Research Article

An Intelligent Framework Based on Deep Learning for Online Quran Learning during Pandemic

Natasha Nigar [iD],[1] **Amna Wajid,**[1] **Sunday Adeola Ajagbe** [iD],[2] **and Matthew O. Adigun** [iD][3]

[1]*Department of Computer Science (RCET), University of Engineering and Technology, Lahore, Pakistan*
[2]*Department of Computer and Industrial Production Engineering, First Technical University, Ibadan, Nigeria*
[3]*Department of Computer Science, University of Zululand, Richards Bay, South Africa*

Correspondence should be addressed to Natasha Nigar; natasha@uet.edu.pk

The COVID-19 pandemic influenced the whole world and changed social life globally. Social distancing is an effective strategy adopted by all countries to prevent humans from being infected. Al-Quran is the holy book of Muslims and its listening and reading is one of the obligatory activities. Close contact is essential in traditional learning system; however, most of the Al-Quran learning schools were locked down to minimize the spread of COVID-19 infection. To address this limitation, in this paper, we propose a novel system using deep learning to identify the correct recitation of individual alphabets, words from a recited verse and a complete verse of Al-Quran to assist the reciter. Moreover, in the proposed approach, if the user recites correctly, his/her voice is also added to the existing dataset to leverage proposed approach effectiveness. We employ mel-frequency cepstral coefficients (MFCC) to extract voice features and long short-term memory (LSTM), a recurrent neural network (RNN) for classification. The said approach is validated using the Al-Quran dataset. The results demonstrate that the proposed system outperforms the state-of-the-art approaches with an accuracy rate of 97.7%. This system will help the Muslim community all over the world to recite the Al-Quran in the right way in the absence of human help due to similar future pandemics.

1. Introduction

The COVID-19 pandemic has significantly impacted the whole world causing both unprecedented disruptions and massive changes in the lives of people. From living and social lifestyles to business models, all have suffered [1]. The geographic distance between countries has reduced while the social distance between humans has increased [2]. To prevent the spread of COVID-19 infection, different countries took diverse measures, e.g., complete lockdown, work from home, both religious and educational institutes closure, and suspension of transport facilities.

Arabic is one of the four major languages which are widely spoken in the world [3]. Despite the fact that a large number of people in the world speak the Arabic language, the problem arises as they speak Arabic in their own accents which makes it difficult to understand as well as it changes the meaning in a specific context. These accents made the appearance of the Arabic language totally different, especially for someone who does not know Arabic well enough to understand due to the difference in speaking styles. Like other major languages of the world, Arabic has its own specific set of rules called Arabic grammar. The correct pronunciation of Arabic alphabets is a real challenge.

Learning Arabic is not necessary for everyone; however, for Muslims, it is necessary to learn due to religious binding to correctly recite Al-Quran which is the sacred book of Islam. It is composed of 30 chapters, 6,236 verses grouped into 114 units called "Surahs." The reciter of Al-Quran needs to follow a set of recitation rules called as Tajweed in Arabic to recite the Al-Quran without errors and ensure the delivery of the correct meaning of the verses. The common practice in Muslim's society is that the students learn the Quran by listening to the voice of their instructor, who is a Tajweed expert (a.k.a Hafiz).

Muslims faced difficulty in learning Al-Quran during the pandemic as it requires face-to-face supervision of teachers a.k.a. Hafiz. It is also possible that these teachers were not available at all the places. Moreover, many Muslims do not have access to Hafiz and there are others who are elder and want to learn Al-Quran but feel embarrassed while going to Hafiz to learn Al-Quran. It is also a fact that all Muslim parents want to make their children an expert reciter of Al-Quran and ensure that they learn the correct rules of Tajweed. However, they do not want to send their children to Hafiz and prefer to have a home tutor. These issues accelerate the need of a computer-aided system which can replace the Hafiz and is available to everyone.

The traditional approaches are limited to recite the Al-Quran using words and verses with Tajweed or speech-to-text translation. To overcome the issues discussed in the previous paragraph, a deep learning approach (LSTM-RNN) using MFCC for voice feature extraction is presented which helps the reciter in learning Al-Quran by alphabets, words, and verses with Tajweed. If the reciter reads Al-Quran in the wrong way, the system notifies the user and asks him/her to re-recite until acceptable pronunciation level is achieved. The proposed system also gives an option to the user if he/she wants to listen to the recitation of an expert. In this paper, we present a novel system with enhanced accuracy and make the following contributions:

(i) New system design: A novel system is designed to ease Al-Quran reciters to learn Quran accurately and quickly. The proposed system helps users to learn Al-Quran recitation by alphabet, word, and verses.

