Volume 11. No 2, May 2024
e-ISSN: 2527-9572 / ISSN: 2354-8924

A Deep learning Approach for Recognizing the Noon Rule for Reciting Holy Quran

Hanaa Mohammed Osman
Computer Science Dept./
College of Computer Science and Mathematics/
University of Mosul, 41002 Mosul, Iraq
hanaosman@uomosul.edu.iq

Ban Sharief Mustafa
Computer Science Dept./
College of Computer Science and Mathematics/
University of Mosul, 41002 Mosul, Iraq
banmustafa66@uomosul.edu.iq

Basim Mahmood
Computer Science Dept./
College of Computer Science and Mathematics/ University of Mosul, 41002 Mosul, Iraq
ICT Research Unit/ Computer Center/ University of Mosul
bmahmood@uomosul.edu.iq

Abstract **– Ahkam Al-Tajweed represents the most precious religious heritage that is in critical need to be preserved and kept for the next generation. This study tackles the challenge of learning Ahkam Al-Tajweed by developing a model that considers one of the rules experienced by early learners in the Holy Quran. The proposed model focuses, specifically, on the "Hukm Al-Noon Al-Mushaddah," which pertains to the proper pronunciation of the letter "Noon" when it is accompanied by a Shaddah symbol in Arabic. By incorporating this rule into the proposed model, learners will benefit the model because it will improve their Tajweed skills and facilitate the learning process for those who do not have access to private tutors or experts. The proposed approach involved three models namely, Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), and Random Forest models in the context of a classification task. The models were evaluated based on their validation accuracy, and the results indicate that the CNN model achieved the highest validation accuracy of 0.8613. The other contribution of this work is collecting a novel dataset for this kind of study. The findings show that the Random Forest model outperformed the other models in terms of accuracy.**

Keywords: Artificial Intelligence, Deep Learning, Quran Recitation, Ahkam Al-Tajweed, Hukm Al-Noon Al-Mushaddah.

Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.

I. Introduction

This section provides a comprehensive description and background about Quran recitation as well as Ahkam Al-Tajweed. It also states the related works of this study and provides the problem considered alongside the contribution of this work.

A. Overview

The Holy Quran is the main sacred book of Islam, composed of 30 chapters and 6236 Verses grouped into 114 groups called "Surahs." Correct pronunciation during recitation is called "Tajweed," and rules must be considered to ensure the correct meaning is delivered. However, there are many issues in teaching Ahkam Al-Tajweed, including the requirement of an expert for private tutoring called "Talqeen," which is not always available [1]. To tackle this issue, researchers have turned to Machine Learning (ML) techniques aiming to develop computerized systems that check the proper application of Ahkam Al-Tajweed based on audio recordings. However, existing systems are limited in the rules they consider or the Quran Verses they cover [2]. Moreover, to detect errors in this rule and other rules, we utilized state-of-the-art deep learning techniques, including Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), to analyze audio recordings of Quran recitations.

Furthermore, one of the remarkable aspects of the Quran is its linguistic structure and the way its Verses are composed. The study of the phonetics of the Quran, also known as the science of Tajweed, involves the analysis of the sounds, rhythm, and melody of the Quranic text [3]. The Quranic text is written in Arabic, a language that is highly complex and rich in its phonetic and linguistic features. The proper recitation of the Quranic text is considered to be of great importance in Islam, and a skilled reciter of the Quran is highly respected and admired within the Muslim community. The phonetics of the Quran are, therefore, an essential part of Islamic scholarship, and the study of Tajweed is highly valued by Muslim scholars and laypeople alike [3]. The phonetics of the Quran are also of interest to linguists and scholars of Arabic language and literature. The rhythmic and melodic patterns of the Quranic text have been the subject of much study and analysis, and the beauty and complexity of these patterns continue to fascinate scholars today. The phonetics of the Quran also offer insights into the development of the Arabic language and its evolution over time [4].

