Bulletin of Electrical Engineering and Informatics
Vol. 11, No. 5, October 2022, pp. 2817~2827
ISSN: 2302-9285, DOI: 10.11591/eei.v11i5.3715

Characteristics with opposite of quranic letters mispronunciation detection: a classifier-based approach

Tareq AlTalmas[1], Salmiah Ahmad[1], Nik Nur Wahidah Nik Hashim[2], Surul Shahbudin Hassan[3], Wahju Sediono[2]
[1]Department of Mechanical Engineering, Faculty of Engineering, International Islamic University Malaysia, Kuala Lumpur, Malaysia
[2]Department of Mechatronics Engineering, Faculty of Engineering, International Islamic University Malaysia, Kuala Lumpur, Malaysia
[3]Centre for Languages and Pre-University Academic Development, International Islamic University Malaysia, Kuala Lumpur, Malaysia

Article Info

Article history:

Received May 8, 2022
Revised Jun 23, 2022
Accepted Jul 24, 2022

Keywords:

Audio analysis
Classification
Feature extraction
Feature selection
Machine learning

ABSTRACT

Reading Quran for non-Arab is a challenge due to different mother tongues. learning Quran face-to-face is considered time-consuming. The correct pronunciation of *Makhraj* and *Sifaat* are the two things that are considered difficult. In this paper, *Sifaat* evaluation system was developed, focusing on *Sifaat* with opposites for teaching the pronunciation of the Quranic letters. A classifier-based approach has been designed for evaluating the *Sifaat* with opposites, using machine learning technique; the k-nearest neighbour (KNN), the ensemble random undersampling (RUSBoosted), and the support vector machine (SVM). Five separated classifiers were designed to classify the Quranic letters according to group of *Sifaat* with opposites, where letters that are classified to the wrong groups are considered mispronounced. The paper started with identifying the acoustic features to represent each group of *Sifaat*. Then, the classification method was identified to be used with each group of *Sifaat*, where best models were selected relying on various metrics; accuracy, recall, precision, and F-score. Cross-validation scheme was then used to protect against overfitting and estimate an unbiased generalization performance. Various acoustic features and classification models were investigated, however, only the outperformed models are reported in this paper. The results showed a good performance for the five classification models.

Corresponding Author:

Tareq AlTalamas
Department of Mechanical Engineering, International Islamic University Malaysia
Kuala Lumpur, Malaysia
Email: tareq.ma@live.iium.edu.my

1. INTRODUCTION

Pronunciation training is a long-time process, which needs a long face-to-face meeting between the trainers and the learners. Teaching the pronunciation of Quranic letters has been performed similar to any other languages. In non-Arab societies, the problem is worsen as the exposure to Arabic sounds is less, making it is difficult for learners to practice. The availability of qualified teachers is another challenge. Thus, automatic detection of the pronunciation errors by computers can be a complementary to the traditional methods. Students can perform self-learning at their places, at any suitable time then the final approval of the pronunciation can be done with fewer meetings with the teachers [1]-[3].

Mispronunciation detection can point out the reason for the mispronunciation [4], which can be achieved using two approaches, firstly, by using confidence measure, where an automatic speech recognition techniques are utilized to measure various statistical scores. Secondly, the classification-based approach, where the classification models are designed with some acoustic features as inputs to evaluate the pronunciations. Those acoustic features are still not accurately determined; therefore, this approach is still an open research area for more development [5].

There are various methods developed to accommodate language learning including the Quranic recitation that use Arabic language. There is a significant argument among the researchers whether the objective of the pronunciation training is to achieve a sound like a native speaker, or only pronounce intelligible sounds [6]. Many have agreed that intelligibility is essential, however, this is accepted when we are talking about speech as a medium of communication. In the case of Quranic recitation, it is compulsory to be pronounced as same as the prophet Muhammed (PBUH). Generally, errors are committed during pronunciation due to two factors, firstly, some sounds are missing from the first language of the learners. Secondly, some sounds origin at similar points of articulations, which caused the pronunciation of one letter to be like another letter pronunciation [7].

There were many multi-media-based systems developed related to Quranic learning; where students listen to the recitations, read the knowledge, and repeat the recitations many times. Thereafter, a quiz-type evaluation of the students is performed. In such systems, only the level of comprehension of the *Tajweed* knowledge is assessed, and not the recitation [8]. While much research was conducted on English and other languages, the Arabic language still lacking in research, and it is an open area for more research. Deep learning approach was proposed by [9] for detecting errors in the pronunciations of the Arabic language learners. The convolutional neural network-based (CNN) system has outperformed the traditional machine learning approach. The data used was clean without background noise, and the system was able to identify the pronunciation errors in 27 Arabic words. Maqsood, [5] proposed a classification-based approach for detecting the Arabic consonants' errors. The size of the Database was about 5600 recordings. A group of discriminative features were identified and used as an input to multilayer artificial neural network. The average accuracy of the system was about 82%. This work was performed on the word level by selecting a group of words that cover most of the Arabic letters, however, it is essential to study the Arabic letter pronunciation according to its approved way of pronunciation based on its *Makhraj* and *Sifaat*. Nazir [10] has investigated the traditional machine learning approach and the CNN for detecting the errors in the pronunciations of the Arabic letters. In the traditional machine learning approach, features were extracted using normal speech processing techniques and by extracting features from the convolutional layers of pre-trained CNN. The research was conducted on the letters that used to be mispronounced by the learners. The CNN-based approach defeated the other approaches and achieved a 92% average accuracy. A classification-based mispronunciation detection system has been introduced, it focused on five phonemes, and the data included 100 Pakistani speakers, ranging from the beginner level to the expert level. The system was aimed to classify the pronunciation either correct or wrong. Various acoustic features were tested, and the support vector machines (SVM) were used as a classification model. The system showed a comparative high overall accuracy, just about 97%. But the research focused only on five phonemes as compared to ours which is focused on all Quranic letters [11].