(ii) Wrong Tajweed detection: A novel approach to detect and inform the reciter for wrong Tajweed.

(iii) Mobile application: A mobile application to help Al-Quran reciter in learning Al-Quran.

(iv) Augmented dataset: If the reciter reads the verse correctly, his/her voice is also added to the database.

This paper is structured in 6 sections. Section 2 presents related work followed by the proposed methodology in Section 3. The experimental design and analysis are presented in Section 4. The threats and validity are discussed in Section 5. This paper concludes with Section 6, where results are discussed and future directions are proposed.

2. Literature Review

2.1. Background

2.1.1. The Arabic Language and Tajweed Rules. The Arabic language is one of the most widely used languages in the world. It has religious importance because the holy book of Muslims Al-Quran is also in Arabic language. It has twenty-eight consonants and three vowels. The vowels are further classified as short and long [4]. There are multiple challenges associated with Arabic speech recognition due to its rich vocabulary, multiple dialects, and different accents. The main objective of Muslims is to read Quran accurately and

with correct Tajweed rules. The recitation of Al-Quran according to correct Tajweed rules is the main difference between Arabic and Quranic ASR. This makes Quranic ASR more challenging as compared to simple Arabic ASR. Moreover, the reciter of Quran can add emotions in recitation due to which the sound of phonemes transit from one acoustic level to the other [4]. There are seven famous recitation styles known as "Qira'at" of Al-Quran. Among these seven "Qira'at" styles, "Nafi" and "Aasim" are the most famous. These styles differ from one another based on lengthening and shortening of words called as "madd" and "qasr," respectively [5].

Tajweed means reciting with correct pronunciation and according to rules at a normal speed. There are well-defined Tajweed rules for Al-Quran recitation. Reading Quran with Tajweed rules makes its recognition challenging as compared to the normal Arabic language. The recitation of the same verse by different reciters also varies from one another due to the recitation style, Tajweed rules, and use of "harakattes." Some of the Tajweed rules, which may affect the recitation, are as follows [5]:

(i) Use of "maqams"

(ii) Pronunciation of some phonemes using nasalization

(iii) Shortening and lengthening of sounds of different phonemes

(iv) Use of "Ghonna" and "Tanween" rules

(v) Rule for combining words

2.1.2. Recurrent Neural Network Model. The recurrent neural network (RNN) is a powerful model for sequential and time series data [6]. It is trained in such a way that their internal state provides a powerful, general framework to model time series. Moreover, they also tend to be robust to temporal and spatial noise [7]. LSTM is a variant of RNN (Figure 1). It is designed to model temporal sequences. It has an extra long-term memory and input, output, and forget gates. At each iteration, these gates try to remember when and how much information in the memory cells should be updated [6]. A single LSTM memory cell is shown in Figure 2.

By keeping in view all the challenges associated with Al-Quran recitation learning, there must be an expert who can teach the correct way of reading Quran. That expert can be some Hafiz, and in the case of online learning, it can be some ASR system. A perfect ASR system for Quran learning is difficult to develop due to the diversity in applying Tajweed rules and different accent of recitation by different reciters.

2.2. Related Works. A lot of work has been done to classify Al-Quran recitation as correct or incorrect based on Tajweed rules using machine learning (ML). In this regard, Nahar et al. [4] developed an Al-Quran reader identification system using MFCC [12] for the feature extraction and support vector machine (SVM) and artificial neural network (ANN) for the recognition of a speaker. Al Anazi and Osama [5]
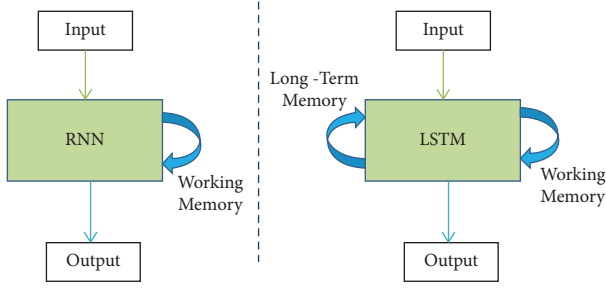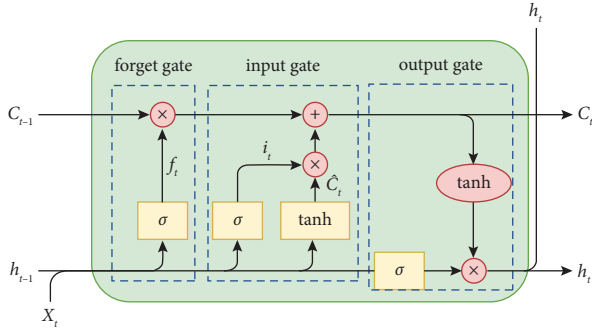
FIGURE 1: RNN vs. LSTM working.

