Indonesian Journal of Electrical Engineering and Computer Science
Vol. 39, No. 3, September 2025, pp. 1562~1570
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v39.i3.pp1562-1570

Enhancing Qur'anic recitation through machine learning: a predictive approach to Tajweed optimization

Mohamed Amine Daoud[1], Nayla Fatima Hadjar Kherfan[2], Abdelkader Bouguessa[1], Sid Ahmed Mokhtar Mostefaoui[1]
[1]LRIAS Laboratory, Department of Computer Siences, Ibn-Khaldoun University of Tiaret, Tiaret, Algeria
[2]Department of Computer Siences, Ibn-Khaldoun University of Tiaret, Tiaret, Algeria

Article Info

Article history:

Received Feb 11, 2025
Revised Apr 6, 2025
Accepted Jul 2, 2025

Keywords:

Machine learning
Optimization
Predictive
Qur'an
Recitation

ABSTRACT

The human voice is a powerful medium for conveying emotion, identity, and intellect. Arabic, as the language of the Qur'an, holds deep spiritual and linguistic importance. Reciting the Qur'an correctly involves following Tajweed rules, which ensure phonetic precision and aesthetic quality. However, mastering these rules is challenging due to complex pronunciation and articulation variations, often requiring expert guidance. Traditional learning methods lack personalized feedback, making it difficult for learners to identify and correct errors. With the rise of machine learning, new opportunities have emerged to support Qur'anic recitation through intelligent analysis of Tajweed patterns and error prediction. This study presents a predictive model that identifies Qur'an reciters using ensemble learning techniques. By incorporating deep learning models like gated recurrent units (GRUs), long short-term memory (LSTM), and recurrent neural network (RNN), the system effectively captures the vocal features unique to each reciter. The model achieves an accuracy rate of 88.57%, demonstrating its potential to support Qur'anic learning and preservation. Nonetheless, its performance may be affected by audio quality and limited training data diversity. To improve adaptability and robustness, future work will focus on enriching the dataset and optimizing the model to generalize better across a broader range of reciters.

Corresponding Author:

Mohamed Amine Daoud
LRIAS Laboratory, Department of Computer Sciences, Ibn-Khaldoun University of Tiaret
Tiaret Algeria
Email: mohamedamine.daoud@univ-tiaret.dz or m_daoud@esi.dz

1. INTRODUCTION

The human voice is a fundamental tool for communication, expression, and identity, shaping civilizations through speech, storytelling, and music [1]. Among global languages, Arabic holds a distinct cultural and spiritual significance, spoken by approximately 400 million people natively and serving as the liturgical language for over 1.5 billion muslims worldwide [2]. The Qur'an, revealed over fourteen centuries ago, is not merely a religious text but a source of divine guidance and spiritual reflection for muslims. Its recitation follows Tajweed, a structured set of phonetic rules ensuring proper pronunciation and rhythm [3]. The Ahkam Al-Tajweed [4] governs these rules, preserving the Qur'an's oral tradition through the practice of Talqeen, which relies on expert instruction. However, mastering Tajweed remains challenging due to its complexity and the need for skilled teachers [5].

The recitation of the Qur'an holds immense spiritual and cultural significance for muslims worldwide. Central to this practice is Tajweed, a precise set of phonetic rules that ensure the correct

pronunciation of Qur'anic verses. Mastering Tajweed requires a deep understanding of articulation points (*makhraj*) [6], phonetic characteristics (*sifat*), and recitation rules such as elongation (*madd*), merging (*idgham*) [7], conversion (*iqlab*) [8], and pauses (*waqf*) [9]. However, traditional methods of Tajweed instruction rely heavily on oral transmission (*Talqeen*), where learners receive direct feedback from qualified teachers. While this approach is effective, it is often constrained by limited accessibility to expert instructors and the need for extensive practice.

In recent years, technological advancements, particularly in machine learning, have created new opportunities for enhancing Tajweed learning. A key challenge remains the lack of accessible, tailored solutions for learners, whether through human or machine-assisted tutoring, to effectively master this art. Innovations in machine learning provide powerful tools to bridge this gap. These technologies analyze complex sound patterns, predict recitation errors, and personalize learning experiences, blending traditional instruction with contemporary needs. Advanced neural network architectures [10], such as convolutional neural networks (CNNs) [11], [12], long short-term memory (LSTM) networks [13], [14], and gated recurrent units (GRUs) [15], are pivotal in this progress. Ensemble learning approaches, which combine multiple architectures—including CNNs, recurrent neural networks (RNNs), and Transformers—further enhance the robustness and accuracy of speech recognition systems. Together, these technologies form the foundation of our proposed solution, enabling efficient analysis of Tajweed recitation and addressing the limitations of traditional methods.