Three types of acoustic features to model the *Sifaat* without opposites the Quranic letters were investigated, which are mel-frequency cepstral coefficient (MFCC), formant frequencies and power spectral density (PSD). Two classification methods were used the linear discriminant analysis (LDA) and quadratic discriminant analysis (QDA). The best performance was achieved when a combination of the three feature vectors was created. The results showed that the QDA classifier has outperformed the LDA classifier, where the average accuracy was about 84% [12]. An integrated system for pronunciation error detection was established by [13]. It is a hybrid system that used speech recognition methods and classification-based methods. The average accuracy of the integrated system in evaluating the word-level pronunciation was about 91%. This system dealt with the Quranic verses as words and checked the pattern of the words, while it did not cater for the Quranic letters from the points of articulation and characteristics.

Secondly, the automatic speech recognition approach was reviewed. Web-application was developed by [14] for teaching children the recitation of Al-Quran. The comparison between the reference feature vector and the unknown feature vector was conducted using the dynamic time wrapping (DTW). The features used in this research were from MFCC. The Euclidean distance and the cosine distance were investigated where the cosine distance showed a better performance. The system was tested using 8 Arabic words, and the result showed the accuracy between 85% to 93%. A method was proposed by [15] to identify the pronunciation mistake in the letter ر while it is in 3 different places, start, middle and end of the words. Mel-frequency cepstrum coefficients (MFCCs) and its dynamic forms; Delta-MFCC and Delta2-MFCC were used as features. Both the Gaussian mixture modelling with a universal background model GMM-UBM and I-vector were tested for categorizing pronunciation errors. I-vector has outperformed in identifying the error when the letter ر is at the beginning of the word. The system was able to evaluate the pronunciation proficiency but not identifying the source of mispronunciation. A confidence measure approach was introduced by [16] to evaluate the proficiency of the pronunciation using the goodness of pronunciation (GOP) method of the five mostly mispronounced phonemes by Indian and Pakistani speakers. The overall system accuracy was between 87% to 100% for each letter of the five. A system was developed to evaluate the proficiency of the pronunciation of the Algerian learners who face difficulties in learning the Arabic language. The global average log-likelihood (GLL) score was used to evaluate the pronunciations. The system showed a good performance in evaluating the pronunciations of three words [17]. On the other hand, the hidden markov model (HMM) log-likelihood probability was used to score Malaysian teachers' pronunciation on their way to learn the Arabic language. Two databases were used in this research; the Database has only native speakers and the Database that includes the non-native speakers. The accuracy of the system with non-native speakers has outperformed and showed an accuracy of 89% [18].

Spectrogram's key feature are the formant frequencies. The tongue position affects the value of the first formant frequency, it also changes the pharyngeal space. There is a negative correlation between the mouth opening and the first formant frequency, where the more opening of the mouth means lesser value of the first formant frequency and vice-versa. The formant frequencies are proven as an efficient and compact form to represent the time-varying characteristics of speech. Mostly, three formants are used, where all are below 4 kHz, where vowels can be identified by the first and second formants, which lies in the range of 0.2 kHz to 0.7 kHz, the second and third formants usually lie in the range of 0.8 kHz to 2.3 kHz and 1.7 kHz to 3 kHz respectively [19]. The *Qalqalah* letters were analyzed using the formant frequencies extracted from spectrogram graphs [20]. Quranic letters can be identified using the formant frequencies, and the PSD can improve the accuracy of the system as compared to the formants [21]. On the other hand, many studies have utilized the PSD as a feature extraction technique for automatic speech recognition [22], [23].

From the literature, the developed systems based on machine learning methods are still new for Arabic language especially for the Quranic recitation application, and more research are needed related to this area. Furthermore, the classifier-based and deep learning systems has proven to outperform the automatic speech recognition approach-based systems, and it can detect the source of pronunciation errors. Therefore, this paper focused on building a new approach for the evaluation the pronunciation of the Quranic letters, to evaluate the *Sifaat* with opposites in the pronounced Quranic letters. The missing of any of these *Sifaat* affects the correctness of the pronunciation.

2. METHODOLOGY

2.1. *Sifaat* of quranic letters

Sifaat (characteristics) are the attributes of the Quranic letters. It helps in distinguishing the letters that share the same point of articulation (*Makhraj*). *Sifaat* are grouped into two main groups: *Sifaat* with opposites and *Sifaat* without opposites. Figure 1 shows the list of *Sifaat* with opposites of the 28 Quranic letters and letters that are classified to every group. Each Quranic letters must hold one *Sifaat* from each row as follow [24].

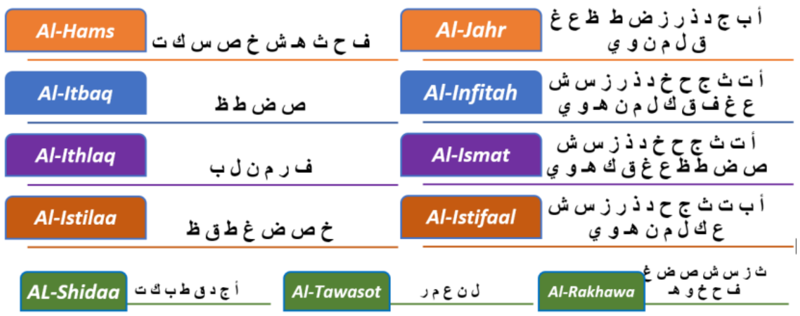

Figure 1. Quranic letters characteristics (*Sifaat*) and the opposites

2.2. Proposed classification-based approach

The selection between machine learning and deep learning approaches is a trade-off, which is affected by various factors, one of the most important factors is the amount of the available data for training the models. Indeed, the traditional machine learning approach has outperformed deep learning approach in the case of the lesser amount of available data for the model training [25]. The traditional machine learning approach has been chosen for this stage due to the limited database that is available. In this paper, a classifier-based approach is proposed to detect the incorrect pronunciation of *Sifaat* with opposites among the Quranic letters.

