Information Sciences Letters
An International Journal

http://dx.doi.org/10.18576/isl/110410

A Machine Learning Model for the Identification of the Holy Quran Reciter Utilizing K-Nearest Neighbor and Artificial Neural Networks

Meshal Mohammed Al Anazi[1,*] *and Osama R. Shahin*[2]

[1]Department of Islamic Studies, College of Science and Arts in Qurayyat, Jouf University, Qurayyat, Saudi Arabia
[2]Department of Computer Science, College of Science and Arts in Qurayyat, Jouf University, Qurayyat, Saudi Arabia

Received: 21 Feb. 2022, Revised: 22 Mar. 2022, Accepted: 3 Apr. 2022.
Published online:1 Jul. 2022.

Abstract: The method of identification of the Holy Quran reciter, which is entered on the various features of the acoustic wave, is referred to as the Holy Quran Reciter Identification. The Muslim community's Holy Book is the Holy Quran. Listening to or reading the Holy Quran is one of the obligatory activities for Muslims. This research proposes a machine learning model for identifying the Holy Quran reciter using a machine learning language. Here, the presented system comprises the essential phases for a voice recognition system encompassing the processes of classification, extraction of features, preprocessing, and data acquisition. Moreover, the voices of ten known reciters are framed as a dataset in this research. The reciters are leaders of prayers in the Holy masjids of Madinah and Makkah. The analysis of the audio dataset is performed using the mel frequency cepstral coefficients (MFCC). The artificial neural network (ANN) and the k-nearest neighbor (KNN) classifiers are employed for classification. The pitch is utilized as features employed to train the KNN and ANN classifiers. The proposed system is validated using two chapters chosen from the Holy Quran. The results revealed an excellent level of accuracy. With the help of the ANN classifier, the proposed system offered 98.5% accuracy for chapter 7 and 97.2% accuracy for chapter 32. On the other hand, while utilizing KNN, the accuracy for chapter 7 is 97.02% and for chapter 32 is 96.07%. Then, the system's performance is compared with the utilization of support vector machines (SVM) in recognition of Quranic voice reciter. The comparison results revealed that ANN is a better machine learning algorithm for voice recognition when compared to SVM.

Keywords: Machine learning, ANN, KNN, Holy Quran.

1 Introduction

For the Muslim community, the Holy Quran is their Holy Book. There are 114 chapters and 30 parts in the Holy Quran. Those chapters are called surah. Every chapter of the Holy Quran is comprised of a number of verses. Some surahs have a large number of verses and are lengthy, like Surah Al-Baqarah and Surah Ali 'Imran, whereas others have a smaller number of verses, such as Surah Al-Ikhlas and Surah An-Nas. Some special mechanisms and rules have to be employed in reading the Holy Quran [1]. This mechanism of reading the Holy Quran is called Tajweed. The recognition of the reciters of the Holy Quran is not well investigated, unlike the recognition of the speaker and speech of the English language. There is a lag in the recognition of Quran reciters compared with the recognition of speech mechanism, which requires a differentiating biometric signal for every reciter, indicating the lagging of uniformity of the datasets for Quran reciters. In the proposed system, the concept of speech recognition is applied for recognizing the reciters of the Holy Quran.

Eminently, recognizing the reciters of the Holy Quran examination lacks common datasets for Quran reciters [2]. Speech recognition encapsulates many fields including but not limited to dialing or voice calling, forensic, tests/exams and speech data indexing. Many of the speech recognition researchers carried out their examination on the datasets of the English language. However, only a few examinations are made in the language of Arabic. There is a limitation on the Arabic speech recognition systems compared with speech recognition systems of other languages.

The recognition of the Quran reciter is a complex task as it is done with the way of Tajweed. Every Quranic reciter has a differentiating signal [3]. This reciter signal encapsulates a dynamic change, which is temporary and altered over time based on the pronunciation basis and the recitation

*Corresponding author-mail: mmhanazi@ju.edu.sa

way called Tarteel. Classifying and recognizing the reciters of the Holy Quran are regarded as parts of the field of speech and voice recognition systems. This article concentrates on the recognition of the reciters of the Holy Quran.

2 Theoretical Background of the Study

This section deals with the Holy Quran reciter types. There are multiple regional accents in different Arab communities, and these accents pose a challenge for the Arab speech recognition applications under consideration.

2.1 The Arabic Language

Arabic is one among the most prevailing languages around the world due to its diverse speakers, and Allah revealed Al-Quran, the Holy Book, in this language to his messenger Prophet Muhammad ﷺ. Moreover, people who are non-Arabic speakers tend to read the Holy Quran correctly with the rules of Tajweed. Their ultimate objective is understanding and reading the Holy Quran properly. The reading of the Holy Quran differs greatly from other speaker or voice identification systems. The difference is the availability of specific acoustic rules known as "Ahkam Al-Tajweed." These rules must be strictly applied during the reading of the Holy Quran. In addition, the reader might add additional emotional features. The recitation or reading type is differentiated with the transition from one acoustic level to another. This type of transition is called "Maqam" [4]. There are seven eminent reading styles that should be applied while reading the Holy Quran. Those ways of reading the Holy Quran are referred to as "Qirā'āt." Prophet Muhammadﷺ mentioned this statement in Hadith 5041. In addition, three new reading styles are embedded with those reading styles. Thus, ten reading styles, which are to be read by the prayer leaders or imams, are mentioned in the reading science of the Holy Quran. "Nafi'" and "Aasim" are popularly used reading styles. Three variations differentiate the reading styles, which are as follows:

- Shortening and lengthening of words, known as "qasr" and "madd," respectively
- Addition of punctuation in the Holy Quran's written text
- The reading and pronunciation of the various sections of the Quran, based on the different recitation styles

Various researchers have defined the complexities that people face when approaching the Arabic language, as well as the intricacies of the Holy Quran. One of the most important challenges is facilitating the Holy Quran recitation to non-Arabic speakers. The following are the complexities faced by beginners in the recitation of the Holy Quran:

- Lack of monitoring and oral practice for correcting errors
- Unfamiliarity faced in learning Arabic by various learners with the revelation of text in the language of Arabic
- Lack of follow-up of the reciters for enhancing their recitation style by most software programs

Thus, a suitable Tajweed teacher must be there to authorize the recitation of the Holy Quran. Thus, Prophet Muhammad ﷺrecited the Holy Quran. Thus, the accuracy of recitation is checked with some Arabic speech recognition (ASR) systems. These Tajweed pronunciation styles are practiced by various reciters, and differentiating them is a complex task. Thus, research is required to solve this imminent issue [4].