In this work, we propose an innovative approach to facilitate Qur'anic recitation learning by leveraging advanced machine learning techniques. Our system utilizes deep learning models, including CNNs, LSTM networks, and GRUs, to analyze Tajweed recitation with high precision. By integrating ensemble learning methods [16], which combine multiple neural architectures, we aim to enhance the robustness of recitation error detection. The ultimate goal is to provide an accessible, technology-driven solution that complements traditional Tajweed instruction, achieving a performance benchmark of approximately 88.57%.

The structure of this paper is organized as follows: section 2 provides a comprehensive literature review, highlighting prior work and advancements in Qur'anic recitation analysis. Section 3 details the dataset utilized in this study, including its preparation and characteristics, and elaborates on the architecture of the proposed predictive model. Section 4 presents the experimental results, comparing the performance of the proposed model with benchmarking techniques to validate its effectiveness. Finally, section 5 concludes the paper, summarizing the findings and discussing potential directions for future research.

2. LITERATURE REVIEW

The integration of machine learning techniques in Qur'anic recitation analysis has significantly advanced the accuracy, accessibility, and overall learning experience, particularly in applying Tajweed rules and detecting recitation errors. Various studies have explored deep learning models for improving Qur'anic recitation [17], yet challenges remain in dataset availability, model robustness, and practical applicability.

Moustafa and Aly [18], the authors propose a deep learning model for Arabic speaker identification using Wav2Vec2.0 and HuBERT, two state-of-the-art self-supervised learning models for speech representation. Wav2Vec2.0 utilizes a transformer-based architecture with masked feature prediction, while HuBERT employs clustering-based pre-training. A multi-layer perceptron (MLP) classifier is used for speaker classification, achieving an impressive accuracy of 98% with Wav2Vec2.0 and 97.1% with HuBERT. The study acknowledges limitations related to dataset specificity, model complexity, accuracy variability, class diversity, and dependence on pre-trained models, which should be addressed in future work.

Osman *et al.* [19] focuses on evaluating the effectiveness of the QDAT dataset for Qur'anic recitation analysis. The proposed system comprises two main stages: feature extraction using mel-frequency cepstral coefficients (MFCC) and classification through machine learning techniques. While the system demonstrates notable improvements in Qur'anic education through automation, it still faces challenges in capturing the contextual nuances of recitations, as automated models may struggle with the complex phonetic and linguistic aspects of Tajweed.

Samara *et al.* [20], a deep learning model utilizing CNNs was introduced for the classification of recitations from seven well-known Qur'anic reciters. The study employed MFCCs for feature extraction and conducted a comparative analysis between CNN-based classification and traditional machine learning approaches. The CNN model demonstrated exceptional performance, achieving an accuracy of 99.66%, surpassing the SVM, which reached 99%. Despite these promising results, challenges remain regarding dialectal variations and the robustness of the model across diverse reciters and pronunciation styles.

A more comprehensive approach to recitation assessment is presented in [21], where a deep learning system is designed to verify the correctness of individual letters, words, and full verses in Qur'anic recitation. This model integrates MFCC-based feature extraction with LSTM networks, a type of RNN, to

classify recitations. The model achieves an accuracy of 97.7%, surpassing previous methods. However, further improvements are needed to refine pronunciation feedback and expand dataset diversity.

Harere and Jallad [22], the authors introduce an end-to-end deep learning model for Qur'anic recitation recognition, incorporating a CNN-Bidirectional GRU encoder trained using the connectionist temporal classification (CTC) objective function. The model utilizes the Ar-DAD dataset, which contains recitations from 30 individuals across 37 chapters, achieving a word error rate (WER) of 8.34% and a character error rate (CER) of 2.42%. Despite its effectiveness, the study highlights the need for improved Tajweed rule integration and broader coverage of diverse Qur'anic recitation styles.

A study on Tajweed error detection is presented in [23], where the authors propose a hybrid approach combining MFCC features with LSTM networks to identify mispronunciations. Experiments conducted on the QDAT public dataset report high accuracies of 96%, 95%, and 96% for the detection of Separate Stretching, Tight Noon, and Hide rules, respectively. However, a key limitation of this work is its reliance on machine learning algorithms without benchmarking against deep learning models, as well as the use of a private dataset, restricting reproducibility and comparison with existing research.

Al-Fadhli *et al.* [24], the authors explore the use of mobile applications to assist learners in correcting their Qur'anic recitation. These applications leverage speech recognition models, yet challenges persist due to the high demand for robust and error-free speech processing systems. The study identifies major gaps in dataset comprehensiveness and the reliance on online services, which may limit accessibility and application scalability.

