the lowest-cost path between the opposite corners of the cost matrix.

5) Find the cost of minimum path cost parameters of the two signals of recitation speech.

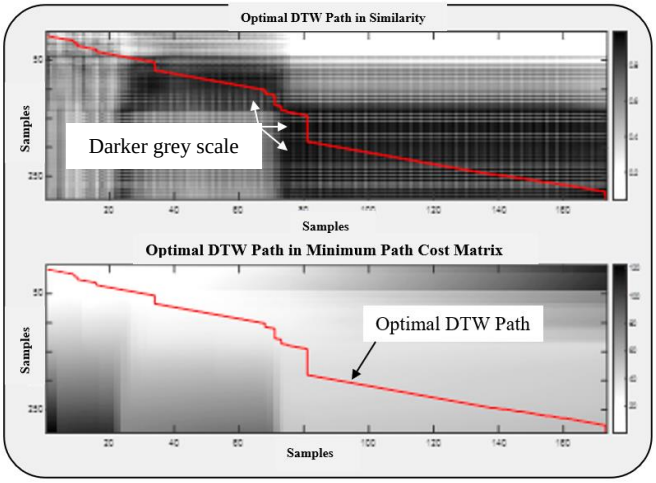

Fig. 1. The minimum path cost (MPC) parameters with optimal path.

The minimum path cost parameters value increases when the duration between two QRSSs is increased. Referring to the optimal DTW path in the similarity matrix, it can be observed that the darker grey scale shows a high similarity between the two QRSS readings. Thus, at the optimal DTW level in the minimum path cost, a lighter grey scale indicates a lower cost path (indicated in red) between the two QRSSs. The minimum path cost parameter is taken at the last value of the lowest cost path.

VI. Experiment

In the context of significant features of QRSS-syllable, the motivation of miniature feature creation remains towards the enrichment of the property characteristics by prosodic features that extracted by DTW. In this section, the *Harakaat* measurement is based on the time length for each syllable measured by using the DTW technique. Firstly, the QRSS-syllable is segmented into a number of frames and uniquely offering the Short Time Frequency Transform (STFT) feature [22] to be extracted using DTW technique. The cosine distance value for each frame is then calculated as the STFT miniature feature property (power energy). Later, the minimum cosine distance is obtained for every frame with the application of dynamic programming.

The minimum cosine distance is finally selected as the minimum path and is represented as the Minimum Path Cost (MPC) value. This means that the MPC parameters are

selected to measure the Harakaat for QRSS-syllables. Fig. 2 shows the different values of MPC for different syllables recited by 54 reciters for the Harakaat measurement performed.

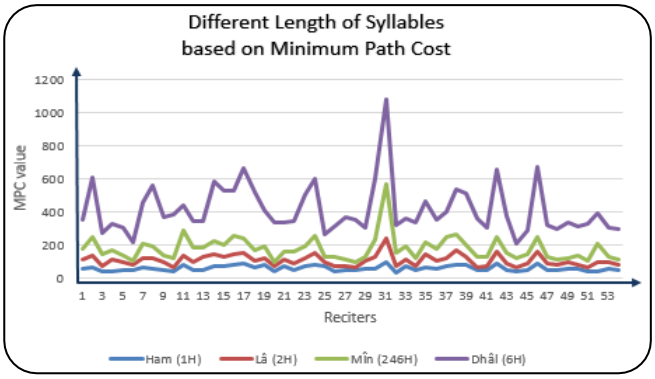

Fig. 2. Different syllables length based on Minimum Path Cost (MPC).

For example, the syllables *Ham*, *Lâ*, *Mîn* and *Dhâl* clearly indicate the lengths of time that differ from each other, which are one Harakaat (1H), two Harakaat (2H), four Harakaat (4H) and six Harakaat (6H) respectively. This is in line with the rule of Harakaat Tajweed for each of these syllables, which have different length of Harakaat respectively.