Various approaches are available in the earlier literature that make use of the features of MFCC (mel frequency cepstrum coefficients) and certain learning algorithms, such as artificial neural network (ANN), k-nearest neighbor (KNN), support vector machines (SVM), and hidden Markov model (HMM). This is made as an effort to recognize various voices of the readers or reciters of the Holy Quran. This research is essential to overcome the drawbacks of the previous recognition systems utilized in recognizing the Holy Quran reciters.

2.2 The Tajweed Art

The Arabic word "Tajweed" means proper recitation with pronunciation at a moderate speed. It refers to a group of rules that governs the recitation of the Holy Quran. Therefore, it is regarded as an art due to the performance of the reciters in a similar form. The Tajweed art refers to the rules that are flexible and well-defined for reciting the Holy Quran. There is a significant difference between the recitation of the Holy Quran verses and the normal speech in Arabic. This creates an interesting outcome, based on the effect of analyzing the art of the automatic recognition process, which most suitably focuses on the acoustic technique. In addition, the Tajweed art indicates the manual technique, which requires a lot of work to make it suitable to the new reciters of the Holy Quran. Still, people believe that for reciting the Holy Quran, the Tajweed art must be taken into account [3].

2.3 Impact of Tajweed Art on the Acoustic Method

The voice of every person is unique. As a result, the voices of the Quran recited by many reciters differed significantly from one another. When a sentence is fetched from the same verse of the Holy Quran, the delivering or recitation way of the Holy Quran is not the same. Various sounds or voices are created by various reciters. Many challenges are encountered while reciting the Holy Quran specialties due to the variations between the Holy Quran recitation and the written language. These similar letter combinations are pronounced in different ways because of the utilization of harakattes. The Tajweed rules, which are believed to affect the recognition aspect of recitation, are as follows:

- The letter "R" is pronounced in an empathetic way.
- Silent pronunciation of the unannounced letters.
- Two vowels are pronounced with nasalization.
- Two, four, and six vowels are pronounced with permissible prolongation.
- Five vowels are pronounced with obligatory prolongation.
- Six vowels are pronounced with necessary prolongation.

These laws are based on specific rules of recitation. The "maqams" are predefined and recited by the reciters for varying their recitation on tones. Ten various sets of laws are certified by certain scholars for the proper recitation of the Holy Quran. For this recitation, the pronunciation of vowels to the n times must be considered and nasalization is essential. Nas.
Alization is concerned with the vowels and consonant combinations, utilization of long and short vowels, pharyngeal and emphatics coarticulation effect, Ghonna and Tanween rules, and combining word rules. When an echo sound is generated during the recitation and recording of the Holy Quran, this echo is regarded as noise. This noise must be eliminated with the help of any filter, which would cancel the noise and the language modeling problem in Arabic. Only the Modern Standard Arabic (MSA) is utilized for formal and written communication. This Arabic is only used as the writing reference standard. The Arabic alphabets of the Holy Quran are comprised of 28 letters, which are called Hijaiyah letters. Three of the letters are vowels and 25 are consonants. Four or two various shapes are encompassed. No capitalization is involved, and most of the letters are connected.

3 Related Works

A machine learning model is utilized to recognize the Holy Quran reciter. The authors employed a database of 12 prayer leaders or Qari who recited the final ten surahs of the Holy Quran [5]. Those 12 people were identified as the class problem. For the audio representation, they utilized two approaches: analysis of the audio in the frequency domain and representation of the audio in the form of spectrogram images. Pitch and MFCC are employed as model features for learning in the initial case. In the next case, audio is represented in the form of images; the features are extracted with the help of auto correlograms. In both cases, conventional machine learning methods were used to learn those features, including Random Forest (RF), J48, and Naïve Bayes. Those classifiers were chosen because of their best performance. It was found that those classifiers could efficiently study the variation between the classes, MFCC audio representation, and the features of the pitch. RF and Naïve Bayes obtained an 88% accuracy of recognition, indicating that the Qaris could be recognized effectively with the recitation of Quranic verses.
In [6], deep learning was utilized to determine automatically the correct execution of the basic recitation rules of the Holy Quran. Their research addressed the complexity of determining the correct use of Ahkam Al-Tajweed of the whole Quran. Primarily, the research focused on eight principles of Ahkam Al-Tajweed, which were overcome by the early reciters of the Holy Quran. In the initial part of this research, traditional techniques of audio processing were employed to extract features, such as RF and SVM, as well as classifiers like KNN, Markov model's spectral peak location, wavelet packet decomposition (WPD), MFCC, and linear predictive code (LPC). Those features were employed in a dataset encompassing many audio recordings that includes all the rules of the whole Holy Quran by various reciters of both genders. Deep learning techniques have been used to enhance the classification accuracy to 97.7%. The extraction of the traditional features was made with the convolution deep belief network (CDBN) and SVM was used for the classification.
In [7], a computer-aided training for the recitation of the Holy Quran was developed. The development of a Computer-Aided Quranic Recitation Training System to detect errors associated with the classifier-based approach and automatic speech recognition for detecting the errors in recitation. The detection of errors was performed in two stages: in HMM-ASR, Quranic recitation is recognized, insertion, substitution, and deletion of phones are detected, and phonetic alignments of time are offered; a classifier-based approach was utilized for differentiating between the confusing phones to achieve enhanced utterances of the detection of the letter "R" in Arabic and the differentiation between confusing letters and closely related pronunciations. Their study results provided a 91.2% accuracy at the word level.
In [4], a Holy Quran reciter/reader identification system using SVM was developed. Their research builds a corpus, which encompasses 15 Holy Quran reciters. The features were extracted from the acoustic input signal with the help of MFCC. The matrix containing the features of a reader is the MFCC utilized for reader recognition with the help of ANN and SVM. The experimental results indicated that 86.1% accuracy was obtained utilizing ANN and 96.59% accuracy was obtained with SVM. Their model was a success with an accuracy of 96%.
In [8], a correction mapping model for the Holy Quran Reading Performance Engine of the evaluation was created. The primary goal of this research is to create that model. Moreover, digital signal processing and machine learning techniques were employed to analyze and represent the speech recitation signal. A recitation form of corrected recitation results was obtained for the performance evaluation in the final stage. In this article, the proposed corrected mapping model is not limited to the classification of the threshold process, estimation of parameters, speaker adaptation, and representation of the recitation and speaker recitation variability. The findings of this experimental study indicated that the automated corrected system of the Holy Quran, known as "Intelligent Quran Recitation