Finally, Salameh *et al.* [25] addresses the unique challenges faced by non-Arabic speakers in learning Qur'anic recitation. The authors developed a crowdsourcing platform integrated into a mobile application to collect and annotate approximately 7,000 recitations from 1,287 participants across 11 non-Arabic-speaking countries. The dataset achieved a crowd accuracy of 0.77 and an inter-rater agreement of 0.63, demonstrating its value for training AI models. However, the study highlights the limited availability of public Qur'anic datasets for benchmarking, which remains a significant barrier to advancing research in this domain.

In our study, we present a predictive model for identifying Qur'an reciters using ensemble learning. By combining deep learning techniques like GRU, LSTM, and RNN, the system captures distinctive vocal features. It achieves 88.57% accuracy, aiding in the preservation of Qur'anic recitation. However, audio quality and dataset diversity may affect generalization. Future work aims to enhance robustness and expand data coverage. Table 1 presents a comparative analysis of recent studies on speech recognition, speaker identification, and audio signal classification for Qur'an recitation. This analysis highlights a diversity of datasets, methodologies, and feature extraction techniques, with varying performances depending on the objectives of the studies.

Table 1. Comparative analysis of existing studies

Paper	Year	Dataset	Methodology	Feature extraction	Focus areas	Performance
[18]	2021	arbitrary audio signals.	Multi-Layer Perceptron	Wav2Vec2.0	Identifying Arabic speakers	Accuracy 98%
[19]	2021	QDAT public dataset	Gradient Boosting	MFCC	Classification	Accuracy 90.37%
[20]	2023	seven well-known reciters	CNN	MFCC	/	Accuracy 99%
[21]	2023	Al-Qur'an dataset	RNN, LSTM	MFCC	Correct recitations	Accuracy 97.7%
[22]	2023	Ar-DAD public	CNN-Bidirectional GRU	/	Error rate	Word Error Rate 8.34%, Character Error Rate 2.42%.
[23]	2023	QDAT public dataset	LSTM	MFCC	Detect mispron-unciations	Accuracy 96%
[25]	2024	Mobile application	Matthews correlation coefficient	/	Prediction reciters	Accuracy 77%

3. PROPOSED METHOD

The methodology proposed in this work is designed to analyze and classify Qur'anic recitations based on Tajweed rules. The overall process is illustrated in Figure 1 and carefully structured to ensure accuracy, robustness, and generalizability. The approach consists of two main layers: a preprocessing layer and a hybrid model layer, each playing a crucial role in achieving the desired classification outcomes. The preprocessing layer consists of multiple components designed to transform raw audio recordings into

structured data suitable for analysis. The process begins with the MFCC module [24], [25], which extracts essential audio features by mimicking the human auditory system. This technique captures phonetic nuances while minimizing background noise, ensuring high-quality feature extraction. The extracted features are then processed through the embedding module, which converts Tajweed rule annotations into numerical representations, making them compatible with machine learning algorithms. Finally, the postprocessing module aligns the extracted audio features with their corresponding Tajweed annotations, ensuring consistency and coherence for effective classification.

The hybrid model layer leverages the strengths of three advanced deep learning architectures: CNN, LSTM, and GRU. Each model is optimized to address specific aspects of audio data processing. CNN is employed to extract hierarchical local features using multiple convolutional layers, max-pooling for dimensionality reduction, and fully connected layers for feature refinement. LSTM specializes in capturing long-term temporal dependencies by processing data across sequential time steps. Meanwhile, GRU, a simplified yet powerful recurrent neural network, complements CNN and LSTM by efficiently handling temporal relationships while reducing computational complexity. To enhance classification accuracy, an ensemble learning approach is applied, combining the predictions of CNN, LSTM, and GRU. The outputs from these models are aggregated using logistic regression, where the final class is determined by averaging probability scores and selecting the highest value. This strategy leverages the complementary strengths of the individual models, resulting in a more robust and accurate classification system. The ensemble method ensures a balanced prediction capability, overcoming the limitations of standalone models and improving overall system performance.

The QDAT dataset serves as the primary resource for training and evaluating the proposed system. It contains over 1,500 expert-annotated audio recordings used to assess Qur'anic recitation accuracy based on three specific Tajweed rules. The recordings, collected via WhatsApp from more than 150 reciters (350 males and 1,159 females), are stored in WAV format with an 11 kHz sample rate, mono channel, and 16-bit resolution. Additional metadata, including age, gender, and adherence to Tajweed rules, is provided in a CSV file. To further validate the model's generalizability, an external dataset, referred to as "Our Dataset," was collected. This dataset includes recordings from 189 reciters, carefully selected to represent a diverse range of recitation styles and Tajweed adherence, ensuring an unbiased evaluation.

By integrating ensemble learning with advanced deep learning architectures, the proposed system effectively enhances classification accuracy and robustness, leading to a more reliable analysis of Qur'anic recitation. The combination of sophisticated preprocessing techniques, state-of-the-art deep learning models, and ensemble learning strategies ensures a highly accurate and resilient classification framework. This innovative approach provides a valuable tool for improving Tajweed education and refining Qur'anic recitation analysis, offering greater precision and reliability as shown in Table 2.