From the graph shown in Fig. 2, the syllable value of 1H has the lowest MPC value compared to 6H. As for 4H, the MPC values are between 2H and 6H, most of which are in the Tadweer category of speed. The speciality of syllable Min in having the multiple choices of Harakaat (2H, 4H and 6H) is limited to the two, four or six Harakaat (4H) for the measurement simplification.

The computational engine score threshold process is used to evaluate the similarity and the dissimilarity based on human-guided threshold process [23]. This successful threshold process shows how a conventional Talaqi process (experts evaluate the recitation by learners based on how to pronounce syllables in the verses of the Quran) are transformed to a machine evaluation approach. Computational engines must have salient features that can distinguish between correct and incorrect readings. Therefore, in determining the score, analysis of salient features, and matching process is used to obtain reading assessment based on the actual assessment by experts called a human-guided or Talaqi-like assessment [24].

Talaqi-like approach and process have been applied in the training and testing phases in the development of computational engine. This is to ensure that the assessment made by the machine are always guided by the human expertise. In this process, MPC parameters are used as representations to each syllable recited by the learners. Initially, the MPC values from recitations by experts were used as primary references in the threshold range assuming that all MPCs generated by expert readings were within acceptable thresholds by the system. Thereafter, learner's recitations were assessed by comparison based on this reference threshold range.

In the training phase, the recitation results by learners were also evaluated by experts to confirm whether learners' recitation is acceptable or not, considering that everyone is still able to comply with the Law of Tajweed despite recitation with different levels of voice. Therefore, the initial value of the threshold originally set for the machine will continue to change (i.e. not necessarily fixed) as long as the expert evaluating the recitation complying with the applicable Tajweed Law.

The training process on this machine with human-guided assessment continue to be repeated until all MPC data have been evaluated by a leading expert (human). After the training process is over, the value of the threshold range has been fully obtained and can be used as a benchmark for the recitations made by learners, that is, whether it is correct or otherwise (within or outside the range). The range for this threshold values is then used and tested for accuracy in the testing phase. The testing performance of this Talaqi-like approach is be measured and justified. Fig. 3 shows the Talaqi-Like assessment flow process.

For Harakaat validation, there are four types of classes are tested, namely as 1H, 2H, 4H and 6H according to their prolongation types.

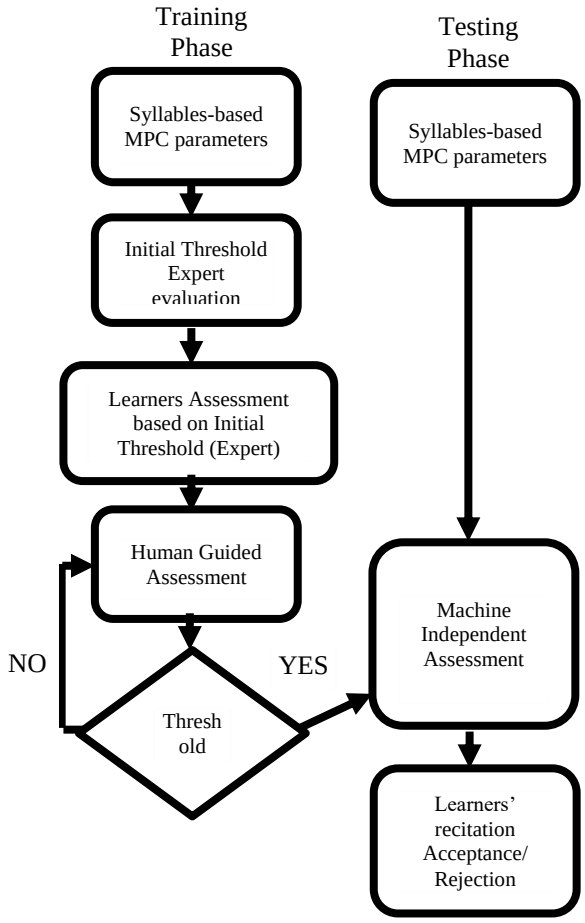

Fig. 3. The Talaqi-like assessment (Human Guided) flow process for Harakaat assessment.