FIGURE 2: An LSTM cell.

proposed a ML-based model for Quran reciter identification using MFCC, K-Nearest Neighbor (KNN), and ANN. They compared the performance of KNN and ANN classifiers for reciter identification. Zerari et al. [13] proposed a system for recognition of Arabic digits and some of the TV commands. They used MFCC and filter bank (FB) both for feature extraction and long short-term memory (LSTM) and neural network (multilayer perceptron (MLP)) as classifiers. All the above works refer to the recognition of simple Arabic language and Arabic reciter identification to classify a recitation as correct or incorrect.

In the literature, there is work on the recognition of basic Arabic, but there are limited studies to teach the Al-Quran using machine learning. Alkhatib et al. [8] developed a mobile application to detect correct or incorrect pronunciation of Arabic words. They used MFCC and dynamic time warping (DTW) algorithm to extract features of sound and compare the sound of the teacher (stored at the server) with the learner. The application classifies the spoken word as correct or incorrect depending on the results of the comparison algorithm.

Al-Ayyoub et al. [9] developed a deep learning-based system to automatically classify a recitation as correct or incorrect based on the Ahkam Al-Tajweed rules. They modeled 8 Ahkam Al-Tajweed rules. The recitation by a user is compared with the audio recording dataset using different feature extraction techniques such as MFCC, linear predictive code (LPC), and convolutional deep belief network (CDBN). Finally, the recitation is classified as correct or incorrect using different classifiers, e.g., SVM and random forest (RF). This is just a classification system for recitation. Farooq and Imran [10] developed a real-time application for mispronunciation detection of Arabic letters according to Tajweed rules. The

decision of correct or incorrect pronunciation is based on articulation points of letters. They considered 29 letters of Arabic language and used RASTA PLP for feature extraction and HMM for training and recognition. Touati-Hamad et al. [14] applied deep learning algorithm (LSTM) to automatically authenticate the integrity of the Quran's verses sequence. They [15] also applied deep learning along with word embedding techniques to distinguish a Quranic verse from a non-Quranic Arabic text. Shahriar and Tariq [16] used various deep learning algorithms (convolutional neural network (CNN), deep ANN, and LSTM) for the classification of the eight popular maqamat (melodies). Al-Zaben et al. [17] analyzed the heart rate variability while listening to Quran recitation. However, these studies do not consider about the accurate recitation of the Holy Quran.

An automatic speech recognition (ASR) to recognize individual words from a selected verse of Quran was developed by Ghori et al. [11] using machine learning. They used MFCC for feature extraction from sound, followed by a four-layer DNN model. This system helps users to listen a selected verse or a selected word from a particular verse and transcribe the user's recitation. The limitation of this system is that it is dependent on the accent of a speaker as this system has been trained using recitation of few speakers. The comparative analysis of these studies is summarized in Table 1. Few studies [18–20] also used deep learning algorithms for Arabic speech recognition and correct pronunciation. Besides these, some studies provide surveys of Al-Quran and Arabic language recognition [21–23].

The above literature highlights the fact that these studies are limited to reciter identification by comparing voice signals and recognition of Arabic language. In this study, we propose a novel deep learning-based system which helps the user in reciting and learning the Al-Quran by alphabets, words, and verses with high accuracy.

3. Proposed Methodology

In this section, the proposed methodology to learn the Al-Quran is presented. The proposed system (Figure 3) consists of 3 main stages, namely, preprocessing, feature extraction, and testing and training.

3.1. Data Preprocessing. Prepossessing of audio signals is considered a crucial step for robust and efficient audio analysis and modeling to transform complicated and noisy raw audio data into a format that is appropriate for further analysis [24]. In this step, the quality of the audio signal is improved by reducing the signal-to-noise ratio, reducing the noise, and filtering the signal. The significant advantage of this step is to determine the beginning and ending boundaries of each word, right organization of information, the simplification of data, and then the simplification of voice recognition.

3.2. Features Extraction. Feature extraction is the process of obtaining different features such as power, pitch, and vocal tract configuration from the audio signal. In this step,

TABLE 1: Comparative analysis of existing studies.