The proposed characteristics (*Sifaat*) errors detection locates the absence of the characteristics from the pronunciations. Here, the evaluation of the *Sifaat* with opposites is defined as classification problems, where each classification model evaluates one group of the five groups. The correctly pronounced letters are correctly categorized to its correct class for each *Sifaat* group. The process begins with choosing the letter to be evaluated, then the true classes for each pair of *Sifaat* will be called from the reference Database, as shown in Figure 2. The recorded sound goes through a successive step, starting from pre-processing to remove the background noise as the recordings were collected in a normal environment such as offices. Then the end-points detection was performed to keep only the segment of the letter from the full recording. Next, the extraction of five feature vectors that are used as input to the five classification models were conducted. Figure 2 shows the five feature vectors above each classification model, these feature vectors were used as inputs for the classification models to assess the *Sifaat* of the pronounced letter. Score 1 to score 5 represents the pass or fail results for each pair of *Sifaat.* Divide-and-conquer was used to divide the problem into several sub-problems by designing four binary classifiers for the first four pairs of *Sifaat* and a three-classes classifier for the fifth group of the *Sifaat*.

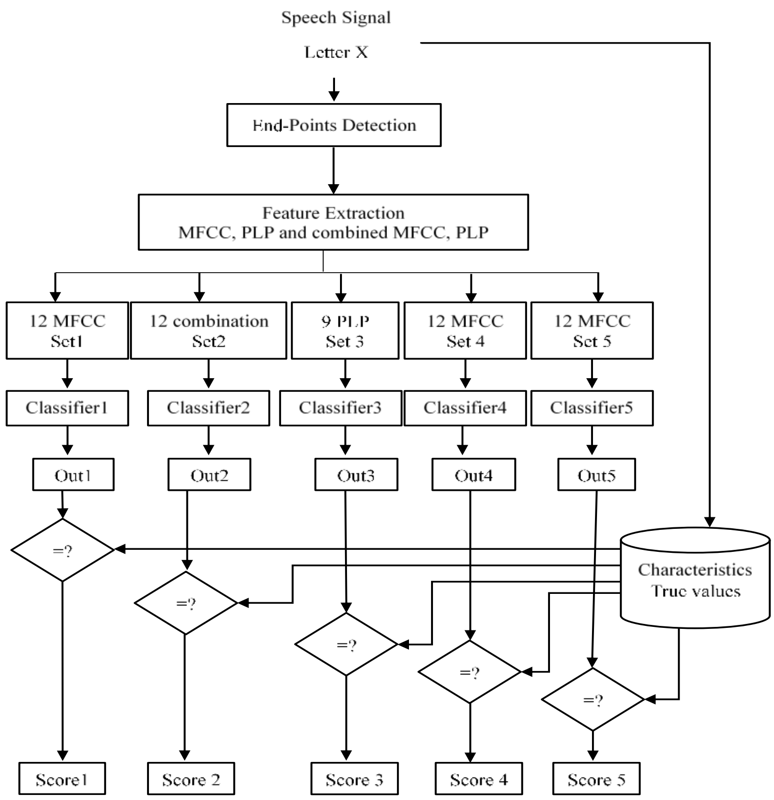

Figure 2. *Sifaat* assessment process block diagram

Figure 3 shows the process used in this paper to find the feature set and classification model to be used in each pair of *Sifaat* mentioned in Figure 1. In this paper, the development of a classification-based approach for evaluating the *Sifaat* with opposites of the Quranic letters' pronunciations is explained. The system is used to evaluate the presence or absence of the *Sifaat* with opposites in the Quranic letters' pronunciations. Designing the *Sifaat* evaluation system for Quranic letter pronunciation was started with collecting the appropriate audio database; unfortunately, there is no Database available for the Quranic letters' pronunciations according to the approved way in *Tajweed* knowledge. The step is followed with identifying the acoustic features that can represent each *Sifaat* group, using feature engineering step, which was implemented to reduce the size of feature vectors, remove the redundancy, and useless parameters. Two groups of acoustic features are extracted and stored in an excel files for each dataset. Various classification algorithms were trained with the full feature vectors and the reduced size feature vectors to find the best classification method for each group of *Sifaat*. The best performing models for each group of *Sifaat* are reported in this paper only.