In this classification approach, cross-validation is used to avoid overfitting of the model or class prediction [25]. Cross-validation is a resampling procedure that used to evaluate machine learning models on a limited data sample. It protects the overfitting by partitioning the data set into the fold and estimating the accuracy of each fold. This classification procedure used is k-fold cross-validation which data is divided into k-number of groups. The number of k-group that used in this classification is 5-fold.

The performance of classification is shown with the use of confusion matrix [26] by finding the True Positive Rate (TPR) and False Negative Rate (FNR). Both TPR and FNR are used to indicate the MPC parameters belong to the class or out from the class for each syllable Tajweed rules group or type. TPR shows that the true parameters of MPCs are included in a positive class. While FNR shows that true parameters of MPCs are included in a negative class. In other words, TPR and FNR indicate that parameters of MPCs are classified in the right or wrong classification, respectively.

Confusion matrix is a relatively popular measurement in summarizing the ability or performance of a group of algorithms in classification tasks [26]. In fact, in many cases, the accuracy of a classification result, can produce confusion if it has inconsistencies in the observations that need to be implemented for each classification. The main idea of employing the confusion matrix remains as commonly practice in any machine learning application. The result of the calculation of the confusion matrix is not only able to provide a better perspective on the modelling of a classification, but also at the same time able to identify the form or type of mistakes that have been made. Table I shows the classification performances by Linear Discriminant (LD), Support Vector Machine (SVM) and K-Nearest Network (KNN) for Harakaat measurement using MPC.

TABLE I. CLASSIFICATION PERFORMANCE OF HARAKAAT USING DIFFERENT CLASSIFICATION

Harakaat	LD (%)	SVM ((%)	KNN(%)
1H	98	99.4	96.8
2H	0	1.4	72.6
4H	7.4	13.5	75.5
6H	84.6	94.2	92.3

VII. RESULT AND DISCUSSION

The acquisition of MPC values can be considered as similarity parameters representing MPC the unique significant features of STFT features. Therefore, the training and testing phase classification are developed to assess the *Harakaat* Al-Quran recitation. However, the assessment of this *Harakaat* is limited to recitation from Malay adults because the data used in establishing the threshold range involves only Malay experts and students. This means that the threshold used is limited to the entire voice of Malay adults only.

In this training phase of the classification stage, the initial parameters of MPC are taken from the calculation of 12 Malay expert recitations to create an initial threshold range along with maximum and minimum values. As a result, the Acceptance Threshold (AT) is defined as,

$$\text{MinMPC} << y << \text{MaxMPC} \quad (6)$$

while y = accepted threshold MPC parameters.

Each syllable is represented by MPC parameters. These parameters are used to determine the pronunciation of the syllable is pronounced correctly or not, based on the rules of Harakaat Tajweed. These AT values are used based on the minimum and maximum value ranges of MPC to evaluate the syllables related to Harakaat. The starting point of training phase is when the input given to this designed system begins to create a change of pattern or minimum and maximum value that limits the correctness of a Tajweed in the reading of the surah. This is seemingly caused by the changes of the acceptable lowest and highest values of MPCs that correspondingly due to the variability demonstrated by various reciters but remains accepted (Acceptance Threshold) by the expert.

The process of correcting (or training) the minimum and maximum values (threshold range) is firstly performed on the group of experts' MPCs data. This is the initial threshold range and used as reference values to be compared with the learner recitations. Secondly, the MPCs values obtained from the recited syllables of 40 learners are matched with the expert threshold range. Besides the setting of minimum and maximum values, the indication of True Acceptance (TA), False Rejection (FR) and False Acceptance (FA) of the calculated MPCs are counted and accumulated.

Table II tabulates the MPCs values of syllables verse-2 of Al-Fatihah recited by 40 learners that have been matched with the threshold range of expert's recitations. Referring to the data shown, performance on Harakaat assessment for each syllable are highly accepted for syllable S5. It shows that the assessment based on expert and learners MPCs parameters are acceptable by using this approached.