Assistant," is capable of fulfilling the future and current trends of the digital world.

In [9], Tajweed checking rules of automation for learning the Holy Quran are developed. This article offered a structural outline of a speech recognition system for creating a model for the recitation of the Quranic verse recognition and checking Tajweed rules. The drawback of their method was that it was found unattractive and less effective for the younger Muslim people. This research proposed software prototype development to construct a system that automatically checks the Tajweed rules during the Quranic recitation. The results indicated that the prototype was a better tool for enhancing the Quranic reading by the Muslim students, where the students could read the Holy Quran with the Tajweed rules, even in the absence of their teachers.

In [10], a technique of speech identification utilizing MFCC-based SVM was proposed. It was a speech recognition technique dependent on text. In this work, the statistical properties of MFCC and SVM were employed as the input characteristics of the neural network. Initially, this work utilized the sequential minimum optimization (SMO) technique of learning for SVM, which significantly enhances the performance compared with traditional techniques. The cepstrum coefficients that indicated the speaker features in a segment of speech were computed with a nonlinear filter bank analysis and the assistance of discrete cosine transform (DCT). The convergence speed and ability of speaker identification by the SVMs were examined for various combinations of characteristics. The results of the experiment, which was conducted on various samples, shed light on the effect of the proposed system.

In [11], a study called smart Tajweed Automatic Recognition of the Arabic Recitation Rules of the Holy Quran was conducted. This paper significantly thoughtfully examined the Tajweed pronunciation and the significant rules of recitation, which offer various rhythms and melodies for the correct pronunciation in each Tajweed rule. Significance was also given to the Holy Quran recitation as multiple meanings for the Arabic words were given. This research study utilized a threshold scoring system and SVM to automatically recognize the Tajweed rules of recitation of the Holy Quran.

4 Methodologies

According to the existing studies, there are four significant stages of speech recognition, which are also included in the present system. The Arabic recitation system of the Holy Quran consists of the following five stages: preprocessing, data acquisition, extraction of features, testing and training, patterns recognition, and features' classification.

4.1 Preprocessing

The significant advantage of the preprocessing stage of the speech recognition system is the right organization of information, the simplification of data, which is the input speech signal, and then the simplification of voice recognition. In the preprocessing stage, the functions performed are silence removal and preemphasis, which is performed on signals that lead to high frequencies in association with the low-frequency magnitude, thereby enhancing the signal-to-noise ratio.

4.2 Data Acquisition

The Holy Quran is the Holy Book for the entire Muslim community. It is a guide for every Muslim, which they follow to have a good life path. It is a verbal collection of revelations to Prophet Muhammad ﷺ over a period of 23 years. It encapsulates 114 chapters or surahs that have verses. The chapters of the Holy Quran, chosen for this research, are chapters 7 and 32, which are Surah Al-A'raf and Surah As-Sajdah, respectively. The Holy Quran has 6348 verses. To date, no corpus exists for the surahs in the Holy Quran, for the performance of the reciters or the readers. Here, the corpus is constructed based on the corpus or verses of each surah. Ten readers were chosen to construct the corpus in the present research.

4.3 Feature Extraction

Extraction of features is the initial step in developing any system of speech recognition. The goal for the features is to identify the important components that highlight the speech signal and eliminate all the redundant data within the speech signal. Knowledge about the speech or audio signal assists in the extraction of robust features [12].

4.4 Training and Testing

The ten feature matrices are combined as one file for testing the features of the ten readers. For the sake of testing and training, a WEKA tool was utilized with 70% training data and 30% testing data, as shown in Figure 1.

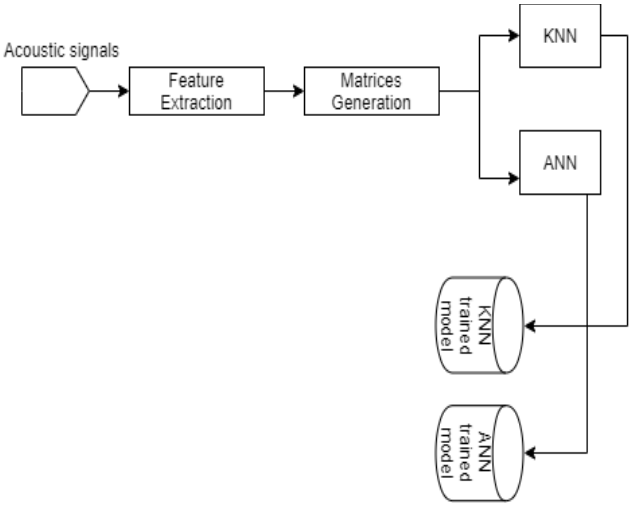

Fig. 1: The 70% training phase.