On the other hand, there are several audio feature extraction techniques such as MFCCs, Spectral Contrast, Chroma STFT, and Spectrograms for speech and music analysis. These features have been shown to be effective in capturing the spectral content of audio signals and can provide valuable information for tasks such as speech recognition, music genre classification, and emotion recognition. The features and their effectiveness in improving the performance of the system are described as follows [5]:

- MFCC (Mel-Frequency Cepstral Coefficients): MFCCs are widely used in speech and music analysis. They are computed by first applying a filter bank to the power spectrum of an audio signal to obtain the mel-scaled power spectrum. This is followed by taking the logarithm of the mel-scaled power spectrum and then performing the Discrete Cosine Transform (DCT) on the resulting coefficients. The first few coefficients are typically the most informative, capturing the overall spectral shape of the signal, while higher coefficients capture more detailed spectral information [6].
- Spectral Contrast: it is a feature that captures the differences in energy between adjacent frequency bands in a spectrogram. It is calculated by dividing the spectrum into sub-bands and computing the difference in energy between the highest and lowest frequencies in each sub-band. Spectral Contrast can be useful for speech and music classification, as it can capture the distinctive spectral features of different types of sounds [7].
- Chroma STFT: Chroma features are a way of summarizing the pitch content of an audio signal. Chroma features are computed by first calculating the short-time Fourier transform (STFT) of the signal, and then projecting the magnitude spectrum onto a set of pitch classes. Each pitch class corresponds to a particular musical note, and the value of the chroma feature for each pitch class is the sum of the magnitudes of the corresponding spectral components. Chroma features are often used in music information retrieval tasks, such as genre classification or chord recognition [8].
- Spectrogram: A spectrogram is a visual representation of the frequency content of an audio signal over time. It is computed by dividing the signal into overlapping frames, computing the magnitude of the Fourier transform for each frame, and then plotting the resulting magnitudes as a function of frequency and time. Spectrograms can be useful for visualizing the spectral content of an audio signal, and can also be used as input to machine learning models for tasks such as speech or music classification [9].

B. Literature Review

The literature includes several studies that aimed to develop an intelligent model for recognizing the rules of Holy Quran recitation and tracking reading errors using automatic speech recognition techniques. One of the early studies was of Muhammad et al., [10], who developed an intelligent system called "Hafize" for tracking Tajweed rules. The system recognized the recitation of 10 different reciters, including men, women, and children, and could identify mistakes at both Verse and word levels. However, it has a limitation as it was based on phonetic rules that are not provided. Another study called "Makhraj" was introduced in [11] to make the recitation of the Holy Quran less dependent on expert reciters. The authors used MFCC for feature extraction and tested the system in two modes: one-to-one and one-to-many. The system achieved a 98% accuracy in the one-to-one mode, which is not considered very accurate due to the utilization of a simple matching technique. Moreover, the authors in [12] introduced an intelligent tutoring system for teaching Tajweed. It was evaluated by reciting teachers and students; the results were promising. However, the system was limited to teaching Tafkhim and Tarqiq in Tajweed for the Holy Quran recitation with the Rewaya of Hafs from 'Aasem. Another intelligent recognition model proposed in [13] to recognize Qira'ah from the corresponding Holy Quran acoustic wave. The model was built upon three phases: 1) feature extraction, 20 training the SVM learning model, and 3) recognizing Qira'ah based on the trained model. The SVM-based recognition model achieved a success rate of 96%.

More studies are available in the literature in this area, for instance, the authors in [14] used traditional audio processing techniques for feature extraction and classification on an in-house dataset of thousands of audio recordings covering all occurrences of the rules under consideration in the entire Holy Quran. The work showed how to enhance the classification accuracy to surpass 97.7% by incorporating deep learning techniques. The researchers in [15] proposed a machine learning approach for recognizing the reciter of the Holy Quran. The system achieved excellent accuracy of 97.62% for chapter 18 and 96.7% for chapter 36 using the ANN classifier. In [16], the authors addressed the problem of identifying the correct usage of Ahkam Al-Tajweed in the entire Quran, focusing on eight Ahkam Al-Tajweed faced by early learners of recitation. The results showed that the highest accuracy achieved was 94.4%, which was obtained when bagging was applied to SVM with all features except for the LPC features. Finally, the work proposed in [17] suggested a deep learning model using MFCCs to distinguish between trustworthy and fraudulent reciters of the Qur'an. It compared the deep learning approach to machine learning methods and determined the optimal segment length and number of features. The proposed model achieved high accuracy and outperformed other models, while a future direction includes creating a dataset encompassing the entire Qur'an for further research on recitation rules using deep learning techniques. The data used in the study is available from the corresponding author upon request.