TABLE II. TABULATES THE MPCS VALUES OF SYLLABLES VERSE-2 OF AL-FATIHAH

Range and Result Indication		Expert Max	Expert Min	TA	FR	FA	Performance (%)
Syllables of Verse-2	**S1**	68	25	34	6	0	85
	S2	77	39	30	10	0	75
	S3	63	16	36	4	0	90
	S4	53	14	31	8	1	78
	S5	69	25	22	17	1	55
	S6	60	22	33	7	0	83
	S7	36	14	31	9	0	78
	S8	75	17	36	4	0	90
	S9	89	44	31	7	2	78
	S10	61	27	29	11	0	73
	S11	152	30	33	6	1	83

In the testing phase, the main objective is to test the computational engine (from the training phase) that has been designed from the context of reliability of the miniature salient feature, extractor and classifier. The trained range of MPC is used to assess the performance of test data. Each syllable is tested according to the threshold determined based on MPC (maximum and minimum value). A total of 40 different learners from the training phase took their readings and the readings of each syllable in Al-Fatihah were extracted and matched with the reference MPC from the training phase. Each test data is also evaluated manually by an expert and the performance of the reading truth that refers to Tajweed rules is calculated in a technical context, namely true and false positive (TP and FP), false rejection (FR) and false acceptance (FA).

The comparison of errors was made and analyzed between the machine evaluation and human evaluation. From the results, the performance of the machine as an evaluator is then compared with respect to human expert performance.

Table III shows the performance of the parameters of MPCs that can be accepted, which consists of true acceptance (TA) and false acceptance (FA) in the testing phase. From that table, each syllable has been evaluated in terms of good of pronunciation (GOP) [25] for the testing data from Al-Fatihah verses based on Human Guided Assessment threshold from training phase.

In the test phase, the average performance show that each verse in the testing data is lower than training data. However, the average performance for testing data is still above 80% for each MPCs parameter and still shows good performance of Al-Quran recitation assessment. Fig. 4 visualizes the comparison between the performance percentages of training phase and testing phase where the testing gives a lower percentage in performance.

TABLE III. SHOWS THE PERFORMANCE OF THE PARAMETERS OF MPCS TESTING PHASE

Syllables verse 2	Human Machines Performance (%)
S1	87
S2	62
S3	92
S4	85
S5	95
S6	92
S7	85
S8	79
S9	64
S10	85
S11	64
Overall Score (%)	81

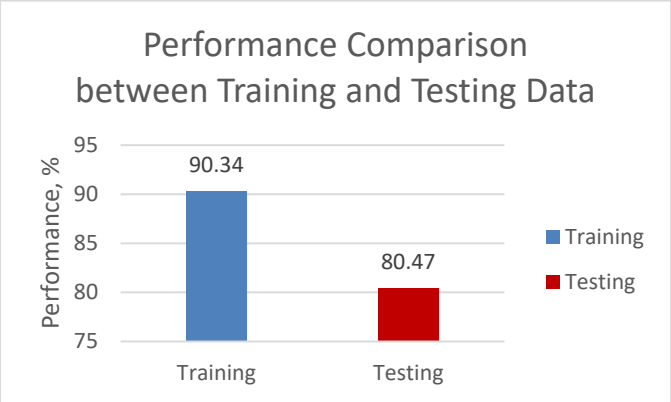

Fig. 4. The performance percentages of training phase and testing phase.

VIII. CONCLUSION

The classification performance using the threshold method for these MPC parameters can be used to evaluate syllable - based Al-Quran recitation where Tajweed Harakaat rules are embedded in the syllable. These features show good performance over 80% in testing phase for representing Tajweed Harakaat rules using DTW distance-features.The threshold method is very reliable in the training and testing phases giving a 90.34% and 80.47% respectively. For future work, the range of thresholds from various countries that are experts in reciting the Qur'an with the best Tajweed-based recitation can be included in experiments to improve computerized evaluations that are guided or learned from human expertise. In addition to that, the threshold for the entire verses of the Qur'an can be expanded.