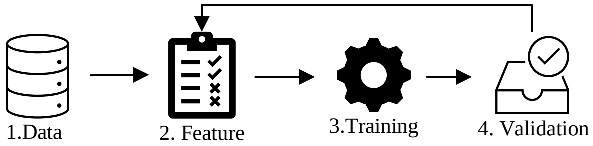

Figure 3. Classification models development block diagram

2.3. Database collection and preparation

Despite Arabic is the 5th spoken language in the world, there is not enough research were conducted in the field of the pronunciation training, and the availability of proper Database is still very minimum [11]. Most of the researchers tend to create their special Database that is suitable for their own purpose. In our research, the Database should be prepared such that the letter should be pronounced with *Sukun* on it and preceded by the *Hamza* (ء), since Quranic letters have its unique way of pronunciation. It is the main challenges in designing a computer-aided pronunciation training for the Arabic language, in having the availability of the Database according to *Makhraj* and *Sifaat* pronunciation. In this research, experts in Quranic recitation were asked to recite the 28 Quranic letters according to the approved way. The data includes male adults within the age of 20 to 35 years old. The samples were among the experts from both native and non-native speakers. A total of 75 experts were asked to recite the 28 letters. Table 1 shows each pair of the *Sifaat*, where the *Sifaat* was given class 0, while its opposites was given class 1, the number near to each *Sifaat* represents the number of samples in training data related to this group of *Sifaat*. However, Group 5 is divided into 3 *Sifaat*; therefore, it has three classes. The Database was recorded in a normal environment such as offices, classes, and rooms, to mimic the actual learning environment. Then, the data were presented to 2 certified experts in *Tajweed* knowledge to manually evaluate it in thorough listening, where the inaccurate pronunciations were taken out from the Database. 1612 recordings out of 2100 recordings passed the expert evaluations. Table 1 shows the total number of all letters that are approved for further used in the analysis. The number of records indicate the number of individual reciters who recited each letter.

Table 1. Labelling of each *Sifaat* to its class

	Sifaat pairs				
Class	Pair 1	Pair 2	Pair 3	Pair 4	Group 5
Class 0	*Al-Jahr (1041)*	*Al-Istifal (1201)*	*Al-Infitah (1377)*	*Al-Ismat (1265)*	*Al-Rakhawa (876)*
Class 1	*Al-Hams (571)*	*Al-Istilaa (411)*	*Al-Itbaq (235)*	*Al-Ithlaq (347)*	*Al-Shida (442)*
Class 2	-	-	-	-	*Al-Tawasot (294)*

Table 1. Number of approved recorded samples for each quranic letter

Letter	Number of records	Letter	Number of records	Letter	Number of records	Letter	Number of records
أ	50	د	58	ض	60	ك	58
ب	55	ذ	60	ط	57	ل	57
ت	50	ر	58	ظ	58	م	59
ث	57	ز	60	ع	60	ن	60
ج	55	س	60	غ	60	ه	54
ح	58	ش	59	ف	58	و	57
خ	57	ص	60	ق	59	ي	58

2.4. Features extraction and selection

Speech raw data is huge, and it contains much redundant and irrelevant information. Thus, the raw audio data must be processed into a new reduced form. Only the distinctive information should be kept in the new structure of data; this new form is called the acoustic features. In this research, two types of features were investigated; MFCCs, and perceptual linear prediction (PLP). They have been chosen as they are widely used in the field of speech processing, and they have shown a robust result [26]. In the traditional method in teaching the pronunciation of the Quranic letters, the *Sifaat* are usually evaluated by the hearing sensed by the teachers. This has led to the assumption that MFCC and PLP techniques, which may perform well in distinguishing the *Sifaat* with opposites, as both mimic the human perception of sounds and have been used for the modelling of the human auditory system.

MFCC is one of the most widely used acoustic features for speech recognition. MFCC is a time-frequency domain analysis, suitable for non-stationary signals. MFCC imitates the human perception sensitivity to the frequencies, where the frequencies below 1 kHz are linearly distributed and above 1 kHz are logarithmically distributed. It has shown robust modelling of speech signals, and it is a kind of human auditory system modelling. Figure 4 shows the block diagram of calculating the MFCC coefficients.

PLP considers the perception qualities of human auditory system same as MFCC. The basic idea of PLP analysis is to estimate the auditory spectrum of the speech using the all-pole model, which is therefore based on linear prediction analysis. Before the estimation, some modifications are made on the spectrum to match the human hearing perceptions. Compared to LPC, the PLP analysis yields better performance in the noisy environments. Figure 5 shows the process of estimating the PLP coefficients.

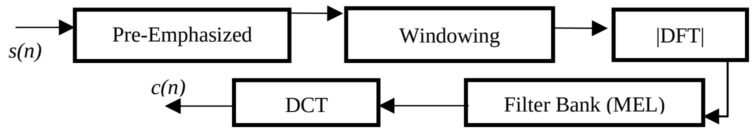

Figure 4. Block diagram of calculating MFCC coefficients

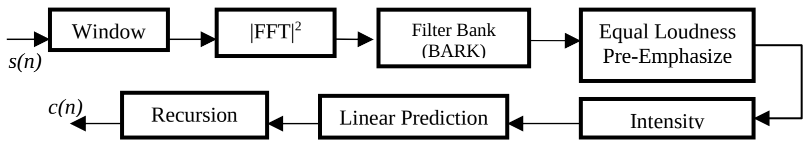

Figure 5. Block diagram of calculating PLP coefficients

In this paper, three feature vectors were investigated, MFCCs which has 14 coefficients, PLP which has 15 coefficients and a combination of both MFCC and PLP which has 29 coefficients. Each pair of *Sifaat* showed different performance with each feature vector. The process of feature selection is very important in building classification models. It is used to reduce the number of parameters in the feature vectors. It removes the redundant parameters and the less important parameters to the classification process. *Relieff* algorithm in MATLAB was used to evaluate the importance of each parameter in the feature vectors to the classification process. *Relieff* is a filter-based feature selection technique, that can be used with classification and regression problems, categorical and numerical data. The algorithm penalizes the predictors that give different values to neighbours of the same class and rewards predictors that give different values to neighbours of different classes [27]. *Relieff* is an extension version to overcome the time complexity of the *Relief* algorithm [28]. It has shown a robust performance for the noisy features. Moreover, it is fast and certainly practical for high-dimensional datasets. The process starts with selecting a random instance from the training data, then finding the *k* nearest neighbors *Hit* and *Miss*. The features of this random instance are being assessed one by one to the same features of the nearest neighbors *Hit* and *Miss*. Two scenarios are expected from this assessment process. First, a random instance feature and the *Hit* feature are different, which means this feature separates instances with the similar classes and this is not desirable, so it reduces the weight of this feature. Second, a random instance feature and the *Miss* feature are different, which means this feature separates instances with different classes that is desirable, so this feature is rewarded by increasing its weight. This whole process is repeated *m* times, where *m* is a user defined parameter.