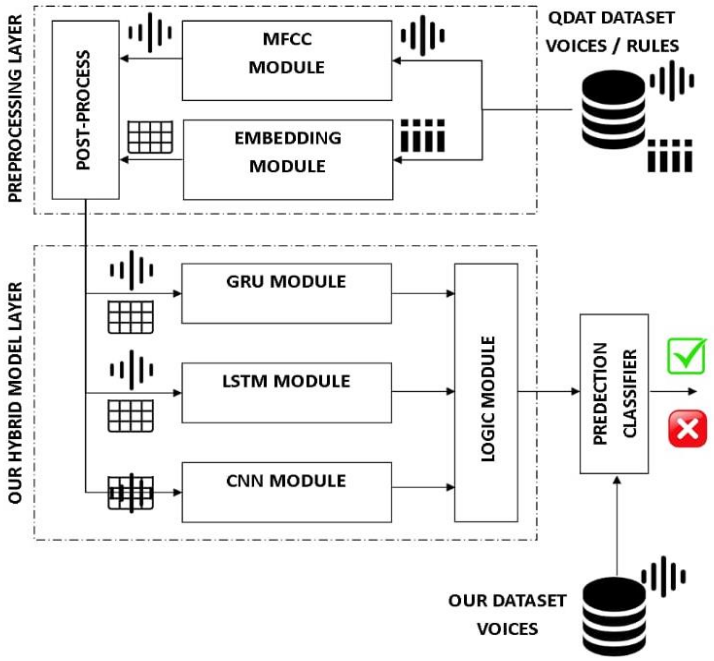

Figure 1. The workflow of the proposed methodology

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

Table 2. The provisions of the Verse

The provisions	Verse	Verse in symbols
Separate stretching of four movements	قَالُوا لَا عِلْمَ لَنَا إِنَّكَ أَنتَ عَلَّامُ ٱلْغُيُوب	قَـــَالُـــوا لَـــا عِـــلْـــمَ لَـــنَـــا إِنَّـــكَ أَنـــتَ عَـــلَّامُ ٱلْـــغُـــيُـــوب
Separate stretching of five movements	قَالُوا لَا عِلْمَ لَنَا إِنَّكَ أَنتَ عَلَّامُ ٱلْغُيُوب	قَـــَالُـــوا لَـــا عِـــلْـــمَ لَـــنَـــا إِنَّـــكَ أَنـــتَ عَـــلَّامُ ٱلْـــغُـــيُـــوب
Tight Noon	قَالُوا لَا عِلْمَ لَنَا إِنَّكَ أَنتَ عَلَّامُ ٱلْغُيُوب	قَـــَالُـــوا لَـــا عِـــلْـــمَ لَـــنَـــا إِنَّـــكَ أَنـــتَ عَـــلَّامُ ٱلْـــغُـــيُـــوب
Hide	قَالُوا لَا عِلْمَ لَنَا إِنَّكَ أَنتَ عَلَّامُ ٱلْغُيُوب	قَـــَالُـــوا لَـــا عِـــلْـــمَ لَـــنَـــا إِنَّـــكَ أَنـــتَ عَـــلَّامُ ٱلْـــغُـــيُـــوب

4. RESULTS AND DISCUSSION

4.1. Results

The results demonstrate that the ensemble learning model, combining RNN, LSTM, and GRU, outperforms individual models across all evaluation metrics—accuracy, recall, and precision. Table 3 presents a detailed comparison of the performance metrics for each model. Below is a detailed discussion of the performances:

a) CNN: the CNN model, while less sophisticated than LSTM, still delivered competitive results with an accuracy of 75.9%, recall of 77.99%, and precision of 79.45% as shown in Figure 2. These limitations underscore the importance of more advanced models for sequence learning. The CNN model shows signs of overfitting, with training accuracy reaching 85% while validation accuracy fluctuates around 75%. A widening gap after epoch 10 suggests poor generalization. Additionally, high variability in validation accuracy indicates instability, possibly due to data imbalance or sensitivity to phonetic variations. The model may struggle with misclassifications in similar sounds.

b) LSTM: the LSTM model exhibited the highest individual accuracy of 84.3%, recall of 87.31%, and a strong precision of 89.1%. This suggests that LSTM effectively captures long-term dependencies, making it more adept at modeling the temporal features of the dataset as shown in Figure 3. The fluctuations in validation accuracy imply some instability, possibly due to dataset variability or imbalanced classes. The model may struggle with sequences containing similar phonetic patterns, leading to misclassifications in certain Tajweed rules.