ACKNOWLEDGMENT

The authors would like to express greatest appreciation to Universiti Teknologi Malaysia (UTM) and Ministry of Higher Education (MOHE), Malaysia for the financial support (grant number: Q.K130000.2456.08G28). We also acknowledged the support from the Cyber-physical Research Group and Advanced Informatics at Razak Faculty of Technology and Informatics (FTIR), UTM in this research work conducting this research.

REFERENCES

[1] T. Ridwan and N. Majid, “Development System for Recognize Tajweed in Qur’an using Automatic Speech Recognition,” in ICSTI 2018, 2018, doi: 10.4108/eai.19-10-2018.2281368.

[2] N. Shafie, M. Z. Adam, H. Abas, and A. Azizan, “Sequential classification for articulation and Co-articulation classes of Al-Quran syllables pronunciations based on GMM-MLLR,” AIP Conf. Proc., vol. 2401, no. October, 2021, doi: 10.1063/5.0072594.

[3] M. M. Al Anazi and O. R. Shahin, “A Machine Learning Model for the Identification of the Holy Quran Reciter Utilizing K-Nearest Neighbor and Artificial Neural Networks,” Inf. Sci. Lett., vol. 11, no. 4, pp. 1093–1102, 2022, doi: 10.18576/isl/110410.

[4] Y. Mohamed, M. Hoque, T. H. S. Bin Ismail, M. H. Ibrahim, N. M. Saad, and N. N. M. Zaidi, “Relationship Between Phonology, Phonetics, and Tajweed: A Literature Review,” in Proceedings of the 4th International Conference on Sustainable Innovation 2020, 2020, vol. 518, no. ICoSIHESS 2020, pp. 407–411, doi: 10.2991/assehr.k.210120.153.

[5] A. M. Alagrami and M. M. Eljazzar, "SMARTAJWEED Automatic Recognition of Arabic Quranic Recitation Rules," AIRCCSE 2020, pp. 145–152, 2020, doi: 10.5121/csit.2020.101812.

[6] D. Darmawan, A. Hanifan, R. Sri, and R. Waluyajati, "Learning to Recite the Qur ' an through Mobile Applications," in 1st International Conference on Islam, Science and Technology, ICONQUHAS, 2020, doi: 10.4108/eai.2-10-2018.2295515.

[7] H. M. A. Tabbaa and B. Soudan, "Computer-Aided Training for Quranic Recitation," Procedia - Soc. Behav. Sci., vol. 192, pp. 778–787, 2015, doi: 10.1016/j.sbspro.2015.06.092.

[8] M. K. Ilya Husna, Hidayati Azkiya, Suci Ramadhananti Febriani, Rahmawati, "The Analysis of Arabic Language Textbookfor Class II Elementary School," Lect. Pendidik., vol. 13, no. 2, pp. 234–245, 2022.

[9] A. Mohammed, M. Shahrizal, B. Sunar, and S. H. Salam, "Recognition of Holy Quran Recitation Rules Using Phoneme Duration," Recent Trends Inf. Commun. Technol., vol. 1, 2018, doi: 10.1007/978-3-319-59427-9.

[10] M. Computing, M. Rashwan, and S. Abdou, "Automatic Speech Recognition (ASR) Systems for Learning Arabic Language and Al-Quran Recitation : A Review," vol. 10, no. 7, pp. 91–100, 2021, doi: 10.47760/ijcsmc.2021.v10i07.013.

[11] B. Yousfi, A. M. Zeki, and A. Haji, "Holy Qur ' an Speech Recognition System Distinguishing the Type of prolongation," Sukkur IBA J. Comput. Math. Sci., vol. 2, no. 1, 2018.

[12] S. S. Alrumiah and A. A. Al-shargabi, "Intelligent Quran Recitation Recognition and Verification : Research Trends and Open Issues," Arab. J. Sci. Eng., 2022, doi: 10.1007/s13369-022-07273-8.