2.5. Classification methods analysis

MATLAB classification learner App was used to investigate the best classification model to separate the Quranic letters to its *Sifaat* with opposites. It offers various features for building classification models, including explore data, select features, two validation schemes, train, and extract models, and various visualization techniques for model evaluations [29]. With the availability of dozens of machine learning algorithms, the selection of the right method is an overwhelming. In fact, no one can claim that a specific algorithm is the best solution for a specific problem, therefore, the optimized algorithm is partly based on trial-and-error approach. The algorithm needs to be tested with the data at the first stage to be judged. Moreover, a trade-off is required between the model speed, accuracy, and complexity in choosing the best model for a specific problem. MATLAB classification learner App helps in testing various types of models with the provided data and reports back the validated model performance. It has a very important feature which is the automated classifier training, where in this process, the App trains different classification models.

In the case of limited data samples, the need for validation scheme is high. In this research, the cross-validation was used during the training of the classification models. Cross-validation works efficiently when the small data problem is presented. The cross-validation scheme used in splitting the data. It was performed by keeping out a group of the data for testing, and the other groups are used for training the model. The process continues with many groups until all data are used for testing and training. The accuracy of the model is calculated as the average value of all trained models. By default, the classification learner App protects against over-fitting. For every validation set, the model is trained using the training part and assessed using the validation part of data. Finally, the reported results reflect the validated model.

2.6. Metrics and evaluations

Assessing the classification models is an important step to find the efficiency of that model to a particular problem to be solved. Confusion matrix was used to show the models performances. Using confusion matrix some other values can be extracted, such as the overall accuracy, recall, precision, and F1-score. Classification accuracy summarizes the total number of correctly classified samples to the total number of samples. It is commonly used to evaluate the classification models for its simplicity; however, it is not enough to be used alone, especially when the distribution of samples between the classes is severely skewed. The recall represents the correctly classified samples of a specific class to the total number of samples to that class, where it shows the model performance in identifying specific class. The precision represents the correctly classified samples of a specific class to the total number of predicted samples to that class. Recall and precision are number between 0.0 to 1.0, and the higher it is, the better is the performance of the classifier. While F1-score is the harmonic average of the precision and recall, and its value range between 0.0 for the worst performance to 1.0 for the best performance. In this paper, four values are used to evaluate model performance; overall accuracy, recall, precision, and F1-score, where recall, precision, and F1-score were calculated for the classes separately. The average of F1-score for all classes was calculated to show the performance of the classification model.

3. RESULTS AND DISCUSSION

This section presents the results for the designed five classification models. First, the feature selection results are presented, which illustrate the selected features set of parameters to be used as feature vectors to the classification models. Then the confusion matrices of the selected models are presented to calculate the overall accuracy, recall, precision, and F1-score. The experiments were conducted on various feature vectors; MFCC, PLP and combination of MFCC and PLP. Various classification methods were tested as well, but only the outperformed models are reported in this paper. Figure 6 shows the result of *Relieff* analysis. MFCCs represent the MFCC and PLPs represents the PLP coefficients. Different set of features has been used for each classification model. These features are the highly ranked features according to the *Relieff* analysis.

		Feature used											
		1	2	3	4	5	6	7	8	9	10	11	12
Classifier #	Pair 1	MFCC1	MFCC2	MFCC3	MFCC4	MFCC5	MFCC6	MFCC7	MFCC8	MFCC9	MFCC11	MFCC12	
	Pair 2	MFCC2	MFCC3	MFCC4	MFCC5	MFCC6	PLP1	PLP2	PLP3	PLP4	PLP5	PLP6	PLP7
	Pair 3	PLP1	PLP2	PLP3	PLP4	PLP5	PLP6	PLP7	PLP13	PLP15			
	Pair 4	MFCC1	MFCC2	MFCC3	MFCC4	MFCC5	MFCC6	MFCC7	MFCC8	MFCC9	MFCC11	MFCC12	MFCC14
	Group 5	MFCC1	MFCC2	MFCC3	MFCC4	MFCC5	MFCC6	MFCC7	MFCC8	MFCC9	MFCC11	MFCC12	MFCC14

Figure 6. Feature set used for each classifier based on the relieff analysis

3.1. *Sifaat* pair 1 (*Al-Jahr* and *Al-Hams*)

For classifying the first pair of *Sifaat* with opposites *Al-Jahr* and *Al-Hams*, the best classification model is the k-nearest neighbour (KNN). Figure 7 illustrates the confusion matrix of the selected model. The model's overall accuracy was 94.8%. The recall of class 0 is 0.96 and for class 1 is 0.93. The precision of class 0 is 0.96 and for class 1 is 0.93. F1-score is 0.96 and 0.93 for class 0 and class 1 respectively. The average F1-score of the classifier is 0.95, which shows an excellent performance.

