World Journal of Advanced Engineering Technology and Sciences
eISSN: 2582-8266
Cross Ref DOI: 10.30574/wjaets
Journal homepage: https://wjaets.com/

(RESEARCH ARTICLE)

Check for updates

Identification of Mad Thabi'i Reading Errors in Surah Al-Kahfi Verses 1–10 Using Long Short-Term Memory: A Case Study of Students at DTA As-Salaam

Raden Marissa Lestari * and Syarif Hidayatulloh

Department of Informatics Engineering, Faculty of Information Technology, Adhirajasa Reswara Sanjaya University, Bandung, Indonesia.

World Journal of Advanced Engineering Technology and Sciences, 2026, 20(02), 079–089

Publication history: Received on 24 June 2026; revised on 05 August 2026; accepted on 07 August 2026

Article DOI: https://doi.org/10.30574/wjaets.2026.20.2.0404

Abstract

Mad Thabi'i is the most basic elongation rule in Quran recitation; however, many students at Diniyah Takmiliyah Awaliyah (DTA) As-Salaam still have difficulty in applying it correctly, while manual evaluation by teachers is limited by time constraints. This study proposes an automated approach to identify Mad Thabi'i recitation errors using a Long Short-Term Memory (LSTM) model. A dataset consisting of 710 valid audio samples representing correct and incorrect pronunciations of the three Mad Thabi'i letters (alif, waw, and ya) was collected from teacher recordings. Audio features were extracted using Mel-Frequency Cepstral Coefficients (MFCC) with delta and delta-delta coefficients, and data augmentation was applied to improve model generalization. A stacked LSTM model containing 61,249 trainable parameters was evaluated using stratified 5-Fold Cross Validation, achieving an average validation accuracy of 84.93% with an F1 score of 0.8453. Evaluation on an independent test set of 107 samples yielded an accuracy of 88.79% and an F1 score of 0.8877. Performance analysis showed the highest accuracy for kasrah (97.3%) and the lowest for fathah (81.0%), indicating that Mad Thabi'i readings involving alif with fathah were the most frequently misclassified among the evaluated categories.

Keywords: Mad Thabi'i; Long Short-Term Memory; Mel Frequency Cepstral Coefficient; Al-Qur'an Recitation; Tajweed

1. Introduction

Traditional Quran memorization still relies heavily on manual evaluation by teachers, although recent advances in deep learning have created significant opportunities to automate recitation assessment, including tajweed error detection. Artificial intelligence-based systems are capable of recognizing phonetic patterns from speech signals, enabling a more objective and consistent evaluation process than subjective assessments conducted by a single teacher.

The main challenge lies in the accurate pronunciation of Mad Thabi'i, the most fundamental rule of elongation in tajweed, where mispronunciation can change the meaning of the verses of the Quran [1]. Initial observations conducted at Diniyah Takmiliyah Awaliyah (DTA) As-Salaam showed that approximately 60% of 20 students still lacked good pronunciation of Mad Thabi'i. In addition, the existing evaluation method is still less than optimal because each student's reading is assessed individually by one teacher within a limited time limit. Because teachers may not be able to evaluate each student's reading consistently and comprehensively, the development of an automated reading assessment system based on artificial intelligence has become an important research issue.

* Corresponding author: Raden Marissa Lestari.

This study addresses the following research questions:

- To what extent can the Long Short-Term Memory (LSTM) model using the Mel-Frequency Cepstral Coefficients (MFCC) feature detect Mad Thabi'i reading errors in Surah Al-Kahfi verses 1–10?
- How does the proposed system perform when simultaneously combining all three Mad letters (alif, waw, and ya), compared to previous studies that only considered the alif letter?

This study makes three main contributions. First, it builds an audio dataset of Mad Thabi'i recitations covering all three Mad letters in Surah Al-Kahf verses 1–10, using recordings from teachers and youth of As-Salaam Mosque as training data and students as target users for system evaluation. Second, it develops a comprehensive MFCC feature extraction and LSTM modeling pipeline, including audio preprocessing, data augmentation, and five-fold cross-validation, ensuring that model predictions serve as the sole criterion for determining whether a recitation is correct or incorrect. Third, the proposed system is implemented in a real-world case study involving students at As-Salaam Mosque, enabling the identification of recitation error patterns in a non-formal Islamic educational environment.