Study	ASR of verse	ASR of word	ASR of alphabets	Feature extraction	Training model
Alkhatib et al. [8]		Comparison-based techniques for Arabic language recognition				
Al-Ayyoub et al. [9]	No	No	No	MFCC, LPC, CDBN	KNN, SVM, RF
Farooq and Imran [10]	No	No	Yes	RASTA PLP	HMM
Ghori et al. [11]	Yes	Yes	No	MFCC, DNN	DNN
Our study	Yes	Yes	Yes	MFCC	LSTM-RNN
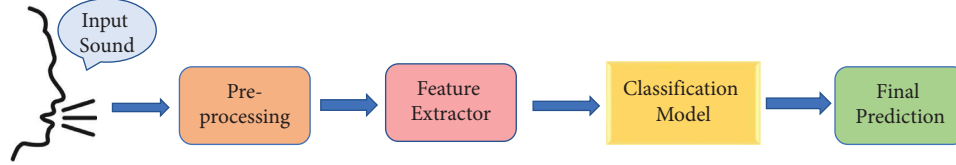
FIGURE 3: The workflow of the proposed methodology.

features are extracted with a specific number of values or coefficients. These are generated by applying the MFCC method on the input audio signal. The benefit of using MFCC is that it effectively reduces errors, eliminates all the redundant data, and generates a strong feature even in situations where noise is interfering with the signal [25, 26]. This method (Figure 4) also converts the unlimited continuous speech to limited sequence. After that, the audio signal is converted from time to frequency domain so that its frequency domain analysis is done. Every human has a different voice frequency, so it is a good option to use the signal's frequency domain properties such as the magnitude and phase to differentiate between two signals and measuring the similarity ratio between them. The fast Fourier transform (FFT) is a function used to convert the signal from time domain to frequency domain. Apart from using FFT, the FFT shift is used to circularly shift the values of FFT around zero so that the FFT is easily understandable.

3.3. Model Training and Testing.
After the features are extracted using the MFCC method, these are mapped into the ML classifier for testing and training to identify the correct recitation of the Al-Quran. We used a deep LSTM-RNN model. The reason for choosing LSTM-RNN is that it has proven for significant performance when dealing with sequence data like speech [27–29]. It is a popular recurrent network to store and retrieve memory by using gates [30]. The LSTM detects the phonemes in individual words with higher accuracy and equal efficiency. 70% of audio signals are used for training of the ML model and 30% for testing. Moreover, if the reciter read the Al-Quran in the correct way, then the audio signal also becomes a part of the training dataset.

4. Experimental Analysis

4.1. Experimental Setup

4.1.1. Dataset.
In this paper, the dataset is divided into chapters, Surahs, and verses. The chapters are stored with their name in Arabic, whereas Surahs are stored with

metadata, i.e., English name, Arabic name, place of revelation, number of rukoos, number of verses, Surah number, and number of Sajdas. The verses are stored with chapter number, Surah number, and verse number as attributes.

4.1.2. Parameters.
Table 2 shows the hyperparameters of the LSTM model used in the experiments.

4.1.3. Performance Measures.
Performance measures are used to evaluate the proposed system, accuracy, precision, and recall. The results for classification problems are either normal case or abnormal, named as positive class or negative class, respectively. The prediction results can also be either true or false, implying correct or incorrect prediction, respectively. Thus, we can categorize classification in the following four possible states:

(i) True positive (TP): correct prediction of positive class

(ii) True negative (TN): correct prediction of negative class

(iii) False positive (FP): incorrect prediction of positive class

(iv) False negative (FN): incorrect prediction of negative class

The precision and recall are calculated as follows:

$$accuracy = \frac{TP + TN}{FP + TN + TP + FN},$$

$$precision = \frac{TP}{TP + FP}, \tag{1}$$

$$recall = \frac{TP}{TP + FN}.$$

4.1.4. Experimental Environment.
All experiments are run on MATLAB using Android to develop a mobile application on a personal computer with Intel Core i5, 3.2 GHz CPU, and 16 GB RAM.
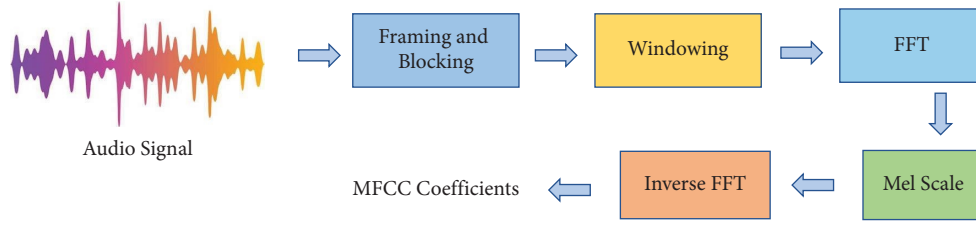
Figure 4: MFCC features' extraction process.

Table 2: Hyperparameters of the LSTM.