C. Problem Statement and Contributions

As mentioned in the previous section, there are many issues in teaching Ahkam Al-Tajweed, including the requirement of an expert for private tutoring called "Talqeen," who is not always available, which is a critical issue when learning Ahkam Al-Tajweed. According to the literature, there is still a lack of stable and reliable models for Ahkam Al-Tajweed. Hence, this study tries to fill this gap and tackles the challenge of learning Ahkam Al-Tajweed by developing a model that takes into account one of the rules experienced by early learners in the Holy Quran. The proposed model involves three AI models; CNN, RNN and Random Forest, and focuses, specifically, on the "Hukm Al-Noon Al-Mushaddah" which pertains to the proper pronunciation of the letter "Noon" when it is accompanied by a Shaddah symbol in Arabic. By incorporating this rule into the proposed model, learners will benefit the model because it will improve their Tajweed skills and facilitate the learning process for those who do not have access to private tutors or experts.

The proposed model considers many aspects for learners' benefit such as the economic aspect in case a learner cannot find a tutor in the local area and it is difficult for the learner to travel. Also, the other benefit aspect is the case of distance learning. It is important to mention that the development of the proposed model represents the beginning of an integrated learning system that relies on machine learning techniques and can improve and develop itself to distinguish correct recitations. The other contribution of this work is building a novel dataset that has not been available in this literature. The dataset is considered comprehensive and appropriate to be used for research purposes.

The organization of this document is summarized as follows: Section 2 describes the dataset used and the details of the proposed model. Section 3 demonstrates the experimental results that are obtained using the proposed and benchmarking models and Section 4 concludes this work.

II. RESEARCH METHOD

This section describes about the dataset used in this work and the method followed in performing this research. Figure 1 demonstrates the flow of the proposed method step by step.

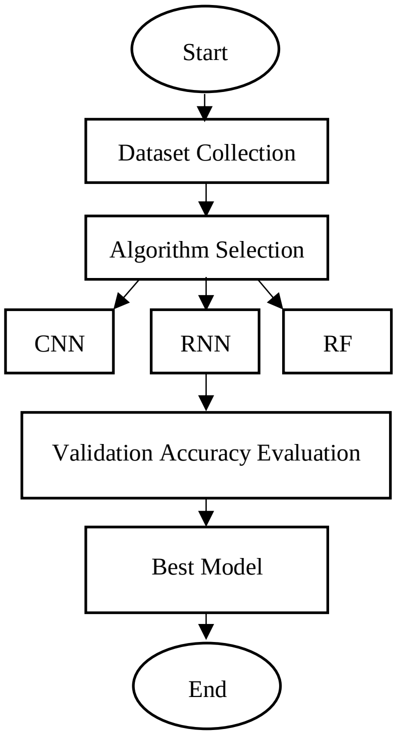


Figure 1: Flowchart of the proposed approach.

A. Dataset Creation

A dataset of over 6000 audio files representing readings of different Verses by volunteers was collected. The dataset was designed to capture the variability in speech patterns and acoustic characteristics across different speakers and Verses and was collected with the aim of building a system for recognizing the correct recitation. Each audio file was annotated with metadata such as the speaker ID, Verse ID, and recording conditions, in order to facilitate analysis and model training. This section describes the process of building the dataset, including data collection, annotation, and preprocessing. A summary statistics and analysis of the dataset were provided as well as highlighting the main characteristics and the potential challenges for model training.