c) GRU: the GRU model achieved an accuracy of 77.03%, with recall and precision values of 79.23% and 80.45%, respectively. This could be attributed to its simplified architecture, which may be less effective for complex temporal datasets as shown in Figure 4. The training and validation curves suggest good generalization, with minor overfitting. However, fluctuations in validation accuracy indicate some instability, possibly due to dataset imbalances or insufficient contextual learning.

d) Ensemble learning: the ensemble model, which integrates predictions from RNN, LSTM, and GRU, achieved superior performance across all metrics. The accuracy reached 88.57%, marking a significant improvement over the best individual model (LSTM at 84.3%). Furthermore, the recall was 90.56%, reflecting enhanced sensitivity in identifying true positives. Lastly, the precision of 94.11% indicated exceptional specificity in correctly identifying relevant instances. These results demonstrate the effectiveness of ensemble learning in combining the strengths of individual models to achieve better generalization and robustness.

Table 3. Performance metrics results for different models

Model	Accuracy	Recall	Precision	F1-Score	Learning time	Error rate
GRU	77.03	79.23	80.45	83.78	25-35 s/ epoch	8.90%
LSTM	84.3	87.31	89.1	84.72	30-46 s/ epoch	17.54%
CNN	75.9	77.99	79.45	87.41	6-7 s/epoch	14.24%
Ens. learning	88.57	90.56	94.11	86.80	60s/epoch	6.24%

In Figure 5, the learning curve shows the evolution of the accuracy of the ensemble model for Qur'an recitation classification on the training and validation sets over 20 epochs. The performance reaches

around **90%** for training and **87%** for validation, indicating a good generalization capability of the model. The rapid convergence of the curves after 10 epochs demonstrates an efficient stabilization of the learning process, with no significant signs of overfitting.