The features are rearranged with a filter property for the

randomly selected rows. Then, 70% of the data are trained with ANN and KNN in the system. Two trained models were obtained as a result: one for ANN and one for KNN. The remaining feature matrices are utilized for testing, as shown in Figure 2.

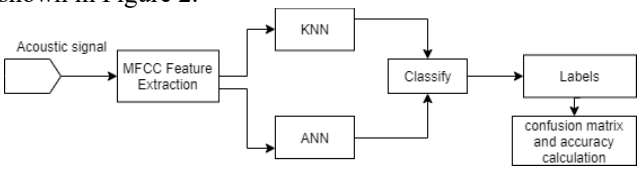

Fig. 2: The 30% testing phase.

Two confusion matrices are obtained for KNN and ANN. The miss and hit classification of the acoustic waves is represented by the confusion matrix. The main diagonal indicates the right classification.

5 The Proposed System

The significant part of the system proposed here is the extraction step of the features. The mapping of these features was made into the KNN or ANN classifiers for recognizing the reciters of the Holy Quran. Figure 3 shows a summary of the model of the proposed system.

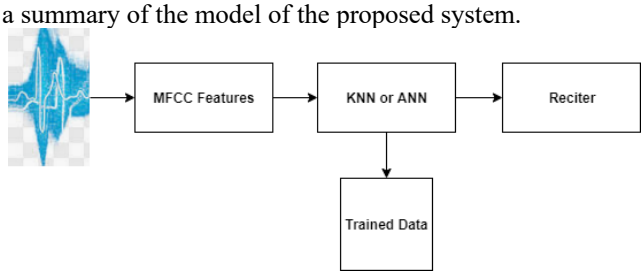

Fig. 3: Block diagram of the proposed system.

The MFCC method is employed for extraction of features. There, those features are mapped into the KNN or ANN classifier for testing and training to identify the reciters of the Holy Quran. Initially, building the training model is performed after the features' extraction. Then, certain samples of about ten reciters were gathered to create the dataset. The prayer leaders of the Holy Mosques of Makkah and Madinah are the reciters in the present system. Listening to or reading the Holy Quran is one of the obligatory activities for the Muslims. For example, it is recommended that every Friday, chapter 18 is recited. Two chapters have been selected to be recited by the ten readers. The chapters of the Holy Quran, which were chosen for the present research, are chapter 7, Surah Al-A'raf, and chapter 32, Surah As-Sajdah. Surah Al-A'raf depicts the events associated with the Holy Prophets (may the peace and blessings of Allah be upon them) and the nations' fate, which encourages good deeds and faith for the current people. It is narrated by the Holy Prophet (ﷺ) that chapter 32 (Surah As-Sajdah) provides shade for its reciters on the Day of Resurrection. The reciters' recitations based on these two chapters was employed for this research. These chapters are commonly recited by Muslims.

In this research paper, the corpus is developed for two chapters and ten reciters. Importantly, each verse is created with a wave file and the chapters are recited by every reciter and are stored as wave files. As a result, 2060 wave files were created for Surah Al-A'raf because it has 206 verses. Similarly, 300 wave files were created for Surah As-Sajdah as it has 30 verses. There are 114 chapters in the Holy Quran, with n number of verses in each chapter. The audio or speech signal for every reciter was classified into parts having frames of 20ms, selected based on the best recognition rate achieved when utilizing it. The larger frame size does not improve the recognition quality. The length of the verses varies in each chapter. Because of this variation, the MFCC technique is employed for features' extraction for every wave file or verse. Fundamentally, each wave file has 20 features, which were extracted for the 2060 and 300 acoustic waves. Essentially, all the features were combined into two different files, which were encompassed in the files of the feature matrix. One file was allotted for chapter 32 and the other file for chapter 7. The feature vector of each verse was mapped within the KNN or ANN classifier, respectively.

Table 1 shows the Holy Quran reciter names employed in the study.

Table 1: The names of Quranic reciters in English and Arabic.

No.	Reciter name in Arabic	Reciter name in English
1	بندر بليلة	Bandar Baleelah
2	عبد الرحمن السديس	Abdul Rahman Al-Sudais
3	ماهر المعيقلي	Maher Al-Muaiqly
4	ياسر الدوسري	Yasser Al-Dosari
5	سعود الشريم	Saud Al-Shuraim
6	عبد الباري الثبيتي	Abdul Bari Ath-Thobaity
7	احمد طالب	Ahmed Taleb
8	عبدالله البيجان	Abdullah Al-Baijan
9	صلاح البدير	Salah Al Budair
10	علي الحذيفي	Ali Al-Huthaify

5.1 Mel Frequency Cepstrum Coefficients (MFCC)

MFCC indicate a power spectrum representing the sound waves over a short period of time. These frequencies are examined based on the known frequencies in the critical bandwidth of the human ear. MFCC approach is the most used extraction method due to its consideration of signal representation [13]. Knowledge about the speech signal requires knowledge about the reciters' sound knowledge. More importantly, the determination of the accuracy of sound shape provides an accurate indication of the phoneme or the wave file. The power spectrum yields a shape within a less period of time. With the exploitation of the MFCC method, that commonly utilized in speech recognition systems, the features of wave files can be extracted. The mel scale frequencies were examined on the bandwidth differentiation that happens in the ears in gathering the speech characteristics [14]. In the Quranic recitation, the pronunciation and the voice alter from one

reciter to another, which provides a tone difference as follows:

$$I = 1, 2, 3\ldots p.$$

The order "p" is encompassed within the cepstral coefficients (C_i) and the magnitude coefficients' number of discrete Fourier transform (DFT). The number of filters is denoted by "N," where N = 20 in the system proposed here. The filters' log energy output is X_k. For every reciter, the maximum frequency cepstral coefficients are denoted as follow:

$$C_i = \sum_{i=1}^{N} cos\left(\frac{(k-0.5)\pi i}{N} * X_k\right). \quad (1)$$

The entire extraction process of MFCC is indicated in Figure 4.