A dataset has been collected containing recitations of Quranic Verses recited according to the rules related to the Hukm Al-Noon Al-Mushaddah, which includes six different Verses. This type of rule includes specific instructions for reciting Al-Noon Al-Mushaddah and presents a challenge for learners to master correct recitation. The Ghunnah must be shown in Hukm Al-Noon Al-Mushaddah when it is pronounced with emphasis through two movements. This ruling is called an “accentuated Ghunnah” letter. The dataset can be used to develop techniques for Quranic recitation based on the rules and to improve the accuracy of the final recitation. The training dataset includes the following Verses:

- {إِنَّكَ أَنتَ عَلَّٰمُ ٱلْغُيُوبِ} [Surah Al-Ma'idah 109]
- {وَٱلنَّٰزِعَٰتِ غَرْقًا}[Surah An-Nāziʿāt: 1]
- {وَٱلنَّٰشِطَٰتِ نَشْطًا}[Surah An-Nāziʿāt: 2]
- {مِن شَرِّ ٱلْوَسْوَاسِ ٱلْخَنَّاسِ}[Surah An-Nās: 4]
- {ٱلَّذِى يُوَسْوِسُ فِى صُدُورِ ٱلنَّاسِ}[Surah An-Nās: 5]
- {مِنَ ٱلْجِنَّةِ وَٱلنَّاسِ}[Surah An-Nās: 6]

And the following Verses in testing dataset:

- {وَلَا ٱلظُّلُمَٰتُ وَلَا ٱلنُّورُ}[Surah Fāṭir: 20]
- {ٱلنَّجْمُ ٱلثَّاقِبُ} [Surah Aṭ-Ṭāriq: 3]
- {عَنِ ٱلنَّبَإِ ٱلْعَظِيمِ}[Surah An-Nabaʾ: 2]
- {وَجَنَّٰتٍ وَعُيُونٍ}[Surah Ash-Shuʿarāʾ: 134]
- {ٱلنَّارِ ذَاتِ ٱلْوَقُودِ}[Surah Al-Burūj: 5]
- {وَٱلنَّٰشِرَٰتِ نَشْرًا}[Surah Al-Mursalāt: 3]

The dataset comprises more than 6000 audio recordings of Quranic recitations. They were performed by more than 750 individuals with varying ages and genders. The dataset is divided into two parts, one for training and the other for testing. The training dataset consists of audio in the form of (.wav) files for six Verses, and the test dataset consists of audio files for another six Verses, all of which have Al-Noon Al-Mushaddada along with a label indicating whether it was recited correctly or not. These recordings were captured using online platforms such as WhatsApp and Telegram channels. The QDAT2 dataset is an updated version of the previously published QDAT dataset available on Kaggle. Each audio sample consists of the recitation of one of the six Verses containing the Al-Noon al-Mushaddada, along with other related features such as age, gender, and accuracy of the recitation of the Ghunnah rule. All audio files are provided in a (.wav) format, and their corresponding links are included in an Excel (.csv) file. The dataset is valuable for analyzing the proper pronunciation and intonation of Al-Noon Al-Mushaddad in Quranic recitation. The readers who participated in the recordings are from the southern regions of Iraq and Syria. The recitation rule of Al-Noon Al-Mushaddad in Quranic recitation is commonly known as (Al-Ghunnah). This rule pertains to the correct pronunciation of the letter (Noon) in Arabic when it appears with a Shadda, and is considered a crucial aspect of Quranic recitation. Regarding the Phonological distribution, Table 1 presents the provisions of the Verse in our dataset.

Table 1: The provisions of the Verse

Provisions	Verse	Verse in symbols
Al_Gunna	(وَلَا ٱلظُّلُمَٰتُ وَلَا ٱلنُّورُ)	وَ-لَ-ظُ-لـُمـِـِت-و-ـِل-ـِن- نـُـُر
Al_Gunna	(ٱلنَّجْمُ ٱلثَّاقِبُ)	ءـِ-ن-ـِج-مـُ-ث-ثـِـِق-ب
Al_Gunna	(عَنِ ٱلنَّبَإِ ٱلْعَظِيمِ)	عـِ-ن-ِ-ن-ـِب-ـِء-ِ-ل-ع-ـِ ظ-ِ-ِ-م
Al_Gunna	(وَجَنَّٰتٍ وَعُيُونٍ)	و-ـِج-ـِن-ن-ـِـِت-ِ-ن-و-ـِ-ع-ـُ يـُـُن
Al_Gunna	(ٱلنَّارِ ذَاتِ ٱلْوَقُودِ)	ءـِ-ن-ن-ـِـِر-ِ-ذ-ـِـِت-ِ-ل-و-ـِ قـُـُد
Al_Gunna	(وَٱلنَّٰشِرَٰتِ نَشْرًا)	و ـِ-ن-ن-ـِـِش-ِ-ر-ـِـِت-ِ-ن-ـِ ش-ر-ـِـِ