		PREDICTED	
		0	1
TRUE	0	998	43
	1	41	530

Figure 7. Confusion matrix of the selected model for pair 1

3.2. *Sifaat* pair 2 (*Al-Istilaa* and *Al-Istifal*)

For classifying the second pair of *Sifaat* with opposites (*Al-Istilaa* and *Al-Istifal*) the weighted KNN classification model was outperformed. The model's overall accuracy was 92%as shown in Figure 8. Class 0, and class 1 represent *Al-Istifal*, and *Al-Istilaa* letters respectively. For recall and precision, the values were 0.95 and 0.94 were for class 0 respectively, which means that the model ability in classifying *Al-Istifal* letters is excellent. Class 1 shows degraded values compared to class 0, but it is still considered good values, where recall is 0.82 and precision is 0.84, these results prove that the model is good in classifying *Al-Istilaa* letters. The average F1-score of the classifier is 0.89, which shows a good performance.

		PREDICTED	
		0	**1**
TRUE	**0**	**1138**	**63**
	1	**73**	**338**

Figure 8. Confusion matrix of the selected model for pair 2

3.3. *Sifaat* pair 3 (*Al-Itbaq* and *Al-Infitah*)

For classifying the third pair of *Sifaat* with opposites (*Al-Itbaq* and *Al-Infitah*) the RUSBoosted showed the best performance. The RUSBoosted a hybrid approach that includes sampling and boosting techniques to solve the problem of imbalanced data distribution and improve the classification performance. The model's overall accuracy was only 81% as shown in Figure 9. The recall and precision are 0.81 and 0.96 for class 0, respectively. Class 1 shows degraded values compared to class 0, where the recall was 0.81, but precision was only 0.42. The average F1-score of the classifier was 0.72, which shows the best-achieved performance among various feature vectors and classification methods. The degraded performance of this classifier is due to the lower number of samples in class 1. The model ability in identifying *Al-Infitah* is excellent, while it needs to be improved for *Al-Itbaq* letters. This is due to the lower number of samples for *Al-Itbaq* letters in the database. The performance can be enhanced in future work by increasing the number of samples in class 1.

		PREDICTED	
		0	**1**
TRUE	**0**	**1116**	**261**
	1	**44**	**191**

Figure 9. Confusion matrix of the selected model for pair 3

3.4. *Sifaat* pair 4 (*Al-Ithlaq* and *Al-Ismat*)

For the fourth pair of *Sifaat* with opposites (*Al-Ithlaq* and *Al-Ismat*), the ensemble random under-sampling boosting (RUSBoosted) classification model outperformed. The model's overall accuracy was 83% as shown in Figure 10. The recall and precision are 0.85 and 0.92 were for class 0, respectively. Class 1 shows degraded values compared to class 0, where the recall was 0.73, but the precision was only 0.58. The average F1-score of the classifier was 0.77, which shows the best-achieved performance among various feature vectors and classification methods. The degraded performance of this classifier is due to the low number of samples in class 1. The probability of the system in identifying *Al-Ismat* letters is excellent, while it is fair for *Al-Ithlaq* letters. The model performance can be improved by increasing the number of samples in *Al-Ithlaq* group.

		PREDICTED	
		0	**1**
TRUE	**0**	**1078**	**187**
	1	**93**	**254**

Figure 10. Confusion matrix of the selected model for pair 4

3.5. Sifaat group 5 (*Al-Rakhawa, Al-Shida* and *Al-Tawasot*)

Group 5 of the *Sifaat* with opposites includes 3 *Sifaat*, *Al-Rakhawa*, *Al-Shida and Al-Tawasot*. Therefore, this is solved as a multi-class classifier. The fifth group of *Sifaat* with opposites, the medium gaussian SVM was used as the classification model. The model's overall accuracy was 86.5% as shown in the Figure 11. The recall and precision are 0.90 and 0.89 for class 0 respectively. Class 1 recall was 0.82, and precision was only 0.79. The Class 2 values were 0.82 and 0.86 for recall and precision, respectively. The average F1-score of the classifier was good where the value was about 0.84, which shows the best-achieved performance among various feature vectors and classification methods.

		PREDICTED		
		0	**1**	**2**
TRUE	**0**	**788**	**62**	**26**
	1	**62**	**366**	**14**
	2	**36**	**18**	**240**

Figure 11. Confusion matrix of the selected model for group 5

4. CONCLUSION

In this paper, a classification-based approach was built to assess the Quranic letters' pronunciations according to its *Sifaat* with opposites groups. Five classifiers were designed for the five groups of *Sifaat*. Three feature vectors were investigated to find the best acoustic features to represent each group of *Sifaat*. The investigated acoustic features were MFCC, PLP, and a combination of both MFCC and PLP. Feature selection was used to reduce the size of the feature vectors, and remove the irrelevant parameters. Three types of classification algorithms have outperformed in identifying the five pairs of *Sifaat*. For Pair 1 and Pair 2, the weighted KNN algorithm, for Pair 3 and Pair 4, the RUSBoost, and for Group 5, the medium gaussian SVM. The designed classification models were used as a part of an automatic real-time evaluation system for Quranic letters' pronunciation according to its *Makhraj* and *Sifaat*.

ACKNOWLEDGMENT

This project is currently sponsored by Majlis Agama Islam dan Adat Istiadat Melayu Perlis (MAIPs) under Jamalullail Grant (JRGS21-009-0009) with the title "Intelligent Modelling System of Characteristics (*Sifaat*) Without Opposite for Tajweed Teaching and Learning", under the management of Research Management Centre, International Islamic University Malaysia (IIUM)

REFERENCES

[1] Z. Ali, M. Alsulaiman, G. Muhammad, I. Elamvazuthi and T. A. Mesallam, "Vocal fold disorder detection based on continuous speech by using MFCC and GMM," *2013 7th IEEE GCC Conference and Exhibition (GCC)*, 2013, pp. 292-297, doi: 10.1109/IEEEGCC.2013.6705792.