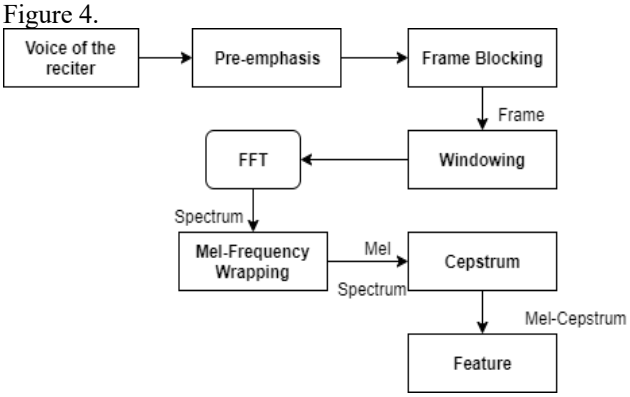

Fig. 4: MFCC features' extraction process.

Preemphasis refers to the filtering process for obtaining a smooth speech signal frequency spectral form and minimizing the noise while capturing a sound. This filtering process is essential after the sampling technique in the speech signal process. The output/input time domain relationship is considered in this preemphasis filtering process, represented in the following equation:

$$Y(n) = x(n) - a\,x(n-1), \quad (2)$$

where "a" is the preemphasis filter constant, whose value usually ranges from 0.9 to 1. The next process in MFCC is frame blocking, which indicates the audio signal's segmentation into overlapped frames. Signal deletion decreases during this process. Here, x(n) is the long audio signal. M overlapping is done on N samples: hence, N = M*2x. Windowing refers to an analytical method of fetching a section of sufficient representation from a long audio signal. This method deletes the aliasing signal due to the signal discontinuity with the finite impulse response (FIR) digital filter technique. These discontinuities are due to the blocking of the frames. Moreover, w(n) is the window, which has an average value of $0 \le n \le N-1$, where N is the number of samples encapsulated within each frame; thus, the signal window result is indicated as follows:

$$y(n) = w(n) * x(n), \text{ where } 0 \le n \le N-1. \quad (3)$$

Here, y(n) is the convolution resultant signal that occurs between the window function and the input signal. x(n) is the signal in the input function that has to undergo convolution. The value of w(n) is given as follows:

$$w(n) = -0.46\cos\left(2\frac{\pi n}{N} - 1\right) + 0.54. \quad (4)$$

A limited period function is represented in the Fourier series. Fourier transforms are utilized to convert the frames into N samples, which are transferred from the domain of time to the domain of frequency and minimize the reputation of multiplication within DFT:

$$X_n = \sum_{K=0}^{N-1} xf e^{-2\Pi kjn/N}, \quad (5)$$

where n = 0, 1, 2, 3… N-1; xf is the frame signal. A periodogram or spectrum is the result of this section's conclusion. The frequency axis unit, which represents the human speech perception, is the mel scale unit. When the frequency is lower, the interval will be narrow; whereas when the frequency is higher, the interval will be wider. At low frequencies, the variation in sound heights is understood by humans, but the high-frequency variation could not be understood by humans. Thus, the mel scale frequency is defined as follows:

$$F_{mel} = \left\{FHz, FHz < 1000\ 2595 * log10\left(1 + \frac{FHz}{700}\right)\right\}, \quad (6)$$

where $F_{Hz} > 1000$.

The human ear filter function is performed by the filter bank. In the mel frequency wrapping, the resultant FFT signal is wrapped as a triangular filter segment. In the frequency domain, the signal wrapping process is performed as follows:

$$X_i = \log_{10}\left(\sum_{K=0}^{N-1} H_i(k).\, X(k)\right), \quad (7)$$

where i = 1, 2, 3, 4…M and M is the number of triangle filters. $H_i(k)$ is the triangle filter that defines the acoustic filter k. The voice information could be listened to base on the signals in the time domain. The conversion of the mel spectrum into the time domain with the help of DCT is dealt with in this section such that the result will be the MFCC. The MFCC, C_j, is represented as follows:

$$C_j = \sum_{j=1}^{k} cos\left(\frac{\frac{j(i-1)}{2\pi}}{K}\right) * X_j, \quad (8)$$

where X_j is the mel frequency power spectrum; j = 1, 2, 3 …K, where K is the number of coefficients and M is the number of filters. The final procedure in the MFCC extraction is fetching the inverse of the DFT.

5.2 KNN Classification Algorithm

The KNN classification algorithm is employed here. Most importantly, the KNN classification is regarded as a rigid fast-speed machine learning technique and is an effective classification method [15]. The KNN classification is utilized to classify the unknown testing parameters, as the supervised learning algorithm is used in the place of the training set. For classifying the reciters, the MFCC characteristics are applied for a particular reciter and are fulfilled with the training set, and a comparison is made for the training characteristics, depending on their distances. Moreover, the testing reciters' prediction class is found based on the minimum length between the samples of the testing reciter and the training reciter, using the Euclidean distance. For instance, provided a query archetype for a

chosen reciter, the KNN to this given reciter is a commonly employed class. The distance function is a base behind this function. Fundamentally, the neighborhood values are fetched as prediction values by the KNN algorithm in the testing set. The KNN algorithm functions well for the lowest distance between the training set samples. The Euclidean distance D between the feature vectors of X and Y is indicated as follows:

$$D = \sqrt{\sum_{i=1}^{N} (xi - yi)^2}, \quad (9)$$

where x_i and y_i indicate the X and Y elements.