Al_Gunna	(وَٱلنَّٰزِعَٰتِ غَرْقًا)	و-ـِ-ن-ن-ـِـِز-ِ-ع-ـِـِت-غ-ـِر-ق-ـِ ـِ
Al_Gunna	(وَٱلنَّٰشِطَٰتِ نَشْطًا)	و-ـِ-ن-ن-ـِـِش-ِ-ط-ـِـِت-ِ-ن-ـِ ش-ط-ـِـِ
Al_Gunna	(مِن شَرِّ ٱلْوَسْوَاسِ ٱلْخَنَّاسِ)	م-ِ-ن-ش-ـِر-ِ-ل-و-ـِس-و-ـِـِ س-ِ-ل-خ-ـِن-ن-ـِـِس
Al_Gunna	(ٱلَّذِى يُوَسْوِسُ فِى صُدُورِ ٱلنَّاسِ)	ءـِ-ل-ل-ـِذ-ِ-ِ-يـُ-وـِ-س-و-ِ- سـُ-ف-ِ-ِ-صـُ-د ـُـُر -ِ-ن-ن-ـِ ـِس
Al_Gunna	(مِنَ ٱلْجِنَّةِ وَٱلنَّاسِ)	م-ِ-ن-ـِل-ج-ِ-ن-ن-ـِت-ِ-و-ـِ ن-ن-ـِـِس

B. Proposed Model

Several models were experimented using different features and input representations. Specifically, the models trained CNN [18], RNN [19], and Random Forest with K-Fold training model [20]. The use of the CNN model started by converting the audio files into spectrogram image files. For the RNN model, a neural network architecture was used with multiple layers. The model was trained using a batch size of 32 and a learning rate of 0.001. For the Random Forest model, K-Fold cross-validation with a k value of 10 was applied. Also, different kernel functions were experimented including radial basis function (RBF), polynomial, and linear. For the CNN model, the audio files were converted into spectrogram image files using the Mel spectrogram representation. Additionally, a neural network architecture was used with multiple convolutional and pooling layers, followed by dense layers. The model was trained using a batch size of 64 and a learning rate of 0.0001.

III. Results and Discussions

Table 1 summarizes the accuracy results for each model on the test set. As can be seen, the CNN model achieved the highest accuracy, followed by RNN model and the Random Forest model. Overall, our results suggest that CNN models with Mel spectrogram features are effective for recognizing the correct recitation of Quranic Verses with the "Al Gunna" rule.

Table 2: Accuracy Results for Different Models.

Model Type	Accuracy	Validation Accuracy
CNN	0.9117	0.8613
RNN	0.8749	0.8121
Random Forest	----------	0.8085

In current practice, predictive models are deployed to estimate the accuracy of a given test set, which was not utilized in the model training or validation process. In this context, we have employed six distinct Verses, each characterized by the Al Gunna recitation rule. However, due to the phonetic variability of the Noon sound across different Verses, it is observed that the test accuracy varies between them. Specifically, the Verse3 and 4 from the test set exhibit a similar phonetic distribution and are therefore considered to be the most closely related. Table 2 shows the Verses test accuracy for the different models. This table shows the accuracy of three different models (CNN, RNN, and Random Forest) for each Verse in the test dataset.

Furthermore, the accuracy of the RNN model on both the training and test data is depicted in Figure 2 (A and B). Figure 3 shows the accuracy and the evaluation of the accuracy of the CNN model. On the other hand, Figure 4 demonstrates the accuracy of the SVM model using K-Fold Cross-Validation. In addition, Table 3 shows the values of implementing the Random Forest.