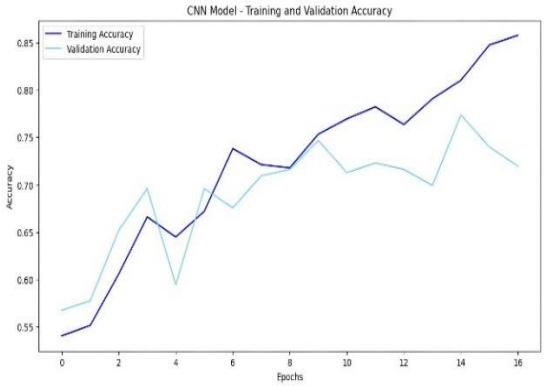

Figure 2. Accuracy of CNN

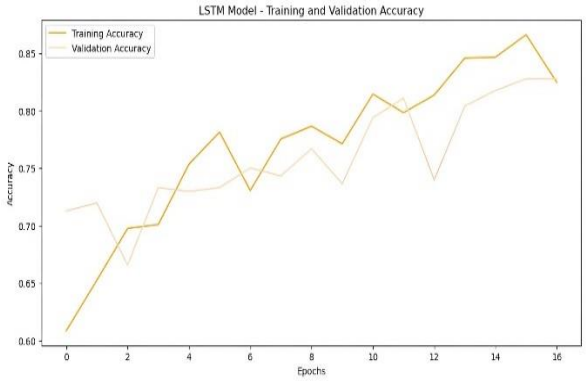

Figure 3. Accuracy of LSTM

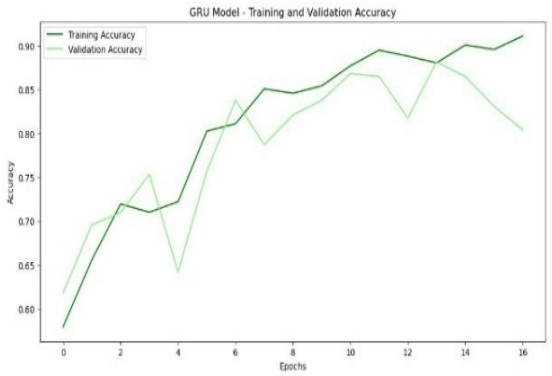

Figure 4. Accuracy of GRU

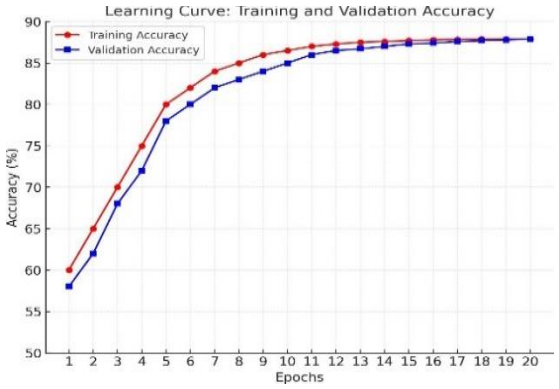

Figure 5. Learning curve of proposed model

4.2. Discussions

Ensemble learning offers improved robustness by combining the strengths of individual models and mitigating their weaknesses. For instance, GRU's relative simplicity complements LSTM's ability to handle long-term dependencies. Meanwhile, RNN contributes by effectively capturing simpler patterns, making the overall system more robust. Additionally, the ensemble approach aggregates diverse decision-making processes, leading to a more balanced and generalized model. As a result, the model performs well across different data patterns, reducing bias and variance. Furthermore, the ensemble model demonstrates a significant precision improvement of 94.11%. This high precision highlights the system's ability to minimize false positives, making it particularly valuable in applications where precise decisions are critical. Moreover, the recall achieved by the ensemble model is 90.56%, indicating its ability to reduce false negatives. Consequently, it ensures comprehensive detection, which is crucial in scenarios requiring high sensitivity. The results validate the effectiveness of ensemble learning in leveraging the complementary strengths of RNN, LSTM, and GRU models. By integrating these architectures, the ensemble achieves the highest performance across all metrics, demonstrating its potential as a robust and reliable solution for the given task as shown in Figure 6.

4.3. Amparative analysis

CoThe results clearly indicate that our study's methodologies (LSTM and ensemble learning) outperform the approach proposed by Osman *et al.* [19]. This improvement is likely due to the ability of our models to better capture complex patterns in the QDAT dataset, particularly when addressing temporal and non-linear relationships inherent in audio data. The highest accuracy achieved by the ensemble learning method underscores its potential as a superior approach for tasks involving the QDAT dataset as shown Table 4.

These findings highlight the importance of selecting advanced techniques tailored to the dataset's characteristics to improve predictive performance in audio-based evaluations.

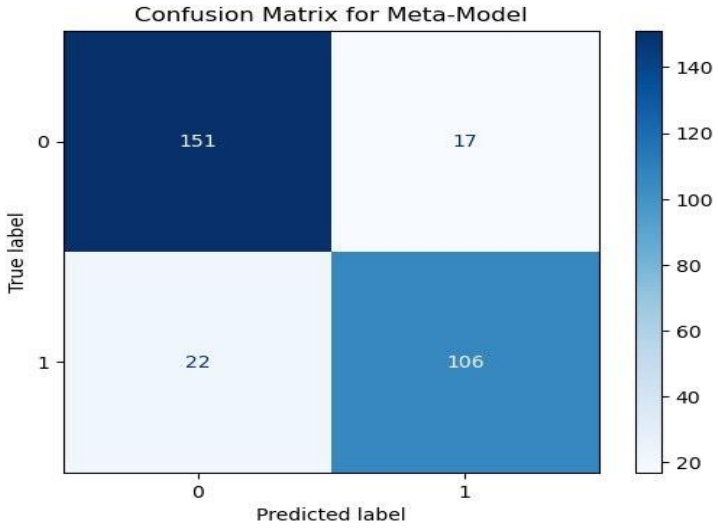

Figure 6. Confusion matrix of our model

Table 4. Summary of performance comparison only QDat dataset

Study	Methodology	Dataset	Accuracy (%)
[19]	Gradient boosting	QDAT	82.09
Our Study	LSTM	QDAT	84.3
Our Study	Ensemble learning	QDAT	88.57

5. CONCLUSION

This study demonstrates the potential of ensemble learning in Qur'anic recitation analysis, surpassing individual deep learning models such as GRU, LSTM, and CNN in terms of accuracy (88.57%), recall (90.56%), and precision (94.11%). The proposed ensemble approach effectively integrates the strengths of different architectures, balancing efficiency, long-term dependency capture, and intricate pattern recognition. Beyond achieving superior classification performance, this work highlights the ability of ensemble techniques to model complex temporal and phonetic structures within the QDAT dataset, setting a new benchmark for automated Tajweed rule evaluation. More importantly, this research underscores the broader implications of ensemble learning in audio-based applications, suggesting its adaptability to various speech processing tasks beyond Qur'anic recitation.

For future work, we aim to enhance the interpretability of our model by integrating attention mechanisms, allowing it to focus on critical phonetic features within Qur'anic recitation. This will help improve the system's ability to capture subtle pronunciation variations and Tajweed rules. Additionally, we plan to develop a benchmarking framework to systematically evaluate model performance on unseen Qur'anic verses, ensuring robustness across different reciters and recitation styles. To enhance generalization and reduce dependence on annotated datasets, we will explore self-supervised learning techniques that leverage large-scale unlabeled data. Furthermore, we intend to optimize inference speed and computational efficiency, making the model suitable for real-time recitation assessment. These advancements will contribute to the development of more precise, adaptive, and intelligent recitation analysis tools, supporting both educational and preservation efforts in Qur'anic studies.