The KNN is utilized to automatically detect the Tajweed rules of the Holy Quran reciter. It also helps in identifying the proper recitation of the reader. The performance of Tajweed rules recognition is also implemented with the KNN. Thus, the recognition of Quranic voice is performed using the KNN classifier. The KNN classification algorithm is a mature algorithm that is embedded in the method of text classification. This algorithm has low classification time and easy interpretation and is simple and effective. The values of data involved in this proposed system are compared utilizing this classifier. The flow of the process within the KNN classification is shown in Figure 5.

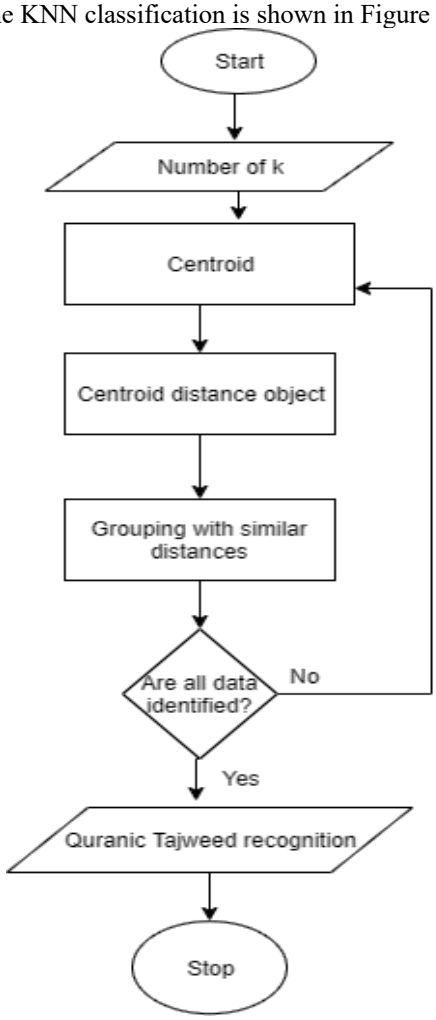

Fig. 5: The flow of process in KNN classification.

Here, the centroid is the object value of the extracted properties. The Euclidean distance formula is utilized for computing the distance. After identifying the entire data, the entire calculation of the system ends. In this section, the data are classified into testing and training data. Then, the performance of the proposed system, which involves the Quranic voice recognition of the reciters utilizing KNN and ANN classification algorithm, is finally evaluated. The performance results of KNN are compared with the actual Tajweed rules, which are incorporated into the input image. The accuracy of the Tajweed rules is evaluated using the following formula:

Accuracy in % = number of true recognition results/total number of testing images * 100%

Additional Tajweed-based recitation of the Holy Quran could also be guided with the KNN classification algorithm.

5.3 Artificial Neural Networks (ANN)

ANNs are mapping structures that are nonlinear and are based on human brain functions [16]. The classification techniques are based on the feature extraction nature. Here, the ANN is in the form of a classifier. This network has been used successfully in speech recognition systems. The ANN is a computational model or a nonlinear system, which encompasses multiple processing parts. The core component behind ANN is the artificial neurons. An input is obtained by one neuron from other neurons and the mathematical output is passed on to other neurons. Generally, these networks encompass three important types of layers: the output layer, hidden layer, and the input layer [17]. The initial data for ANN is in the input layer, which are the features to be extracted. The intermediate layer or the hidden layer lies between the input and output layers. The necessary computations of ANN are performed under the intermediate layer. The classes are generated as output in ANNs. Depending on the predefined activation function, the output of ANNs takes place. There are multiple varieties of ANNs, such as modular neural networks, convolutional neural networks, and multilayer perceptron (MLP). In the proposed system, MLP combined with a backpropagation learning algorithm is employed. ANN is utilized for recognizing and training the reciters of the Holy Quran. Figure 6 indicates the architecture of ANNs [18-19].

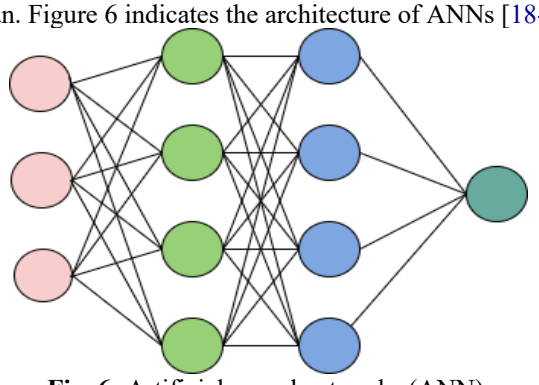

Fig. 6: Artificial neural networks (ANN).

The following are the parameters, which are encompassed within the MLP topology, which are designed for the reciters of the Holy Quran:

- The features are encompassed within the input layer.
- Twenty neurons are encompassed within the first intermediate layer.
- Twenty neurons are encompassed within the second intermediate layer.
- Fifteen neurons are encompassed within the output layer.

The backpropagation learning algorithm was employed for the training and performance. There are some limitations associated with the use of ANNs in voice recognition as follows. First, ANN is not a solution to the voice recognition problem and lacks a structured methodology. Second, sometimes, the output quality offered by ANN is unpredictable. Third, there is no accurate description of the problem-solving methodology of ANN. Fourth, it has a block box and an empirical nature.

5.4 Support Vector Machines (SVM)

SVM is a famous discriminative and powerful classifier. The objective of SVM is to maximize the margin of separation between two different classes that are competing with each other, and the definition of margin is that it is the distance between the closest training archetypes and the decision hyperplane. Moreover, it is defined as the determination of the line that separates the classes with the maximum distance between the samples and the line and the sample. This line is referred to as a hyperplane. Support vectors are the samples that are nearer to the hyperplane. The largest distance between the hyperplane and the support vectors is called the margin [20-22].