According to the aforementioned results, the accuracy of three different models (CNN, RNN, and Random Forest) for each Verse in the test dataset represents evidence that the Random Forest model outperformed the other models since it achieved the highest accuracy for Verse 3 and Verse 4 with scores of 0.7063 and 0.6956, respectively. However, for Verse 1, Verse 2, and Verse 6, the CNN model outperformed both RNN and Random Forest models, with accuracy scores of 52.0, 55.47, and 61.67, respectively. On the other hand, the RNN model achieved the highest accuracy score for Verse 4 with a value of 0.6782.

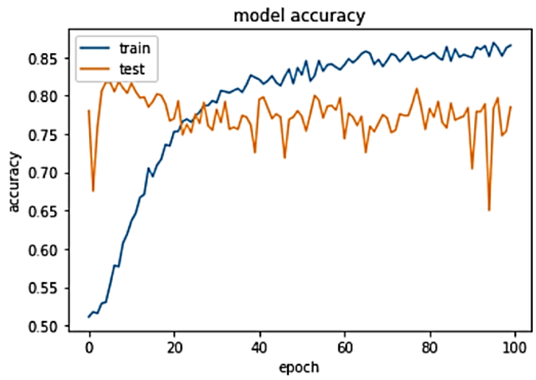
The results suggest that the performance of the models varies across different Verses, indicating that the phonetic distribution of Noon sound plays a crucial role in determining the accuracy of the models. As we have discussed earlier, the phonetic variability of Noon sound within each Verse needs high attention when evaluating the performance of predictive models in Arabic recitation. The findings also highlight the importance of carefully selecting test Verses to ensure that they are representative of the broader phonetic variation in the recitation rules being modeled. Finally, the results provide insights into the challenges of modeling Arabic recitation accurately.

Table 3: Accuracy of the Verses using each of the three models used.

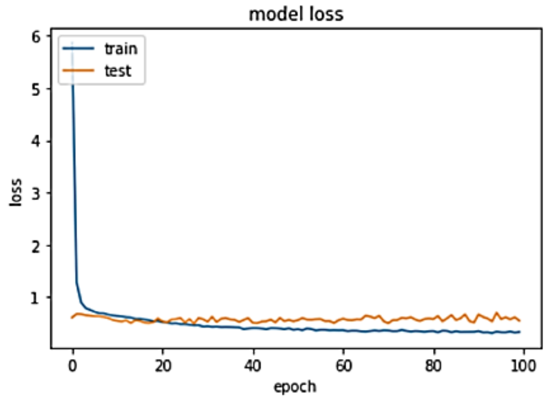
Model	Verse 1	Verse 2	Verse 3	Verse 4	Verse 5	Verse 6
CNN	52.0	55.47	63.49	52.17	69.35	61.67
RNN	0.59	0.625	0.7063	0.6956	0.6854	0.6000
Random Forest	0.6016	0.6328	0.6507	0.6782	0.6129	0.6105

Table 3: Random Forest validation details.

Test accuracy	Validation accuracy	Validation Loss
0.6119402985074627	0.8303519595120427	5.859523780066005

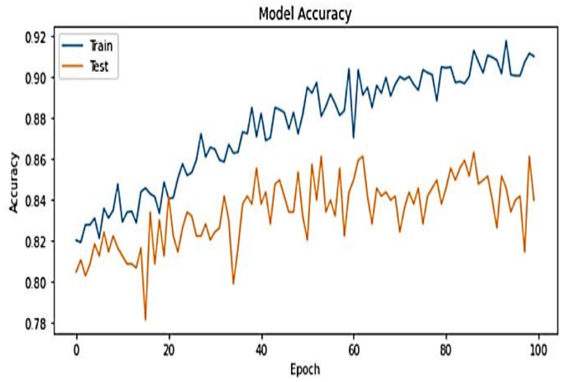


(A)

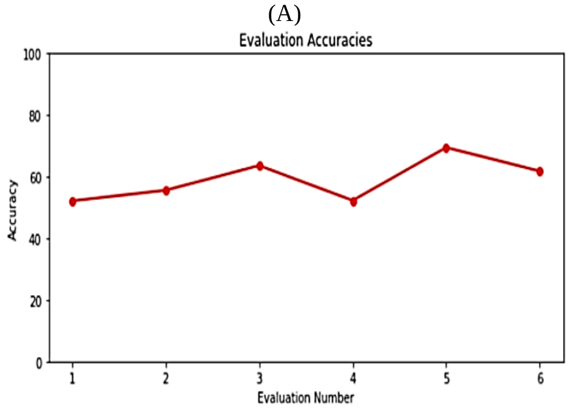


(B)

Figure 2: Random Forest model on both the training and testing data where (A) is the accuracy and (B) is the loss.