Parameters	Values
Batch size	32
Learning rate	0.001
Activation	Tanh
Optimizer	Adam
Epochs	30
Dropout rate	0.5
Train/validation/test split	70%/20%/10%

4.2. Results and Discussion. The user recites the Al-Quran and an audio signal is generated (Figure 5). This audio signal has infinite voice sequence (Figure 6). To address this issue, windowing is performed using MFCC to convert unlimited continuous speech into limited sequence. The waveform which is produced after windowing and the corresponding finite sequence are shown in Figures 7 and 8.

After windowing of the signal, it is converted from time to frequency domain for analysis. The fast Fourier transform (FFT) is a function used to convert the signal from time domain to frequency domain. Apart from using FFT, the FFT shift is used to circularly shift the values of FFT around zero to extract the voice features. The signal after applying FFT is shown in Figure 9.

The extracted features are given to LSTM to train the model. The results in Table 3 show that LSTM achieves an accuracy rate of 97.7% and 97.6% and F1-score of 97.4%. The results are compared to the existing state-of-the-art algorithms and it can be observed that the proposed model outperforms the existing models.

The proposed system notifies the user by highlighting the text in red color, if the user recites (Al-Quran alphabet and verse) in a wrong way, as shown in Figures 10 and 11, and in green color if vice versa, as shown in Figures 12 and 13.

5. Threats to Validity

This study aims to help Al-Quran reciter to learn the Quran in the right way. Acquiring a sufficient amount of high-quality data encompassing different recitation styles, accents, and pronunciations is a challenging task. Also, designing and training a deep learning model that can accurately recognize and differentiate various recitation styles requires a complex architecture. Therefore, there are few limitations to this work. First, the proposed system is only able to compare the user's voice on certain Surahs.
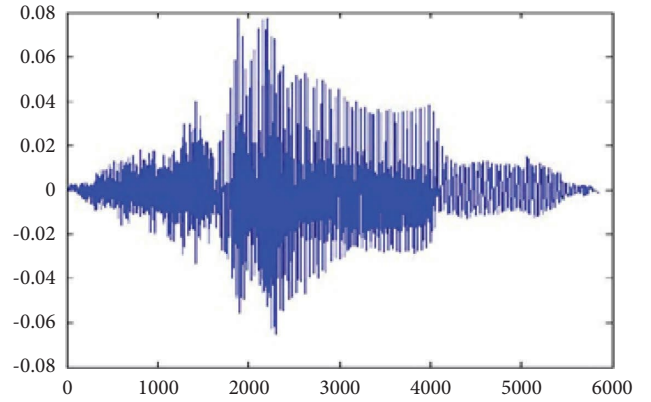
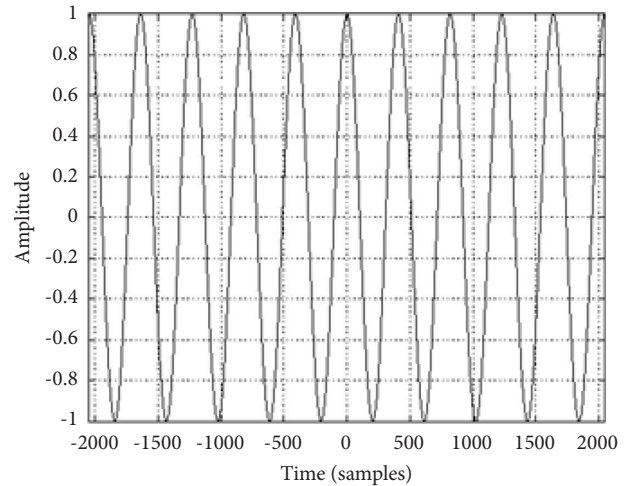

Figure 5: Input audio signal.

Figure 6: Infinite voice sequence.

Second, every Arabic word has its own tone due to which it is a complex task to design a system which can detect different accents. At present, the proposed system is capable of detecting a single accent which is stored in expert's voice in the database. Third, our system processes single alphabet isolated words and small verses of selected Surahs of Al-Quran rather than long verses. Furthermore, the system could also be trained on a larger scale by adding Arabic language learning data for pronunciation accuracy. For this, availability of high-quality, diverse, and well-annotated data for training deep learning models can be a significant challenge.
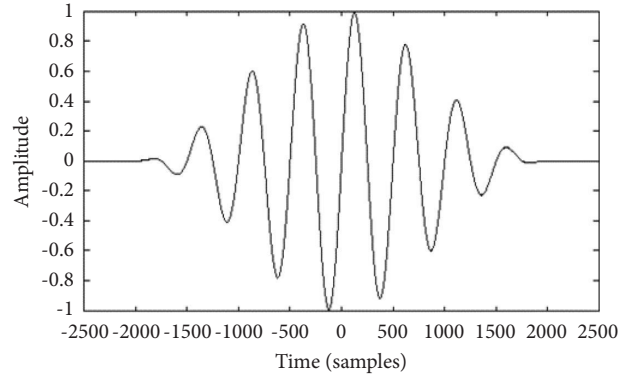
Figure 7: Audio signal windowed sequence.