[13] B. Claire, M. Sawalha, T. Islam, J. Dickins, and E. Atwell, "Automatic Extraction of Quranic Lexis Representing Two Different Notions of Linguistic Salience : Keyness And Prosodic Prominence," J. Semit. Stud., 2018, doi: 10.1093/jss/fgy005.

[14] F. Chao, T. Lo, T. Wu, Y. Sung, and B. Chen, "Multi-aspect Modeling Approach to English Pronunciation Assessment."

[15] N. Bettayeb and M. Guerti, "Speech Synthesis System for the Holy Quran Recitation," Int. Arab J. Inf. Technol., vol. 18, no. 1, pp. 8–15, 2021.

[16] J. Zhao, X. Mao, and L. Chen, "Speech emotion recognition using deep 1D & 2D CNN LSTM networks," Biomed. Signal Process. Control, vol. 47, pp. 312–323, 2019, doi: 10.1016/j.bspc.2018.08.035.

[17] N. Shafie, M. Z. Adam, and H. Abas, "Al-Quran Recitation Speech Signals Time Series Segmentation for Speaker Adaptation using Dynamic Time Warping," J. Fundam. Appl. Sci., vol. 10, pp. 126–137, 2018, doi: 10.4314/jfas.v10i2s.11.

[18] M. Mundada and S. N., "Analysis of Speech and its Fluency Disorders," Int. J. Appl. Inf. Syst., vol. 10, no. 4, pp. 10–12, 2016, doi: 10.5120/ijais2016451484.

[19] T. Li, X. Wu, and J. Zhang, "Time series clustering model based on DTW for classifying car parks," Algorithms, vol. 13, no. 3, 2020, doi: 10.3390/a13030057.

[20] B. J. Mohan and R. Babu, "Speech Recognition using MFCC and DTW," in 2014 International Conference on Advances in Electrical Engineering (ICAEE), 2014, vol. 4, no. 3, pp. 1–4, doi: 10.1109/ICAEE.2014.6838564.

[21] N. Shafie, M. Z. Adam, S. Mohd Daud, and H. Abas, "A Model of Correction Mapping for Al-Quran Recitation Performance Evaluation Engine," Int. J. Adv. Trends Comput. Sci. Eng., vol. 8, pp. 208–213, 2019.

[22] A. H. A. Absa, M. Deriche, M. Elshafei-Ahmed, Y. M. Elhadj, and B. H. Juang, "A hybrid unsupervised segmentation algorithm for arabic speech using feature fusion and a genetic algorithm (July 2018)," IEEE Access, vol. 6, no. July, pp. 43157–43169, 2018, doi: 10.1109/ACCESS.2018.2859631.

[23] M. Maqsood, A. Habib, and T. Nawaz, "An efficient mispronunciation detection system using discriminative acoustic phonetic features for Arabic consonants," Int. Arab J. Inf. Technol., vol. 16, no. 2, pp. 242–250, Mar. 2019.

[24] M. Z. Adam, N. Shafie, H. Abas, and A. Azizan, "Analysis of Momentous Fragmentary Formants in Talaqi-like Neoteric Assessment of Quran Recitation using MFCC Miniature Features of Quranic Syllables," Int. J. Adv. Comput. Sci. Appl., vol. 12, no. 9, pp. 533–540, 2021, doi: 10.14569/IJACSA.2021.0120960.

[25] S. Kanters, C. Cucchiarini, and H. Strik, "The goodness of pronunciation algorithm: a detailed performance study.," Speech Lang. Technol. Educ. -SLaTE, no. 2, pp. 2–5, 2009, [Online]. Available: http://www.eee.bham.ac.uk/SLaTE2009/papers%5CSLaTE2009-33.pdf.

[26] S. Z. H. Naqvi, S. Aziz, M. H. Tariq, M. U. Khan, H. A. Aslam, and M. A. Imtiaz, "Effect of Al-Quran Recitation on Human Physiology," 2nd Int. Conf. Electr. Commun. Comput. Eng. ICECCE 2020, no. June, 2020, doi: 10.1109/ICECCE49384.2020.9179455.