(A)



(B)

Figure 3: (A) CNN model accuracy, and (B) evaluation accuracy.

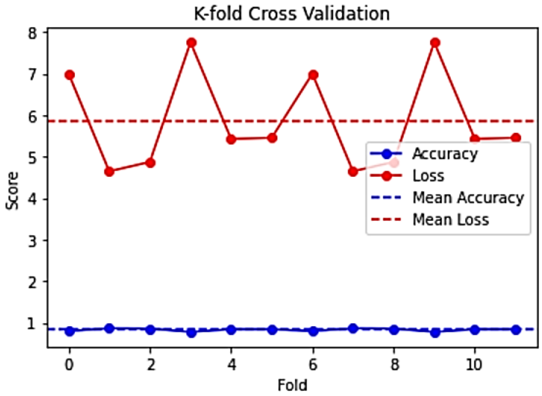


Figure 3: Accuracy of the SVM model on both the training and test data.

According to the obtained results in Figures 2, 3, and 4, it can be observed that CNN is better for achieving the purpose of this work. While RNN that is used for sequential data processing underperformed CNN. Also, since Random Forest is used for classification and regression tasks it is also underperformed CNN. This is because CNN uses convolutional layers to efficiently extract features from data, while RNN uses recurrent layers to maintain a memory of the previous inputs and outputs. On the other hand, Random Forest uses decision trees to make predictions. However, it is observed that Random Forest is a simpler algorithm that can be used for smaller datasets or when computational resources are limited [21-22].

IV. Conclusions

This work tried to address the issue of learning Ahkam Al-Tajweed by developing a model that considered one of the rules experienced by early learners in the Holy Quran. The proposed model focuses, specifically, on the "Hukm Al-Noon Al-Mushaddah," which pertains to the proper pronunciation of the letter "Noon" when it is accompanied by a Shaddah symbol in Arabic. By incorporating this rule into the proposed model, learners will benefit the model because it will improve their Tajweed skills and facilitate the learning process for those who do not have access to private tutors or experts. The proposed approach involved three AI models namely, CNN, RNN, and Random Forest. The other contribution of this work is collecting a novel dataset for this kind of study. The collected data was used by the three models. The findings show that the CNN model outperformed the other models in terms of validation accuracy. Also, the test accuracy varies from verse to verse in different models, some test verse showed promised result against others. Finally, this study is ongoing and will continue until it covers all the available aspects of Ahkam Al-Tajweed.

Furthermore, this work is considered an approach for cultural and religious heritage preservation, which contributes to having sustainable communities as the United Nations declares its goals in promoting our communities to be sustainable. Future works can build upon these findings to develop more effective predictive models that account for the phonetic variability of Arabic recitation, ultimately facilitating the preservation and transmission of this rich cultural heritage.

Acknowledgment

We would like to appreciate the support of the Computer Science Department/ College of Computer Science and Mathematics/ University of Mosul and the ICT Research Unit/ Computer Center/ University of Mosul, Iraq. We are also grateful to the participants who volunteered to help us in collecting the dataset.

Data Availability

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

References

[1] Samara G, Al-Daoud E, Swerki N, Alzu'bi D. The Recognition of Holy Qur'an Reciters Using the MFCCs' Technique and Deep Learning. Advances in Multimedia. 2023 Mar 21;2023. https://doi.org/10.1155/2023/2642558

[2] Al-Ayyoub M, Damer NA, Hmeidi I. Using deep learning for automatically determining correct application of basic quranic recitation rules. Int. Arab J. Inf. Technol.. 2018 Apr;15(3A):620-5.

[3] Nasallah MK. The Importance Of Tajweed In The Recitation Of The Glorious Qur'an: Emphasizing Its Uniqueness As A Channel Of Communication Between Creator And Creations. IOSR Journal Of Humanities And Social Science (IOSR-JHSS). 2016;21(2):55-61.