Previous studies have explored deep learning-based approaches to detecting tajweed errors. Athiyah (2024) developed an MFCC-LSTM-based detection system for *Iqra learning materials* and achieved a test accuracy of 90.0%. However, the study only considered Mad alif letters with *fathah* and excluded waw and ya [2]. Min Gali (2024) expanded its scope to three types of Mad in Surah Al-Fatihah using 100 audio samples collected from five female reciters, achieving a best test accuracy of 90.0%. However, the model failed to predict the class with the fewest samples due to data imbalance [3]. Both studies relied on a limited number of controlled audio recordings. From an architectural perspective, LSTM, introduced by Hochreiter and Schmidhuber (1997), was designed to address the vanishing gradient problem in conventional recurrent neural networks (RNNs). Its combination with MFCC features has been shown to outperform traditional machine learning algorithms for tajweed error detection (Al Jallad & Al Harere, 2023) [4].

This study differs from previous research by simultaneously covering all three Mad letters in Surah Al-Kahf verses 1–10 and directly applying the proposed system in a real-world case study involving students at DTA As-Salaam. Consequently, this study addresses the research gap identified and recommended by previous studies.

2. Material and Methods

2.1. Research Objects and Data Sources

This study examines the misreading of Mad Thabi'i in verses 1 to 10 of Surah Al-Kahf, which collectively contain three mad letters of concern: alif (ا) following fathah; waw (و) following the dhammah; and ya (ي) after kasrah. These verses were chosen because they represent all three letter variants and are frequently recited in Islamic educational practices. This research was conducted at Diniyah Takmiliyah Awaliyah (DTA) As-Salaam, Bandung, a non-formal Islamic educational institution that teaches reading and memorizing the Qur'an to elementary school students (santri).

The data sources and test subjects were intentionally separated to avoid mixing model training with real-world implementation evaluation. The audio dataset used for model training and testing was collected from 25 contributors, consisting of 10 DTA As-Salaam teachers and 15 members of the As-Salaam mosque youth group (mosque youth), and included both correct and deliberately controlled incorrect recitations, each labeled by correctness class and harakat category. After the model was trained and evaluated on this dataset, the resulting system was then tested on third- to sixth-grade students as end-users to examine its performance on the different voice characteristics of the adult and adolescent contributors used for training.