FUNDING INFORMATION

Authors state no funding involved.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Mohamed Amine Daoud	✓	✓	✓	✓	✓	✓		✓	✓	✓	✓		✓	
Nayla Fatima Hadjar Kherfan		✓	✓		✓	✓	✓	✓				✓		
Abdelkader Bouguessa	✓	✓		✓		✓			✓				✓	
Sid Ahmed Mokhtar Mostefaoui		✓		✓							✓	✓		

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

DATA AVAILABILITY

Data availability is not applicable to this paper as no new data were created or analyzed in this study.

REFERENCES

[1] V. Verma *et al.*, "A novel hybrid model integrating MFCC and acoustic parameters for voice disorder detection," *Scientific Reports*, vol. 13, no. 1, p. 22719, 2023, doi: 10.1038/s41598-023-49869-6.

[2] M. A. Shakeel, H. A. Khattak, and N. Khurshid, "Deep acoustic modelling for Qur'anic recitation–current solutions and future directions," *IPSI Transaction on Internet Research*, vol. 20, No. 2, pp. 61–73, Jul. 2024, doi: 10.58245/ipsi.tir.2402.07.

[3] A. M. Alagrami and M. M. Eljazzar, "Smartajweed automatic recognition of arabic Qur'anic recitation rules," arXiv: arXiv:2101.04200, Dec. 26, 2020, doi: 10.48550/arXiv.2101.04200.

[4] M. Rababah, "Algorithms as a method for teaching fifth graders the rules of recitation 'Ahkam Al-Tajweed,'" *Dirasat: Educational Sciences*, vol. 33, no. 2, 2006, Accessed: Jan. 28, 2025. [Online]. Available: https://archives.ju.edu.jo/index.php/edu/article/view/1901

[5] M. K. Nasallah, "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)*, vol. 21, no. 2, pp. 55–61, 2016, doi: 10.9790/0837-21275561.

[6] R. Hamid, F. Naim, and N. Z. A. Naharuddin, "Makhraj recognition for Al-Qur'an recitation using MFCC," *International Journal of Intelligent Information Processing*, vol. 4, no. 2, pp. 45–53, 2013, doi: 10.4156/IJIIP.VOL4.ISSUE2.5.

[7] M. S. bin Sahmat and F. A. binti Zamri, "Enhancing Al-Qur'an reading proficiency in higher education: the implementation of the focused mad and idgham technique," *Journal of Cognitive Sciences and Human Development*, vol. 10, p. 1, 2024, doi: 10.33736/jcshd.6599.2024.

[8] B. Yousfi, A. M. Zeki, and A. Haji, "Isolated Iqlab checking rules based on speech recognition system," in *2017 8th International Conference on Information Technology (ICIT)*, IEEE, 2017, pp. 619–624, doi: 10.1109/ICITECH.2017.8080068.

[9] M. Lagarde, "La pause (waqf) et la reprise (ibtidāʾ)," in *Le parfait manuel des sciences coraniques al-Itqān fī ʿulūm al-Qurʾān de Ǧalāl ad-Dīn as-Suyūṭī (849/1445–911/1505)(2 vols)*, Brill, 2017, pp. 292–316, doi: 10.1163/9789004357112.

[10] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Advances in Neural Information Processing Systems*, vol. 25, no. 6, pp. 84–90, 2012, doi: 10.1145/3065386.

[11] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998, doi: 10.1109/5.726791.

[12] M. Daoud, Y. Dahmani, M. Bendaoud, A. Ouared, and H. Ahmed, "Convolutional neural network-based high-precision and speed detection system on CIDDS-001," *Data and Knowledge Engineering*, vol. 144, p. 102130, 2023, doi: 10.1016/j.datak.2022.102130.

[13] R. C. Staudemeyer and E. R. Morris, "Understanding LSTM -- a tutorial into long short-term memory recurrent neural networks," *arXiv: arXiv:1909.09586*, Sep. 12, 2019, doi: 10.48550/arXiv.1909.09586.

[14] A. Sherstinsky, "Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network," *Physica D: Nonlinear Phenomena*, vol. 404, p. 132306, 2020, doi: 10.1016/j.physd.2019.132306.

[15] K. Cho *et al.*, "Learning phrase representations using RNN encoder-decoder for statistical machine translation," *arXiv: arXiv:1406.1078*, Sep. 03, 2014, doi: 10.48550/arXiv.1406.1078.

[16] X. Dong, Z. Yu, W. Cao, Y. Shi, and Q. Ma, "A survey on ensemble learning," *Frontiers of Computer Science* vol. 14, no. 2, pp. 241–258, Apr. 2020, doi: 10.1007/s11704-019-8208-z.