[4] Ahmad M. Literary Miracle of the Quran. Ar-Raniry: International Journal of Islamic Studies. 2020 Jul 28;3(1):205-20.

[5] Nogueira AF, Oliveira HS, Machado JJ, Tavares JM. Sound Classification and Processing of Urban Environments: A Systematic Literature Review. Sensors. 2022 Nov 8;22(22):8608. doi: 10.3390/s22228608

[6] Ayvaz, Uğur, et al. "Automatic speaker recognition using mel-frequency cepstral coefficients through machine learning." *CMC-Computers Materials & Continua* 71.3 (2022).

[7] Su, Yu, et al. "Performance analysis of multiple aggregated acoustic features for environment sound classification." *Applied Acoustics* 158 (2020): 107050.

[8] Kumar, Nagendra, et al. "CNN based approach for Speech Emotion Recognition Using MFCC, Croma and STFT Hand-crafted features." *2021 3rd International Conference on Advances in Computing, Communication Control and Networking (ICAC3N)*. IEEE, 2021.

[9] Gong, Yuan, et al. "Ssast: Self-supervised audio spectrogram transformer." *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 36. No. 10. 2022.

[10] Muhammad A, ul Qayyum Z, Tanveer S, Martinez-Enriquez A, Syed AZ. E-hafiz: Intelligent system to help muslims in recitation and memorization of Quran. Life Science Journal. 2012 Oct;9(1):534-41.
[11] Akkila AN, Abu-Naser SS. Teaching the right letter pronunciation in reciting the holy Quran using intelligent tutoring system.
[12] Nahar KM, Al-Khatib RM, Al-Shannaq MA, Barhoush MM. An efficient holy Quran recitation recognizer based on SVM learning model. Jordanian Journal of Computers and Information Technology (JJCIT). 2020 Dec 1;6(04):394-414.
[13] Nahar KM, Al-Khatib RM, Al-Shannaq MA, Barhoush MM. An efficient holy Quran recitation recognizer based on SVM learning model. Jordanian Journal of Computers and Information Technology (JJCIT). 2020 Dec 1;6(04):394-414. doi: 10.5455/jjcit.71-1593380662
[14] Al-Ayyoub M, Damer NA, Hmeidi I. Using deep learning for automatically determining correct application of basic quranic recitation rules. Int. Arab J. Inf. Technol.. 2018 Apr;15(3A):620-5.
[15] Alkhateeb JH. A machine learning approach for recognizing the Holy Quran reciter. International Journal of Advanced Computer Science and Applications. 2020;11(7).
[16] Damer NA, Al-Ayyoub M, Hmeidi I. Automatically determining correct application of basic quranic recitation rules. InProceedings of the International Arab Conference on Information Technology 2017 Dec 22.
[17] Samara G, Al-Daoud E, Swerki N, Alzu'bi D. The Recognition of Holy Qur'an Reciters Using the MFCCs' Technique and Deep Learning. Advances in Multimedia. 2023 Mar 21;2023.
[18] Demir, Fatih, Daban Abdulsalam Abdullah, and Abdulkadir Sengur. "A new deep CNN model for environmental sound classification." *IEEE Access* 8 (2020): 66529-66537.
[19] Ou, Zhenyi, Ke Qu, and Chen Liu. "Estimation of sound speed profiles using a random forest model with satellite surface observations." *Shock and Vibration* 2022 (2022).
[20] Park, Ji Soo, et al. "A machine learning approach to the development and prospective evaluation of a pediatric lung sound classification model." *Scientific Reports* 13.1 (2023): 1289.
[21] Al-Ameri, Mohammed Abdulbasit Ali, et al. "Unsupervised Forgery Detection of Documents: A Network-Inspired Approach." *Electronics* 12.7 (2023): 1682.
[22] Sultan, Ahmed, and Basim Mahmood. "A Novel Network Model for Artificial Intelligence Algorithms." *2022 8th International Conference on Contemporary Information Technology and Mathematics (ICCITM)*. IEEE, 2022.