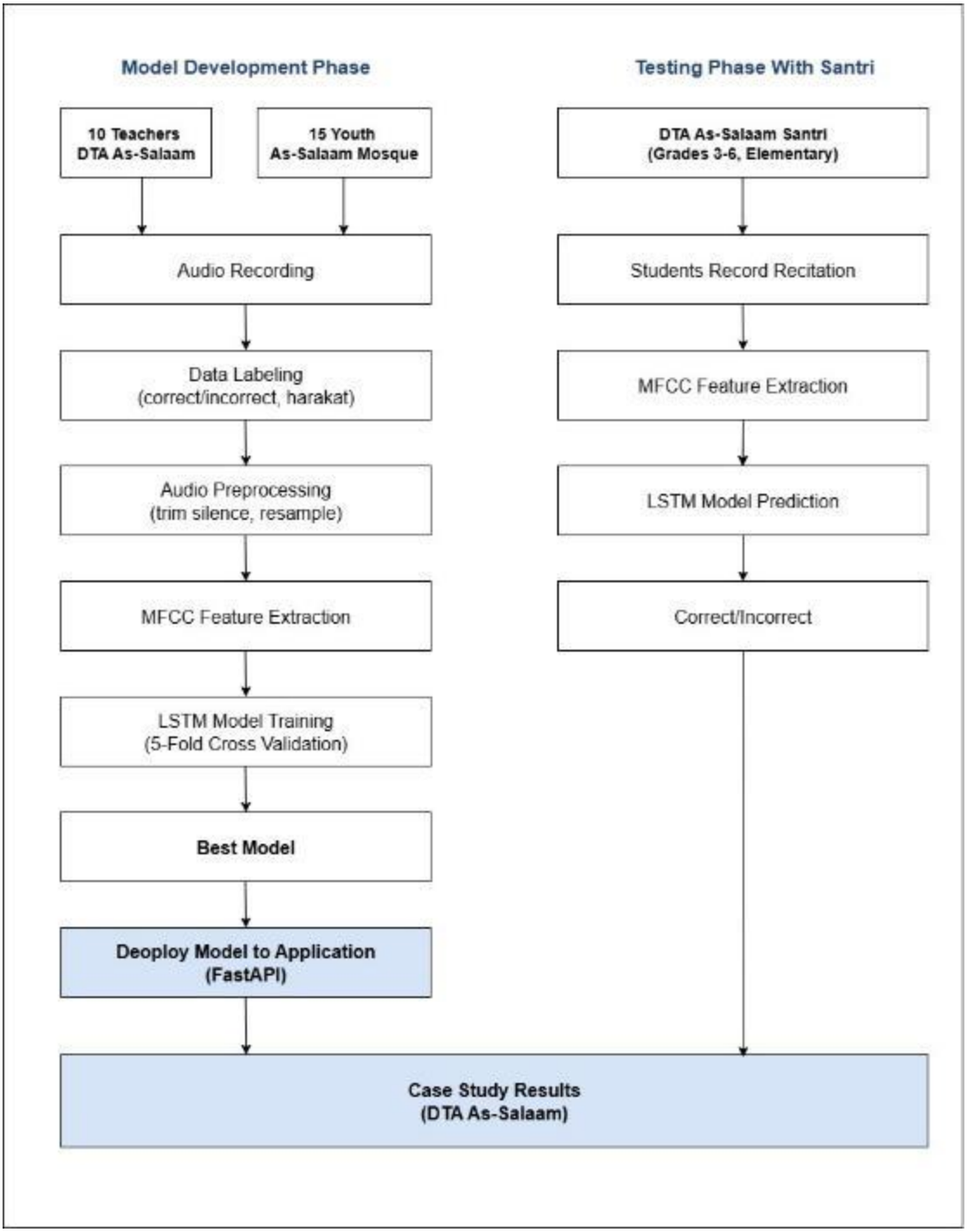

Figure 1 Scheme of training data sources (teachers and mosque youth) and test subjects (students)

2.2. Audio Preprocessing and Feature Extraction

Audio originally recorded in OGG format was converted to WAV, after which the beginning and end portions containing silence were trimmed using a 20-decibel threshold. Acoustic features were extracted as 40-coefficient Mel-Frequency Cepstral Coefficients (MFCC), combined with Delta and Delta-Delta derivatives (40 coefficients each), resulting in 120 feature dimensions per frame. Extraction was performed at a sampling rate of 22,050 Hz, with a Fourier transform window length (n_fft) of 2,048 samples and a hop length of 512 samples. The sequence length was standardized to 130 frames through padding or truncation, resulting in a fixed feature matrix of 130 × 120 for each sample.

The dataset was enriched through audio augmentation, namely mild noise injection (SNR ± 35 dB), time shifting (± 5%), and pitch shifting (one semitone up), all chosen because they preserve phoneme duration. Duration is a key

distinguishing feature in Mad Thabi'i recitation. Each original sample generated three additional augmented samples. Features were normalized using the mean and standard deviation calculated exclusively from the original (unaugmented) data to prevent data leakage. The original data were then split, stratified, into a 15% test set and an 85% training-validation set; the training-validation set was further partitioned using Stratified 5-Fold Cross-Validation. The augmented samples were included only in the training fold, while the validation set and the final test set consisted only of the original, unaugmented recordings.

2.3. LSTM Architecture and Model Training

The classification model was built using a compact Long Short-Term Memory (LSTM) architecture, designed to minimize the risk of overfitting. This architecture consists of two stacked LSTM layers (64 and 32 units), each followed by Batch Normalization and a Dropout rate of 0.40, followed by a Dense layer of 32 neurons with ReLU activation and a Dropout rate of 0.20, and one output neuron with sigmoid activation for binary (true/false) classification. L2 regularization was applied to the kernel and recurrent regularization of each LSTM layer.