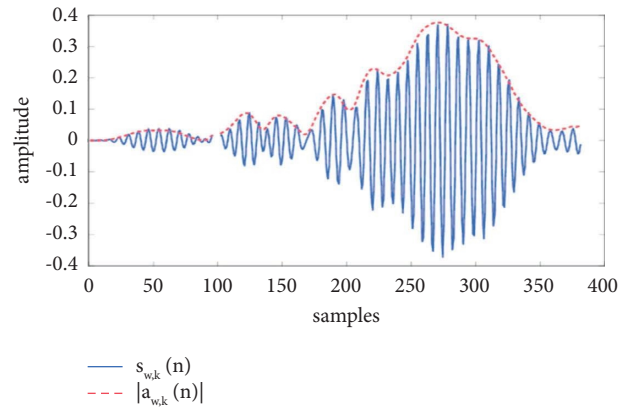—— $s_{w,k}(n)$
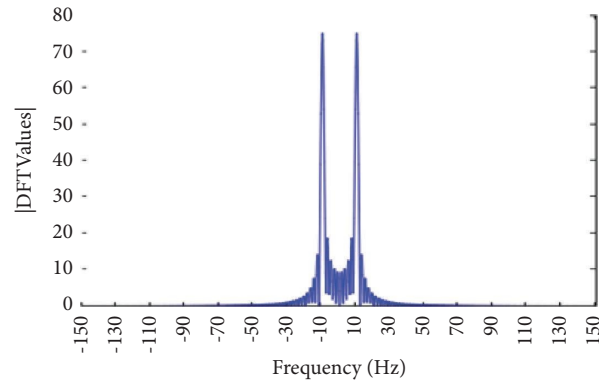- - - $|a_{w,k}(n)|$

Figure 8: Finite voice sequence.

Figure 9: Audio signal after applying FFT and FFT shift.

Table 3: Comparative analysis of performance measures with state-of-the-art algorithms.

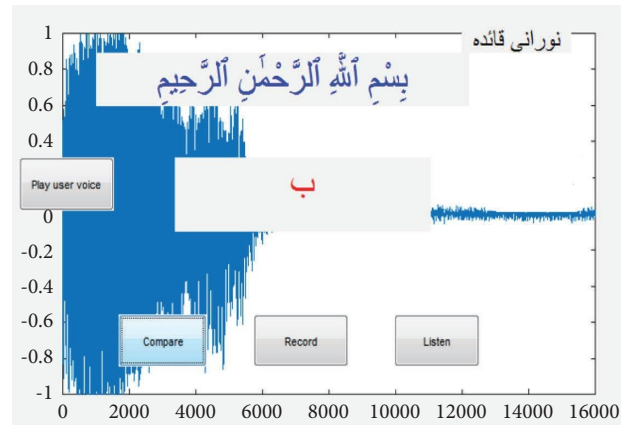|	Accuracy	Precision	F1-score
Al-Ayyoub et al. [9]	96.4%	96.6%	96.3%
Farooq and Imran [10]	98% (only words)	—	—
Ghori et al. [11]	97.5%	—	—
Proposed system	97.7%	97.6%	97.4%

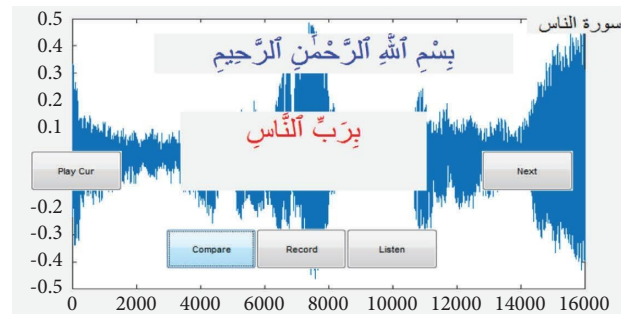FIGURE 10: Wrong alphabet recitation illustration.