6 Experimental Evaluation and Results

The present system of recognizing the voice of the Quranic reciters focuses significantly on the extracted features of the MFCC from the classification, the learning algorithm, and the acoustic wave files. About 20 MFCC features are extracted for ten readers of the study, which is combined into one file. Moreover, two significant files were combined: one file is allotted for chapter 7 and the other for chapter 36. In this section of the experimental evaluation, a cross-validation is utilized to verify the performance of the proposed system. Moreover, 70% of the data are utilized as the training data and 30% are utilized as the testing data. The testing and training data are selected randomly away from the file combined. Thus, the training and testing experiments are repeated randomly for about six times, which include the selection of data for testing and training, 30% and 70%, respectively. Table 2 indicates the average KNN and ANN recognition results for the reader Abdul Rahman Al-Sudais for chapters 7 and 32.

Table 2: Accuracy of chapters 7 and 32 with KNN and ANN classification.

Experiment No.	Accuracy of chapter 7, %		Accuracy of chapter 32, %	
	KNN	ANN	KNN	ANN
1	97.2	97.7	96.2	96.6
2	97.1	97.3	96.7	97.1
3	97.4	98.1	96.4	96.9
4	97.1	97.8	96.8	97.4
5	97.8	98.2	96.5	97.1
6	96.9	97.5	97.1	96.8
Average	97.25	97.8	96.62	96.9

Figure 7 summarizes the recognition rate for six various experiments and their associated results for the above-mentioned reciter of the Holy Quran.

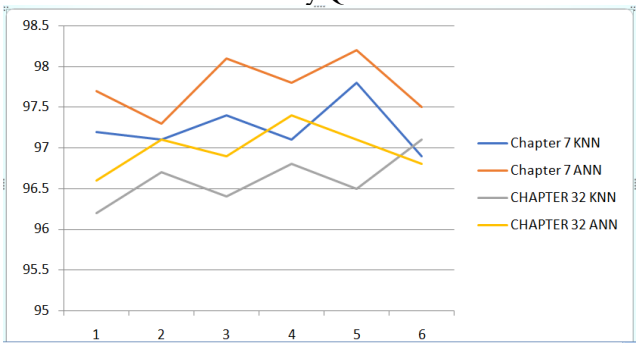

Fig.7: Recognition rate for two chapters read by Abdul Rahman Al-Sudais.

Table 3: The average value of accuracy of ANN.

No.	Name of the reciter	Accuracy of chapter 7, %		Accuracy of chapter 32, %	
		KNN	ANN	KNN	ANN
1	Bandar Baleelah	97.1	97.5	95.9	96.7
2	Abudl Rahman Al-Sudais	96.9	97.4	95.4	96.2
3	Maher Al-Muaiqly	97.1	97.9	96.1	96.9
4	Yasser Al-Dosari	97.1	97.8	95.8	96.9
5	Saud Al-Shuraim	97.2	97.9	97.5	97.2
6	Abdul Bari Al-Thubaity	96.8	97.3	95.7	96.3
7	Ahmed Taleb	97.2	97.6	95.9	96.8
8	Abdullah Al-Baijan	97.3	97.6	95.8	96.5
9	Salah Al Budair	97.1	97.4	95.1	96.6
10	Ali Al-Huthaify	97.25	97.8	96.62	96.9
Average		97.03	97.62	96.08	96.7

Table 3 indicates that the average value of accuracy of ANN is 97.62% and 96.7% for chapter 7 and chapter 32, respectively. Almost 3.5% and 2.5% of the total data are misclassified. The average accuracy of KNN is 97.03% and and 96.08% for chapter 7 and chapter 32, respectively. The important reason for misclassification is the way for reciting and reading various verses. The expressed reciter

emotion and the tone variation are critical components. This is because of the Tajweed reciting rules, where all the reciters follow the same rules. When the current system was compared to the existing systems, determining the performance of the proposed system proved difficult and the comparison was challenging, which is due to the different datasets and criteria used in the already existing systems. Moreover, different reciters were used in the existing works.

For the purpose of comparison, the same procedure was repeated with SVM, which is another machine learning algorithm. The same audio files were given as input to the SVM with 70% training ratio and 30% testing ratio. The comparison results indicated a convergence observed in ANN, where it stopped its progression when a certain accuracy level was reached. When the input voice data are taken from Surah Al-A'raf and are given as an outlet data to the ANN classifier, the accuracy percentage drops slightly. When comparing these two algorithms, ANN is observed to be a better algorithm, as the performance of SVM with a greater number of features and a smaller number of samples is not as good as that of ANN. Moreover, the implementation of SVM requires more effort than that for ANN.

7 Conclusions

This research paper proposes a machine learning model for identifying the Holy Quran reciter using the ANN and KNN classifiers. The performance of both ANN and KNN classifiers was evaluated. Fundamentally, the extraction of the MFCC features and the mapping of these features in ANN and KNN for testing and training are carried out. Two Quranic chapters were chosen to be recited by the prayer leaders of the Holy Mosques in Makkah and Madinah. The average accuracy value of ANN is 97.62% for chapter 7 and 96.7% for chapter 32. The average accuracy of KNN is 97.03% for chapter 7 and 96.08% for chapter 32. The future research of the proposed study opts for enhancing the performance of the system and increasing the accuracy value of the system. Other machine learning algorithms, such as SVM and HMM, could be utilized to enhance accuracy of the proposed system. Furthermore, the wave signals can be incorporated with sliding window characteristics to enhance the system's accuracy.

Acknowledgement

The authors would like to thank the Deanship of Scientific Research at Jouf University for supporting this work, Grant Code: DSR-2021-04-0319.