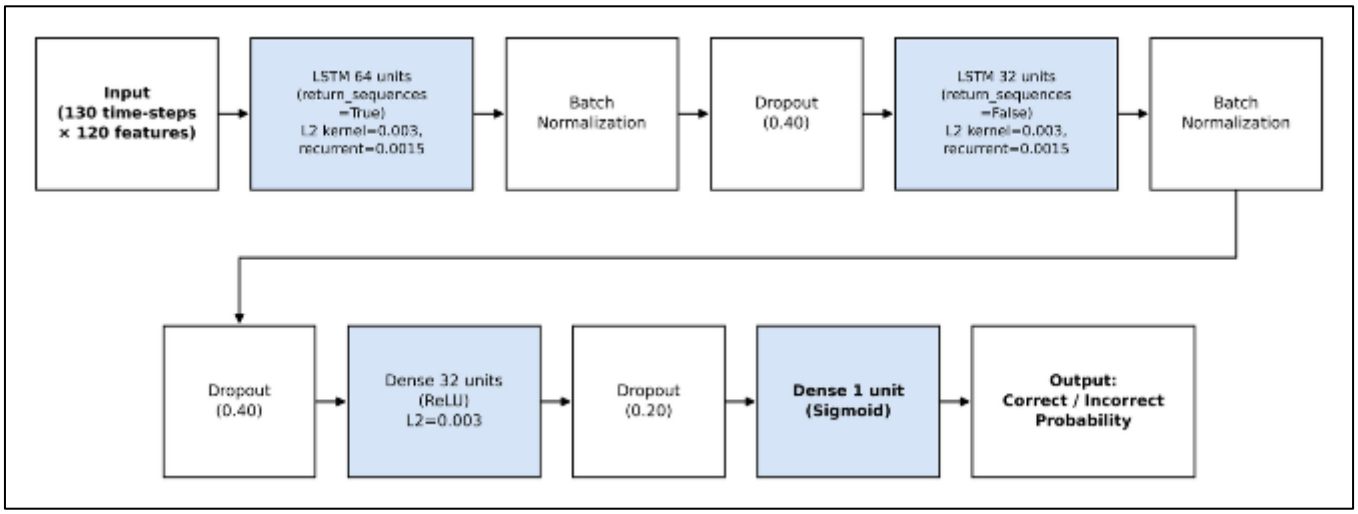

Figure 2 LSTM Model Architecture

This model was compiled with the Adam optimizer (learning rate 0.001) and binary cross-entropy loss, trained with batch sizes of 32 to 100 epochs under Stratified 5-Fold Cross-Validation. EarlyStopping (monitoring val_accuracy, with a patience of 15 epochs), ReduceLROnPlateau, and class weight balancing were applied to maintain training stability and efficiency. The model that achieved the best validation accuracy among the five folds was retained as the final model. The LSTM model output served as the sole determinant of correct/incorrect classification, while the recitation duration served only as supporting information and not as the primary determinant.

2.4. System Evaluation and Implementation Protocol

Model performance is assessed on a test dataset, which has been set aside at the outset, using the evaluation metrics in the following table.

Table 1 Model performance evaluation metrics

Evaluation Metrics	Purpose of This Study
Accuracy	Measures the proportion of correct predictions across the entire test dataset.
Precision, Recall, F1 Score	Evaluate the model performance for each class (true/false) in more detail, especially considering the potential for class imbalance in the dataset.
Confusion Matrix	Provides a comprehensive overview of the distribution of correct and incorrect predictions to uncover misclassification patterns.
Per-Harakat Accuracy	Evaluate the consistency of model performance for each Mad Thabi'i letter (fathah, dhammah, kasrah) separately.

The best-performing model, along with all preprocessing parameters (normalized mean and standard deviation, label encoder, and feature extraction parameters), was integrated into a FastAPI-based backend that exposed a health check endpoint, a prediction endpoint that accepted uploaded WAV audio files, and a model information endpoint. This integrated system was then piloted with DTA As-Salaam students as end-users to assess its performance under realistic usage conditions.

2.5. Ethical Considerations

All audio recordings were obtained with the consent of DTA As-Salaam and were used solely for research purposes and for the development of a Qur'an recitation evaluation system. The audio data were stored and processed in a controlled research environment; no personal data other than that required for the research was collected, and the data were not used for any commercial purposes outside the scope of this study.

3. Results

3.1. Description of Research Data Set

This dataset was constructed from 722 rows in CSV metadata containing recitation labels (correct/incorrect), harakat types (fathah, kasrah, dhammah), and audio file locations for mad thabi'i recitations. After checking file availability, 12 missing audio files were removed, leaving 710 usable samples with a balanced label distribution (355 correct / 355 incorrect) and a harakat distribution of 276 fathah, 234 kasrah, and 200 dhammah, which served as the basis for this study.

3.2. Audio Feature Extraction and Data Augmentation

Features were extracted using MFCC (40 coefficients) along with delta and its delta-delta derivatives, resulting in a 120-dimensional feature vector per frame (sample rate 22,050 Hz, hop length 512, N_FFT 2,048). The frame length was fixed at 130 (MAX_LEN, based on a mean of 131.17 and a median of 108 frames), with zero-padding for shorter clips and truncation for longer clips, resulting in a model input shape of (130, 120). Due to the limited original data, five augmentation techniques namely noise addition, time shift, pitch shift, acceleration, and deceleration were applied, expanding the original 710 samples to a total of 4,260 samples (feature array shape (4,260, 130, 120)), with 0 files failing to process. Each sample was labeled with the is_orig flag so that the cross-validation folds were only validated on the original data, preventing data leakage from the augmented derivatives.

3.3. Distribution of Training, Validation, and Testing Data

The original 710 samples were divided, using a stratification method, into 15% test data (107 samples: 54 correct / 53 incorrect) and 85% training-validation data (603 samples: 301 correct and 302 incorrect). This division was established at the outset, so the test data set was never used in training or cross-validation, ensuring that the final evaluation reflects true generalization to previously unseen data.