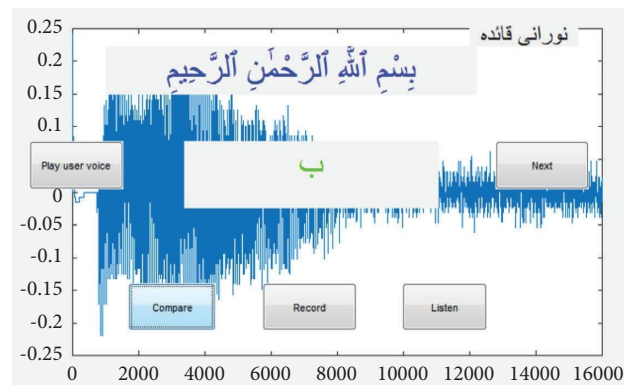
FIGURE 11: Wrong verse recitation illustration.

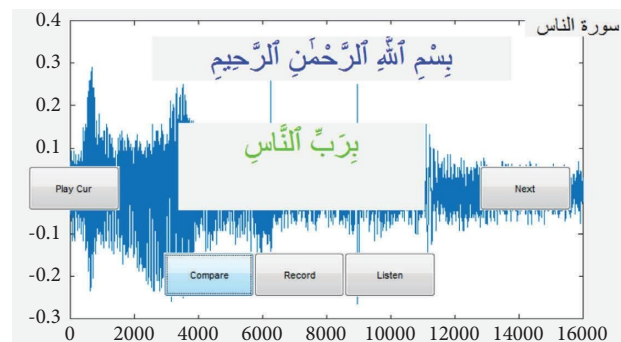FIGURE 12: Correct alphabet recitation illustration.

FIGURE 13: Correct verse recitation illustration.

6. Conclusion

The COVID-19 resulted in the closure of both religious and educational institutions. Islam is one of the biggest religions in the world. It is the obligation of every Muslim to learn and read the Al-Quran being the sacred book. Muslims have to learn and read the Al-Quran in an accurate way a.k.a Tajweed, to ensure the delivery of the correct meaning of Quranic verses. This requires learning the Al-Quran with the help of an expert which is not physically available during pandemic.

In this paper, a deep learning-based approach LSTM-RNN is used to develop the model with MFCC for voice features extraction which helps the users to learn the Al-Quran in a correct and efficient way. The users can learn by alphabets, words, and verses. The results show that the proposed system gives the highest accuracy rate of 97.7%, precision of 97.6%, and F1-score of 97.4% as compared to existing algorithms. Also, it diagnoses the wrong Tajweed accurately and notifies the user to relearn again.

In future, the proposed approach will be processed with long verses while considering processing of whole Al-Quran. Moreover, efforts will be made to include more accents to make our system more reliable and efficient. We will also train the system for Arabic language learning with accurate pronunciation.

Data Availability

The data used in this study are available upon request from the corresponding author.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

Natasha Nigar contributed to the design and development and original draft writing of this study. Amna Wajid worked on the proposed methodology. Sunday Adeola Ajagbe contributed in original draft writing and analyzed the results. Matthew O. Adigun reviewed the article.

References

[1] S. J. Daniel, "Education and the covid-19 pandemic," *Prospects*, vol. 49, no. 1-2, pp. 91–96, 2020.

[2] T. P. Velavan and C. G. Meyer, "The covid-19 epidemic," *Tropical Medicine and International Health*, vol. 25, no. 3, pp. 278–280, 2020.

[3] W. Wright and C. Paul Caspari, *A Grammar of the Arabic Language*, Cosimo, Inc, Duisburg, Germany, 2011.

[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*, vol. 9, no. 4, pp. 458–464, 2019.

[5] M. M. Al Anazi and R. S. Osama, "A machine learning model for the identification of the holy quran reciter utilizing k-nearest neighbor and artificial neural networks," *Information Sciences Letters*, vol. 11, no. 4, pp. 1093–1102, 2022.

[6] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[7] A. Graves, F. Santiago, F. Gomez, and J. Schmidhuber, "Connectionist temporal classification: labelling unsegmented sequence data with recurrent neural networks," in *Proceedings of the 23rd international conference on Machine learning*, pp. 369–376, Pittsburgh, PA, USA, June 2006.

[8] B. Alkhatib, M. Kawas, A. Alnahhas, B. Rama, 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, p. 478, 2017.

[9] M. Al-Ayyoub, N. A. Damer, and H. Ismail, "Using deep learning for automatically determining correct application of basic quranic recitation rules," *The International Arab Journal of Information Technology*, vol. 15, no. 3A, pp. 620–625, 2018.

[10] J. Farooq and M. Imran, "Mispronunciation detection in articulation points of Arabic letters using machine learning," in *Proceedings of the 2021 International Conference on Computing, Electronic and Electrical Engineering (ICE Cube)*, pp. 1–6, IEEE, Quetta, Pakistan, October 2021.

[11] A. F. Ghori, A. Waheed, M. Waqas, A. Mehmood, and S. A. Ali, "Acoustic modelling using deep learning for quran recitation assistance," *International Journal of Speech Technology*, vol. 26, pp. 113–121, 2022.