[2] A. Maier *et al*., "Automatic detection of articulation disorders in children with cleft lip and palate," *The Journal of the Acoustical Society of America*, vol. 126, no. 5, Nov. 2009, doi: 10.1121/1.3216913.

[3] J. Lee, C. H. Lee, D. Kim and B. Kang, "Smartphone-Assisted Pronunciation Learning Technique for Ambient Intelligence," in *IEEE Access*, vol. 5, pp. 312-325, 2017, doi: 10.1109/ACCESS.2016.2641474.

[4] A. Hanani, M. J. Russell, and M. J. Carey, "Human and computer recognition of regional accents and ethnic groups from British English speech," *Computer Speech and Language*, vol. 27, no. 1, pp. 59–74, Jan. 2013, doi: 10.1016/j.csl.2012.01.003.

[5] M. Maqsood, A. Habib, and T. Nawaz, "An Efficient Mispronunciation Detection System Using Discriminative Acoustic Phonetic Features for Arabic Consonants," *International Arab Journal of Information Technology*, vol. 16, no. 2, pp. 242–250, Mar. 2019.

[6] S. M. Witt, "Automatic Error Detection in Pronunciation Training: Where we are and where we need to go," *International Symposium on automatic detection on errors in pronunciation training*, vol. 1, 2012.

[7] M. Alsulaiman *et al*., "Pronunciation errors of non-Arab learners of Arabic language," *2014 International Conference on Computer, Communications, and Control Technology (I4CT)*, 2014, pp. 277-282, doi: 10.1109/I4CT.2014.6914189.

[8] S. M. Abdou and M. Rashwan, "A Computer Aided Pronunciation Learning system for teaching the holy quran Recitation rules," *2014 IEEE/ACS 11th International Conference on Computer Systems and Applications (AICCSA)*, 2014, pp. 543-550, doi: 10.1109/AICCSA.2014.7073246.

[9] S. Akhtar *et al*., "Improving Mispronunciation Detection of Arabic Words for Non-Native Learners Using Deep Convolutional Neural Network Features," *Electronics*, vol. 9, no. 6, 2020, doi: 10.3390/electronics9060963.

[10] F. Nazir, M. N. Majeed, M. A. Ghazanfar and M. Maqsood, "Mispronunciation Detection Using Deep Convolutional Neural Network Features and Transfer Learning-Based Model for Arabic Phonemes," in *IEEE Access*, vol. 7, pp. 52589-52608, 2019, doi: 10.1109/ACCESS.2019.2912648.

[11] M. Maqsood, H. A. Habib, T. Nawaz, and K. Z. Haider, "A Complete Mispronunciation Detection System for Arabic Phonemes using SVM," *Computer Science*, 2016.

[12] S. Khairuddin, “Formant Frequency Analysis of Quranic Alphabets in Sukoon Pronunciation Between Adult Male and Female Experts,” *QURANICA - International Journal of Quranic Research*, vol. 11, no. 1, pp. 37–52, Jun. 2019.
[13] H. M. A. Tabbaa and B. Soudan, “Computer-Aided Training for Quranic Recitation,” *Procedia - Social and Behavioral Sciences*, vol. 192, pp. 778–787, Jun. 2015, doi: 10.1016/j.sbspro.2015.06.092.
[14] B. Alkhatib, M. Kawas, A. Alnahhas, R. Bondok, and R. Kannous, “Building an assistant mobile application for teaching arabic pronunciation using a new approach for arabic speech recognition,” *Journal of Theoretical and Applied Information Technology*, vol. 95, no. 3, Feb. 2017.
[15] A. Hanani, M. Attari, A. Farakhna, A. Joma’A, M. Hussein, and S. Taylor, “Automatic Identification of Articulation Disorders for Arabic Children Speakers,” in *Proc. 5th Workshop on Child Computer Interaction (WOCCI 2016)*, pp. 35-39, 2016. doi: 10.21437/WOCCI.2016-6.
[16] A. Al Hindi, M. Alsulaiman, G. Muhammad and S. Al-Kahtani, "Automatic pronunciation error detection of nonnative Arabic Speech," *2014 IEEE/ACS 11th International Conference on Computer Systems and Applications (AICCSA)*, 2014, pp. 190-197, doi: 10.1109/AICCSA.2014.7073198.
[17] K. Necibi and H. Bahi, “An Arabic Mispronunciation Detection System by means of Automatic Speech Recognition Technology,” in *Proceedings de 13th International Arab Conference on Information Technology ACIT’2012*, 2012.
[18] A. F. Ahmad Khan, O. Mourad, A. M. K. Bin Mannan, H. B. A. M. Dahan and M. A. M. Abushariah, "Automatic Arabic pronunciation scoring for computer aided language learning," *2013 1st International Conference on Communications, Signal Processing, and their Applications (ICCSPA)*, 2013, pp. 1-6, doi: 10.1109/ICCSPA.2013.6487246.
[19] B. S-Pathak and A. Panat, “Extraction of Pitch and Formants and its Analysis to identify 3 different emotional states of a person,” *International Journal of Computer Science Issues*, vol. 9, no. 4, July 2012, doi: 10.1.1.402.3342.
[20] T. Altalmas, S. Ahmad, W. Sediono, and S. S. Hassan, “Quranic Letter Pronunciation Analysis based on Spectrogram Technique: Case Study on Qalqalah Letters,” *The 11th International Conference on Artificial Intelligence Applications and Innovations (AIAI'15)*, 2015.
[21] S. Khairuddin *et al.*, “Classification of the Correct Quranic Letters Pronunciation of Male and Female Reciters,” *IOP Conference Series: Materials Science and Engineering*, vol. 260, p. 012004, 2017, doi: 10.1088/1757-899x/260/1/012004.
[22] H. Azam *et al.*, "Classifications of clinical depression detection using acoustic measures in Malay speakers," *2016 IEEE EMBS Conference on Biomedical Engineering and Sciences (IECBES)*, 2016, pp. 606-610, doi: 10.1109/IECBES.2016.7843521.
[23] N. N. W. N. Hashim, D. M. Wilkes, and R. Salomon, “Timing Patterns of Speech as Potential Indicators of Near-Term Suicidal Risk,” *International Journal of Multidisciplinary and Current Research*, vol. 3, pp. 1104-1116, Nov. 2015.
[24] Kareema Czerepinski, “Tajweed Rules of the Qur’an Part 2,” *King Fahd National Cataloging in Publication Data*, Jeddah: Dar Al-Khair Islamic Books Publisher, ISBN: 9660-9343-4-9, 2003.
[25] Md. Z. Alom *et al.*, “A State-of-the-Art Survey on Deep Learning Theory and Architectures,” *Electronics*, vol. 8, no. 3, p. 292, Mar. 2019, doi: 10.3390/electronics8030292.
[26] M. N. M. Arafa, R. Elbarougy, A. A. Ewees, and G. M. Behery, “A Dataset for Speech Recognition to Support Arabic Phoneme Pronunciation,” *International Journal of Image, Graphics and Signal Processing*, vol. 10, no. 4, pp. 31-38, Apr. 2018, doi: 10.5815/ijigsp.2018.04.04.
[27] “Introduction to Feature Selection-MATLAB & Simulink.” [Online], Available: https://www.mathworks.com/help/stats/feature-selection.html (accessed Dec. 26, 2020).
[28] J. H. Moore and B. C. White, “Tuning ReliefF for Genome-Wide Genetic Analysis,” in *Evolutionary Computation, Machine Learning and Data Mining in Bioinformatics*, pp. 166–175, 2007, doi: 10.1007/978-3-540-71783-6_16.
[29] “Train Classification Models in Classification Learner App-MATLAB & Simulink,” [Online], Available: https://www.mathworks.com/help/stats/train-classification-models-in-classification-learner-app.html (accessed Dec. 26, 2020).