3.4. LSTM Model Architecture and Number of Parameters

A compact architecture was used to limit the risk of overfitting given the small dataset size: two LSTM layers (64 and 32 units), each followed by Batch Normalization and Dropout, a Dense feature pooling layer (32 units, ReLU), and a Dense sigmoid output layer for binary classification. The model has a total of 61,249 parameters (61,057 trainable, 192 non-trainable). Manual calculations using $4 \times [(\text{unit} \times (\text{input\_dim} + \text{unit})) + \text{unit}]$ confirmed 47,360 parameters for LSTM 1 and 12,416 for LSTM 2, in agreement with the model summary and validating the implemented architecture, which is much smaller than previous studies using 128/64 unit LSTMs.

3.5. Model Training Results (5-Fold Cross-Validation)

Training used Stratified 5-Fold Cross-Validation on 603 training-validation samples, with class weighting applied per fold and validation folds restricted to the original data only. The results are in the following table.

Table 2 5-fold cross-validation results

Folding	Val. acc.	Val. F1	Train (peak)	Overfitting gap	Era
1	90.91%	0.9091	100.00%	9.1%	51
2	71.90%	0.6999	96.07%	24.2%	32
3	80.99%	0.8094	94.60%	13.6%	33
4	87.50%	0.8748	100.00%	12.5%	39
5	93.33%	0.9333	99.20%	5.9%	45
Average	84.93%	0.8453	97.97%	13.0%	40

The average validation accuracy was 84.93% (F1 = 0.8453; overfitting difference 13.0%; SD ±8.63%), with Fold 2 performing the worst (71.90% accuracy, 24.2% difference) and Fold 5 performing the best (93.33% accuracy, 5.9% difference); the Fold 5 model was kept as the best model. Early stopping (15 epochs of patience) consistently stopped training between epochs 32–51, keeping the training and validation curves close together and limiting overfitting.

3.6. Model Evaluation on Test Data

The Fold 5 model was evaluated on 107 held-out test samples, achieving an accuracy of 88.79% and an F1 score of 0.8877.

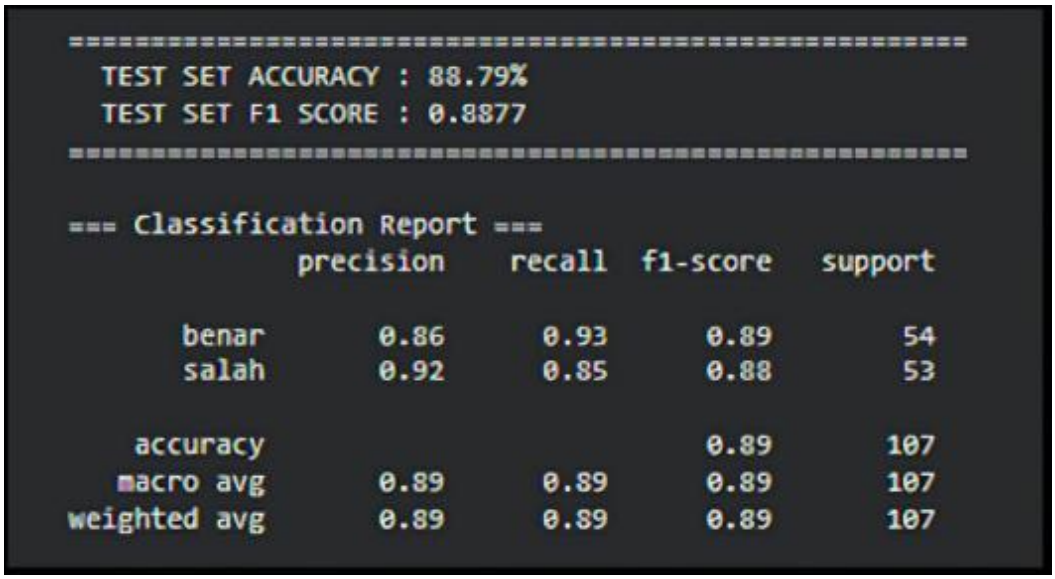

Figure 3 Results of Model Testing on Test Data

Based on Figure 3, the results show that the model is quite good at finding almost all correct reading data, but there are still a number of incorrect readings that are mistakenly classified as correct. Conversely, in the "incorrect" class, precision is higher (92%) than recall (85%), which indicates that when the model predicts a reading as "incorrect", the prediction tends to be reliable, although there are still some reading errors that have not been successfully detected (still classified as correct). The average F1-score value of 0.89 (89%) indicates that overall the model has a fairly good balance between precision and recall in both classes.

The confusion matrix shows that of the 54 samples labeled "true," 50 were correctly classified (4 false negatives); of the 53 samples labeled "false," 45 were correctly classified (8 false positives). This gives a precision/recall of 0.86/0.93 for "true" and 0.92/0.85 for "false" (F1 of 0.89 and 0.88, respectively; macro-average/weighted F1 = 0.89), confirming that the model reliably detects true recitations while producing more reliable, albeit slightly less complete, predictions when marking recitations as false.