[12] B. Logan, *Mel Frequency Cepstral Coefficients for Music Modeling*, Ismir, Plymouth, MA, USA, 2000.

[13] N. Zerari, S. Abdelhamid, H. Bouzgou, and C. Raymond, "Bidirectional deep architecture for Arabic speech recognition," *Open Computer Science*, vol. 9, no. 1, pp. 92–102, 2019.

[14] Z. Touati-Hamad, M. R. Laouar, and I. Bendib, "Authentication of quran verses sequences using deep learning," in *Proceedings of the 2021 International Conference on Recent Advances in Mathematics and Informatics (ICRAMI)*, pp. 1–4, Tebessa, Algeria, September 2021.

[15] Z. Touati-Hamad, M. Ridda Laouar, I. Bendib, and S. Hakak, "Arabic quran verses authentication using deep learning and word embeddings," *The International Arab Journal of Information Technology*, vol. 19, no. 4, pp. 681–688, 2022.

[16] S. Shahriar and U. Tariq, "Classifying maqams of qur'anic recitations using deep learning," *IEEE Access*, vol. 9, pp. 117271–117281, 2021.

[17] A. Al-Zaben, H. Hamad, A. Alfahoum, and W. Saefan, "Heart rate variability while listening to quran recitation," *Arabian Journal for Science and Engineering*, vol. 39, no. 2, pp. 1129–1133, 2014.

[18] M. Nasef, R. Elbarougy, A. A. E. 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, 2018.

[19] N. F. Hmad, *Deep Neural Network Acoustic Models for Multi-Dialect Arabic Speech Recognition*, Nottingham Trent University (United Kingdom), Nottingham, UK, 2015.

[20] H. Qasim and H. A. Abdulbaqi, "Arabic speech recognition using deep learning methods: literature review," *AIP Conference Proceedings*, vol. 2398, no. 1, 2022.

[21] N. O. Balula, M. Rashwan, and S. Abdou, "Automatic speech recognition (ASR) systems for learning Arabic language and Al-quran recitation: a Review," *International Journal of Computer Science and Mobile Computing*, vol. 10, no. 7, pp. 91–100, 2021.

[22] M. H. Bashir, A. M. Azmi, H. Nawaz et al., "Arabic natural language processing for qur'anic research: a systematic review," *Artificial Intelligence Review*, vol. 56, no. 7, pp. 6801–6854, 2022.

[23] N. J. Ibrahim, Z. Razak, M. Y. I. Idris, Z. Mohd Yusoff, E. M. T. Noorzaily Mohamed Noor, and N. N. A. Rahman, "Mobile application of al-quran and Arabic language for interactive and self learning assistant for support in j-qaf learning: a survey," in *Seminar Kebangsaan Penyelidikan J-QAF*, pp. 205–218, Citeseer, Princeton, NJ, USA, 2013.

[24] A. Keerio, B. Mitra, P. Birch, R. Young, and C. Chatwin, "On preprocessing of speech signals," *World Academy of Science, Engineering and Technology*, vol. 47, pp. 317–323, 2008.

[25] I. Kamarulafizam, S. Hussain Salleh, J. M. Najeb, A. K. Ariff, and A. Chowdhury, "Heart sound analysis using mfcc and time frequency distribution," in *World Congress on Medical Physics and Biomedical Engineering 2006*, pp. 946–949, Springer, Singapore, 2007.

[26] P. Pal Singh and P. Rani, "An approach to extract feature using mfcc," *IOSR Journal of Engineering*, vol. 4, no. 8, pp. 21–25, 2014.

[27] H. Sak, A. Senior, K. Rao, and F. Beaufays, "Fast and accurate recurrent neural network acoustic models for speech recognition," 2015, https://arxiv.org/abs/1507.06947.

[28] A. Shewalkar, D. Nyavanandi, and S. A. Ludwig, "Performance evaluation of deep neural networks applied to speech recognition: rnn, lstm and gru," *Journal of Artificial Intelligence and Soft Computing Research*, vol. 9, no. 4, pp. 235–245, 2019.

[29] Z. Albert, P. Doetsch, V. Paul, R. Schlüter, and H. Ney, "A comprehensive study of deep bidirectional lstm rnns for acoustic modeling in speech recognition," in *Proceedings of the 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 2462–2466, New Orleans, LA, USA, March 2017.

[30] Y. Yu, X. Si, C. Hu, and J. Zhang, "A review of recurrent neural networks: lstm cells and network architectures," *Neural Computation*, vol. 31, no. 7, pp. 1235–1270, 2019.