[17] J. Farooq and M. Imran, "Mispronunciation detection in articulation points of Arabic letters using machine learning," in *2021 International Conference on Computing, Electronic and Electrical Engineering (ICE cube)*, IEEE, 2021, pp. 1–6. Accessed: Mar. 25, 2025. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9628251/

[18] A. Moustafa and S. A. Aly, "Towards an efficient voice identification using Wav2Vec2.0 and HuBERT based on the Qur'an reciters dataset," *arXiv*: *arXiv:2111.06331*, Nov. 11, 2021, doi: 10.48550/arXiv.2111.06331.

[19] H. M. Osman, B. S. Mustafa, and Y. Faisal, "QDAT: a data set for reciting the Qur'an," *International Journal on Islamic Applications in Computer Science And Technology*, vol. 9, no. 1, pp. 1–9, 2021.

[20] G. Samara, E. Al-Daoud, N. Swerki, and D. Alzu'bi, "The recognition of holy Qur'an reciters using the MFCCs' technique and deep learning," *Advances in Multimedia*, vol. 2023, pp. 1–14, Mar. 2023, doi: 10.1155/2023/2642558.

[21] N. Nigar, A. Wajid, S. A. Ajagbe, and M. O. Adigun, "An intelligent framework based on deep learning for online Qur'an learning during pandemic," *Applied Computational Intelligence and Soft Computing*, vol. 2023, pp. 1–9, Dec. 2023, doi: 10.1155/2023/5541699.
[22] A. A. Harere and K. A. Jallad, "Qur'an recitation recognition using end-to-end deep learning," *arXiv*: arXiv:2305.07034, May 10, 2023, doi: 10.48550/arXiv.2305.07034.
[23] A. A. Harere and K. A. Jallad, "Mispronunciation detection of basic Qur'anic recitation rules using deep learning," *arXiv*: arXiv:2305.06429, May 10, 2023, doi: 10.48550/arXiv.2305.06429.
[24] S. Al-Fadhli, H. Al-Harbi, and A. Cherif, "Speech recognition models for holy Qur'an recitation based on modern approaches and Tajweed rules: a comprehensive overview," *International Journal of Advanced Computer Science and Applications*, vol. 14, no. 12, pp. 1–16, 2023, doi: 10.14569/IJACSA.2023.0141297.
[25] R. Salameh, M. A. Mdfaa, N. Askarbekuly, and M. Mazzara, "Qur'anic audio dataset: crowdsourced and labeled recitation from non-arabic speakers," *Procedia Computer Science*, vol. 246, pp. 2684–2693, 2024, doi: 10.48550/arXiv.2405.02675.

BIOGRAPHIES OF AUTHORS

Dr. Mohamed Amine Daoud he is a lecturer at Ibn-Khaldoun University in Tiaret and a member of the LRIAS laboratory in Algeria. He obtained his Ph.D. in 2024 from the Higher National School of Computer Science (ESI) in Algiers. In 2016, he earned a Magister degree in Computer Science, specializing in Distributed and Mobile Computing (IRM), also from ESI. In 2010, he obtained his State Engineer degree in Computer Science, specializing in Advanced Information Systems, from Tiaret University. His research interests include machine learning, deep learning, security, modeling, and voice processing. He can be contacted at mohamedamine.daoud@univ-tiaret.dz or m_daoud@esi.dz.

Hadjar Kherfane Fatima Naila is a Master's student in Artificial Intelligence with a background in Information Systems and Software Engineering, earning her license from the University of Ibn Khaldoun Tiaret in 2023. She focuses on Speech Recognition and Deep Learning, developing innovative solutions for Qur'anic recitation recognition using advanced AI models like CNN and LSTM. In addition to her studies, Hadjar enjoys writing and is currently working on topics such as digital transformation. She can be contacted at email: hadjarnayla00@gmail.com.

Dr. Abdelkader Bouguessa is a lecturer at Ibn-Khaldoun University in Tiaret and a member of the LRIAS laboratory in Algeria. He obtained his Ph.D. in 2021 from the University of Science and Technology Oran – Mohammed Boudiaf (USTO-MB). In 2013, he earned a Master's degree in Computer Science, specializing in Information Systems and Web Technology (SITW), from Tiaret University. His research interests include machine learning and deep learning, computer vision, security, modeling, and programming. He can be contacted at abdelkader.bouguessa@univ-tiaret.dz.

Dr. Sid Ahmed Mokhtar Mostefaoui is a Senior lecturer at Ibn-Khaldoun University in Tiaret and the Director of the LRIAS laboratory in Algeria. He obtained my Ph.D. in 2021 from University of Science and Technology Mohamed Boudiaf, (USTO-MB), in Oran. In 2009, He earned a graduation-Port degree in Computer Science, specializing Information and knowledge systems also from USTO-MB. In 2002, He obtained his State Engineer degree in Computer Science, specializing in Software Engineering, from USTO-MB University. His research interests include machine learning and deep learning, security and Drug Discovery. He can be contacted at email: mokhtar.mostefaoui@univ-tiaret.dz.