3.7. System Test Results

Black-box testing on a fully integrated application (feature extraction → normalization → classification) using 35 mad thabi'i words/phrases in all three harakat, each tested in correct and incorrect reading scenarios (total 70 samples).

Table 3 System test results

Reading	Input	Output	Information
"الْكِتٰبَ عَبْدِهِ عَلٰى "	Correct Reading: 'alaa 'abdihil kitaaba	Correct	In accordance
	Wrong Reading: 'ala 'abdihil kitaba	Wrong	In accordance
"شَدِيْدًا "	Correct Reading: syadiidan	Correct	In accordance
	Wrong Reading: syaadiidaa	Correct	It is not in accordance with
"الَّذِيْنَ وَيُبَشِّرَ God"	Correct Reading: wa yubassyiral mu'miniina allaziina	Correct	In accordance
	Wrong Reading: waa yubassyiral mu'miniinal lazina	Wrong	In accordance
"يَعْمَلُوْنَ "	Correct Reading: ya'maluuna	Correct	In accordance
	Wrong Reading: ya'maalunaa	Correct	It is not in accordance with
"الصّٰلِحٰتِ نَ"	Correct Reading: naṣ ṣaali ḥ aati	Correct	In accordance
	Wrong Reading: naṣ ṣaaali ḥ aati	Wrong	In accordance
"مٰ "	Correct Reading: maa	Correct	In accordance
	Wrong Reading: ma	Wrong	In accordance
"فِيْهِ كِتٰبٌ"	Correct Reading: ki ṡ iina fiihi	Correct	In accordance
	Wrong Reading: ki ṡ ina fihi	Wrong	In accordance
"الَّذِيْنَ "	Correct Reading: allaziina	Correct	In accordance
	Wrong Reading: allazinaa	Wrong	In accordance
"قَا "	Correct Reading: qaa	Wrong	It is not in accordance with
	Wrong Reading: qa	Wrong	In accordance
"لُوا "	Correct Reading: luu	Correct	In accordance
	Wrong Reading: lu	Wrong	In accordance
"لَهُمْ مَّا"	Correct Reading: maa lahum	Correct	In accordance
	Wrong Reading: ma lahum	Wrong	In accordance
"لِاٰبَآىِٕهِمْ وَّلَا"	Correct Reading: wa laa li aabaa'ihim	Correct	In accordance
	Wrong Reading: waa laaa li aabaaa'ihim	Wrong	In accordance
"اَفْوَاهِهِمْ "	Correct Reading: afwaahihim	Correct	In accordance
	Wrong Reading: afwahihim	Wrong	In accordance
"يَّقُوْلُوْنَ اِنْ"	Correct Reading: in yaquuluuna	Correct	In accordance
	Wrong Reading:in yaaquuluunaa	Wrong	In accordance
"كَذِبًا اِلَّا"	Correct Reading: illaa kazibaa	Correct	In accordance
	Wrong Reading: illa kaziba	Wrong	In accordance
	Correct Reading: baakhi'un	Correct	In accordance

"بَاخِعٌ "	Wrong Reading: bakhi'un	Wrong	In accordance
"اٰثَارِهِمْ عَلٰى "	Correct Reading: 'alaa aatsaarihim	Correct	In accordance
	Wrong Reading: 'ala atsarihim	Wrong	In accordance
"يُؤْمِنُوْا لَّمْ اِنْ "	Correct Reading: in lam yu'minuu	Correct	In accordance
	Wrong Reading: in lam yuu'minu	Wrong	In accordance
"بِهٰذَا "	Correct Reading: bihaazaa	Wrong	It is not in accordance with
	Wrong Reading: bihaza	Wrong	In accordance
"الْحَدِيْثِ "	Correct Reading: hadith	Correct	In accordance
	Wrong Reading: Hadith	Wrong	In accordance
"عَلَى Mo جَعَلْنَا اِنَّا "	Correct Reading: innaa ja' alnaa maa 'alaa	Correct	In accordance
	Wrong Reading: inna ja' alna maa 'ala	Wrong	In accordance
"لَّهَا زِيْنَةً "	Correct Reading: ziinatal lahaa	Correct	In accordance
	Wrong reading: ziinatal laha	Wrong	In accordance
"اِنَّا "	Correct Reading: inna	Correct	In accordance
	Wrong Reading: ina	Wrong	In accordance
"لَجِمِ "	Correct Reading: lajaa'i	Wrong	It is not in accordance with
	Wrong Reading: lajaa'i	Wrong	In accordance
"عِلُوْنَ "	Correct Reading: 'iluuna	Correct	In accordance
	Wrong Reading: I'ilunaa	Wrong	In accordance
"مَا "	Correct Reading: maa	Correct	In accordance
	Wrong Reading: ma	Wrong	In accordance
"صَعِيْدًا "	Correct Reading: ṣa' iidan	Correct	In accordance
	Wrong Reading: ṣa' idan	Wrong	In accordance
"اَصْحٰبَ "	Correct Reading: aṣ ḥ aaba	Wrong	It is not in accordance with
	Wrong Reading: aṣ ḥ abaa	Wrong	In accordance
"وَالرَّقِيْمِ "	Correct reading: war raqiimi	Correct	In accordance
	Wrong reading: war raqiimii	Wrong	In accordance
"كَا "	Correct Reading: kaa	Wrong	It is not in accordance with
	Wrong Reading: ka	Wrong	In accordance
"مِنْ نُوْا "	Correct Reading: nuu min	Correct	In accordance
	Wrong Reading: nu min	Wrong	In accordance
"اٰيٰتِنَا "	Correct Reading: aayaatinaa	Correct	In accordance
	Wrong Reading: ayatina	Wrong	In accordance
"فَقَا "	Correct Reading: faqoo	Correct	In accordance
	Wrong Reading: faqo	Wrong	In accordance
"اٰتِنَا "	Correct Reading: aatinaa	Correct	In accordance
	Wrong Reading: atina	Wrong	In accordance
	Correct Reading: lanaa min amrinaa rushadaa	Correct	In accordance