Funding Statement

This work was funded by the Deanship of Scientific Research at Jouf University under Grant No. DSR-2021-04-0319.

Conflicts of Interest

The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] B.Putra , B.T. Atmaja ,and D.Prananto. Putra, Budiman, B. Atmaja, and D. Prananto, Developing speech recognition system for Quranic verse recitation learning software, *International Journal on Informatics for Development.*, **1(2)**, 14-21 (2012).

[2] R. Khan, M. Hadwan, and A. M. Qamar, Quranic Reciter Recognition: A Machine Learning Approach, *Advances in Science, Technology and Engineering Systems Journal.,* **4(6)**, 173-176 (2019).

[3] J. H. Alkhateeb, A Machine Learning Approach for Recognizing the Holy Quran Reciter, *International Journal of Advanced Computer Science and Applications(IJACSA).*, **11(7)**, 268-71 (2020).

[4] K. M. O. Nahar, M. Al-Shannaq, A. Manasrah, R. Alshorman, and I. Alazzam, A Holy Quran Reader/Reciter Identification System Using Support Vector Machine, *International Journal of Machine Learning and Computing.,* **9(4)**, 458-464 (2019).

[5] Ahmed. I. Taloba, A. A. Sewisy, and Y. A. Dawood, Accuracy enhancement scaling factor of Viola-Jones using genetic algorithms, *14th International Computer Engineering Conference (ICENCO).*, 209-212 (2018).

[6] M. Al-Ayyoub, N. A. Damer, and I. Hmeidi, Using Deep Learning for Automatically Determining Correct Application of Basic Quranic Recitation Rules, *Int. Arab J. Inf. Technol*, **15**, **3**, 620-625 (2018).

[7] H. M. A. Tabbaa, and B. Soudana, Computer-Aided Training for Quranic Recitation, *Procedia-Social and Behavioral Sciences*, **192**, 778-787 (2015).

[8] N. Shafie, M. Z. Adam, S. M. Daud, and H. Abas, A Model of Correction Mapping for Al-Quran Recitation Performance Evaluation Engine, *Int. J. Adv. Trends Comput. Sci.*, **8**, 208-13 (2019).

[9] I. N Jamaliah, M. Y. I. Idris, Z. Razak, and N. N. Abdul Rahman, Automated tajweed checking rules engine for Quranic learning, *Multicultural Education & Technology Journal.*, **7(4)**, 275-287 (2013).

[10] S. M. Kamruzzaman , A. N. M. Rezaul Karim , M. Saiful Islam and M. E. Haque, Speaker Identification Using MFCC-Domain Support Vector Machine, *International Journal of Electrical and Power Engineering.*, **1(3)**, 274-278 (2007).

[11] A. M. Alagrami, and M. M. Eljazzar, Smartajweed Automatic Recognition of Arabic Quranic Recitation Rules, *Computer Science & Information Technology (CS & IT).*, **10(18)**, 145 (2020).

[12] S. Khairuddin, S. Ahmad, A. Embong, N. N. W. N. Hashim, and T. M. K. Altamas, Syarifah Nuratikah Syd Badaruddin, Surul Shahbudin Hassan. Classification of the Correct Quranic Letters Pronunciation of Male and Female Reciters, *IOP Conference Series: Materials Science and Engineering.*, **260(1)**, 012004 (2017).

[13] L. Marlina, C. Wardoyo, W. M. Sanjaya, D. Anggraeni, S.

F. Dewi, A. Roziqin, and S. Maryanti, Makhraj recognition of Hijaiyah letter for children based on Mel-Frequency Cepstrum Coefficients (MFCC) and Support Vector Machines (SVM) method, *International Conference on Information and Communications Technology (ICOIACT).*, 935-940 (2018).

[14] C. K. On, P. M. Pandiyan, S. Yaacob, and A. Saudi, Mel-Frequency Cepstral Coefficient Analysis in Speech Recognition. *International Conference on Computing Informatics.*, 1-5 (2006).

[15] G. Guo, H. Wang, D. Bell, Y. Bi, and K. Greer, KNN Model-Based Approach in Classification, *OTM Confederated International Conferences "On the Move to Meaningful Internet Systems".,* 986-996 (2003).

[16] S. Lek, and Y. S. Park, Artificial Neural Networks, *Encyclopedia of Ecology.*, 237–245 (2008).

[17] B. Dickson, What Are Artificial Neural Networks (ANN)?, *TechTalks.*, (2019).

[18] M. M. Abdelgwad, T. H. A Soliman, Ahmed I.Taloba, and M F. Farghaly, Arabic aspect based sentiment analysis using bidirectional GRU based models, *Journal of King Saud University - Computer and Information Sciences.*, (2021).

[19] I. Abdalla Mohamed, A. Ben Aissa, L. F. Hussein, Ahmed I. Taloba, and T. kallel, A new model for epidemic prediction: COVID-19 in kingdom saudi arabia case study, *Materials Today: Proceedings.*, (2021).

[20] Ahmed I. Taloba and S. S. I. Ismail, An Intelligent Hybrid Technique of Decision Tree and Genetic Algorithm for E-Mail Spam Detection, *Ninth International Conference on Intelligent Computing and Information Systems (ICICIS).*, 99-104 (2019).

[21] Ahmed I. Taloba, M. R. Riad and T. H. A. Soliman, Developing an efficient spectral clustering algorithm on large scale graphs in spark, *Eighth International Conference on Intelligent Computing and Information Systems (ICICIS).*, 292-298 (2017).

[22] O. R. Shahin, R. M. Abd El-Aziz, and Ahmed I. Taloba. Detection and classification of Covid-19 in CT-lungs screening using machine learning techniques, Journal of Interdisciplinary Mathematics.,1-23 (2022).