BIOGRAPHIES OF AUTHORS

Tareq AlTalmas received the B.Sc. degree in Islamic University of GAZA, (Electrical Engineering), M.Sc and Ph.D. degree from International Islamic University-Malaysia, (Mechatronics Engineering. He was appointed as a research assistant in between 2012 to 2017 in two different projects; designing a two-wheeled wheelchair for disabled, and development a real-time system for evaluating the Quranic letters pronunciation. He used to teach different courses at the international Islamic university–Malaysia. He can be contacted at email: tareq.ma@live.iium.edu.my.

Salmiah Ahmad Associate Professor Ir. Dr. Salmiah Ahmad, started her career as a Project Engineer at Toyo Engineering and Construction (M) Sdn. Bhd, after received her B.Eng in Mechatronics Engineering from International Islamic University Malaysia in 2001. Later she received her M. EngSc in Electrical Engineering from Curtin University of Technology, Australia in 2004. She was then completed her Ph.D in Automatic Control and Systems Engineering from The University of Sheffield in 2010, and currently she is an Associate Professor at the Department of Mechanical Engineering, International Islamic University Malaysia, Kuala Lumpur. Her current research interests include machine learning for signal processing, electric vehicle, intelligent systems and control, mobile robot, rehabilitation and bio-signals, instrumentation and signal processing. She can be contacted at email: salmiah@iium.edu.my.

Nik Nur Wahidah Nik Hashim received her BS, MS, and PhD in Electrical Engineering at Vanderbilt University, Nashville, Tennessee, USA beginning 2005 to 2014. Currently an assistant professor in the Department of Mechatronics, Kuliyyah of Engineering at the International Islamic University Malaysia (IIUM). Her passion in research is studying the field of mental health through speech analysis. Other specific field of research interest include speech processing, machine learning, deep learning, digital signal processing and natural language processing. She can be contacted at email: niknurwahidah@iium.edu.my.

Ustaz Dr. Surul Shahbudin Hassan al-Hafiz is an expert in (*Tajweed*), Tafsir, Hadith and Fiqh Syafie. He obtained his first degree (B.A) in Islamic Science and Arabic Language and Literature from Dar Al Ulum, Nadwatul 'Ulama, India in 1996 after he obtained his Tahfiz from Ma'had Tahfiz Nadwatul Ulama, Lucknow. He later pursued his master's degree (M.A) in Fiqh and Usul Fiqh in 2000 from the International Islamic University Malaysia. He completed his Ph.D in Islamic Education in 2015 from University Malaya and currently he is a permanent senior lecturer at Tilawah Division, Centre For Languages And Pre-University Academic Development (CELPAD), International Islamic University Malaysia. He can be contacted at email: surur@iium.edu.my.

Wahju Sediono completed his Diplom-Ingenieur from RWTH Aachen University, Germany in Electrical Engineering and Information Technology in 1997. He received his Ph.D. in Electrical Engineering from Unversität Karlsruhe (TH), Karlsruhe, Germany in 2003. He joined the Department of Mechatronics Engineering, International Islamic University of Malaysia (IIUM) in 2011. His current research interests include signal and image processing, machine vision, radar and navigation, intelligent systems, and biomedical instrumentation. He is a senior member of the IEEE. He can be contacted at email: wsediono@iium.edu.my.