" رَشَدًا اَمْرِنَا مِنْ لَنَا"	Wrong Reading: lanaa min amrinaa rasyada	Wrong	In accordance

The system correctly classified 63/70 samples (90% accuracy), slightly exceeding the internal testing accuracy (88.79%) and exceeding the research target of 80%. The seven misclassifications followed two patterns: some incorrect readings were classified as correct, and more commonly short words ending in fathah (e.g., "qaa," "kaa") whose correct readings were classified as incorrect, consistent with lower accuracy per fathah class and reduced temporal cues in very short audio segments.

3.8. System implementation

The best models were deployed in a web application with a FastAPI backend (health-check, POST /predict, and GET /model-info endpoints) that handled feature extraction, normalization, and inference, and a three-screen frontend (Home, Record, Results) for microphone-based live reading testing. Initial discrepancies between the silence-trimming logic used in training versus inference were identified and fixed.

Figure 4 Front view and Voice record view

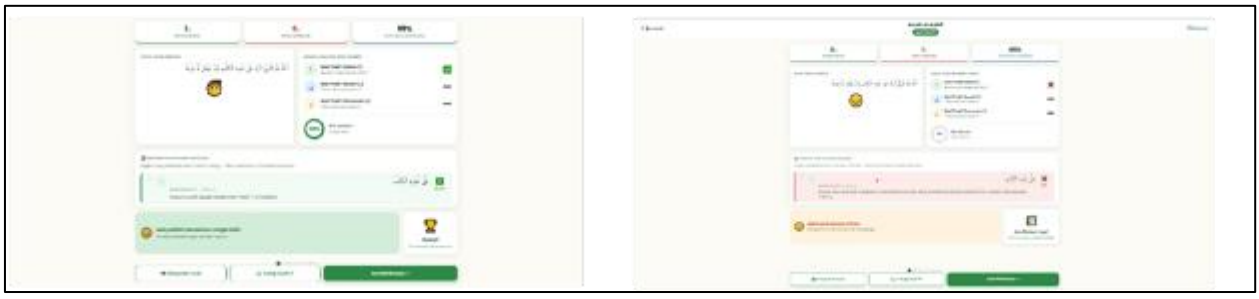

Figure 5 Display of the predicted sound results of the model

The current implementation is temporary, using a local server intended solely for research testing. It does not yet support simultaneous multi-user access and would require a hosting solution or a cloud server for broader public implementation.

3.9. Threats to validity

Several factors limit the generalizability of these findings. First, the dataset is still limited, with 710 original recordings from 25 contributors, and the cross-validation variance is nontrivial, as Fold 2 dropped to 71.90% accuracy compared to a peak of 93.33% in Fold 5, indicating that the model's performance estimates carry significant uncertainty despite the reported average. Second, the training and validation data were drawn from teachers and youth group members, while the final system testing involved students with different vocal characteristics. Although 90% black-box accuracy suggests that the model generalizes reasonably well across these shifts, a larger and more age-diverse training sample would strengthen this claim. Third, this study is limited to Surah Al-Kahf verses 1–10 and three categories of harakat, so the results cannot be extended to other verses, reciters, or Tajweed rules beyond Mad Thabi'i. Finally, the implemented system was only evaluated as a single-user local instance; concurrent and multi-user testing was not conducted.

4. Discussion

The research objective of 80% accuracy was successfully surpassed by the model, which achieved 88.79% on a distinct test set and 90% in an independent black-box assessment. The alignment between these figures indicates that the results are robust rather than a byproduct of specific data partitioning. By utilizing a streamlined two-layer LSTM architecture with 61,249 parameters, the study effectively addressed overfitting concerns despite the small dataset, as evidenced by a 5.9% reduction in the training-validation gap during the optimal fold. In line with existing MFCC-LSTM research on tajwid, classification errors were most prevalent for fathah. This is likely due to fathah possessing the briefest temporal duration of the three harakat, providing the model with minimal acoustic data to process—a trend confirmed by both black-box discrepancies in short words and per-harakat metrics. While prior research often focused on a limited Mad Thabi'i, this work encompasses all three relevant surahs and utilizes authentic student recitations, offering a more practical, though early, evaluation of its viability for educational settings.

5. Conclusion

This study demonstrates that an MFCC-based LSTM model can automatically identify Mad Thabi'i misreadings across all three Mad letters with 88.79% accuracy, and that the resulting system performs well when tested with actual students, not just readers from the training population. This approach offers a practical complement to manual teacher evaluation in non-formal Quranic education settings. Future work should expand the dataset, address the relatively weak fathah classification, and extend its applicability to a scalable, multi-user platform suitable for routine classroom use.

Compliance with ethical standards

Disclosure of conflict of interest

The authors declare no conflict of interest, and informed consent was obtained from DTA As-Salaam and all recording contributors prior to data collection, and all recordings were used solely for research and system development purposes.

References

[1] Heti Sumi Lestari, Neuneu Nurlaela Hasanah, Revika Khoerunnisa, Iin Indriani, and Muhamad Ibnu Malik, "Characteristics of Mad Asli and Mad Far'i Laws in the Study of Tajwid Science," *QAYID: Journal of Islamic Education*, vol. 1, no. 2, pp. 252–264, Dec. 2025, doi: 10.61104/qd.v1i2.631.

[2] D. Detection and K. Reading, "APPLICATION OF MEL-FREQUENCY CEPSTRAL COEFFICIENT AND LONG SHORT-TERM MEMORY METHOD."

[3] SO Muhamad, F. Akter, M. Gali, F. Sains, and D. Teknologi, "CLASSIFICATION OF TAJWID MAD LAWS IN SURAH AL-FATIHAH USING LONG SHORT-TERM MEMORY ALGORITHM WITH MEL-FREQUENCY CEPSTRAL COEFFICIENT FEATURE EXTRACTION IN INFORMATICS ENGINEERING STUDY PROGRAM."

[4] A. Al Harere and K. Al Jallad, "Quran Recitation Recognition Using End-to-End Deep Learning."

[5] M. Aqilah bint Muhaiyat Center for Al-Quran and Sunnah Studies and F. Islamic Studies, "The Approach of the Speed Tahfiz Program in Enhancing Al-Quran Memorization at the Amalillah Al-Quran Academy, Shah Alam," 2026.

[6] O. Virgantara Putra, F. Reza Pradana, and J. Istiqlal Qalbi Adiba, "Mad Reading Law Classification Using Mel-Frequency Cepstal Coefficient (MFCC) and Hidden Markov Model (HMM)," 2021.

[7] S. Md Salleh and S. Mohamad, "THE LAW OF MAD IN TAJWID SCIENCE: LITERATURE HIGHLIGHTS AND ANALYSIS OF QURAN READING MASTERY," *International Journal of Modern Education*, vol. 7, no. 28, p. 487, Dec. 2025, doi: 10.35631/ijmoe.728036.

[8] AA Elmasallati, FS Alahjal, and AM Essa, "Hybrid MFCC-LSTM Models for Automatic Evaluation of Three Primary Tajweed Rules in Holy Quranic Recitation," *Journal of Technology Research*, vol. 3, no. 2, pp. 126–132, Dec. 2025, doi: 10.26629/jtr.2025.15.

[9] A. Zaki and M. Hamidah, "EFFORT TO IMPROVE TAJWID LEARNING ON THE LAWS OF MAD READING IN DKM MUJAHIDDIN KP. BUNGBULANG VILLAGE, BUNGBULANG DISTRICT, GARUT REGENCY." [Online]. Available: https://jurnal.pustakaturats.com/index.php/edupesantren

[10] SS Alrumiah and AA Al-Shargabi, "A Deep Diacritics-Based Recognition Model for Arabic Speech: Quranic Verses as a Case Study," *IEEE Access*, vol. 11, pp. 81348–81360, 2023, doi: 10.1109/ACCESS.2023.3300972.