Indonesian Journal of Electrical Engineering and Computer Science
Vol. 39, No. 2, August 2025, pp. 914~926
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v39.i2.pp914-926

Detection of the Tajweed rules in the Qur'anic recitations

Karim Aly Mohammad[1], Ahmed Hisham Kandil[1], Ahmed Mohamed El-Bialy[1,2], Sahar Ali Fawzi[1,3]
[1]Department of Systems and Biomedical Engineering, Faculty of Engineering, Cairo University, Giza, Egypt
[2]Department of Systems and Biomedical Engineering, Faculty of Engineering, El-Shorouk Academy, Cairo, Egypt
[3]School of Information Technology and Computer Science, Nile University, Giza, Egypt

Article Info

Article history:

Received Oct 16, 2024
Revised Apr 20, 2025
Accepted Jul 2, 2025

Keywords:

Quranic audio processing
Speech pattern matching
Tajweed database
Tajweed learning
Tajweed rules detection

ABSTRACT

Tajweed is the science of reciting the Holy Quran, focusing on the clarity and correctness of recitation. This paper aims to accurately detect the spoken Tajweed rules applied during Quranic recitation, providing a well-structured Tajweed rules database for further analysis, Tajweed learning, and the training of advanced classification models. The main contribution of this work is to identify a high-accuracy approach for Tajweed rules detection and analysis. An improved template matching approach is introduced to enhance detection accuracy by matching the Quranic verse audio file with multiple speech patterns of a specific rule and selecting the best match. The Quranic audio file is segmented into smaller patterns by finding the correlation between the adjacent audio frames. Then, the template matching is applied to these segmented patterns to identify the best-matching ones. The template matching technique relies on a Tajweed database of 487 patterns of the Madd, Noon Sakinah, Tanween, and Meem Sakinah rules. An overall detection accuracy of 97.1% is achieved, and the Tajweed-pattern database is expanded to include the newly detected rules, increasing their total count to 2,583. Furthermore, an application based on the detected rules in this study was developed to enhance the performance of new Tajweed learners.

Corresponding Author:

Karim Aly Mohammad
Department of Systems and Biomedical Engineering, Faculty of Engineering
Cairo University
12613 Giza, Egypt.
Email: karimaly08@gmail.com

1. INTRODUCTION

Reciting the Holy Quran correctly is a religious obligation for all Muslims, who make up 25% of the world's population [1]. Tajweed rules were established to provide guidelines for the accurate recitation of the Holy Quran [2]. As the number of non-Arabic Muslim speakers has increased, there has been a growing demand to develop computer-aided pronunciation training (CAPT) systems [3] aimed at instructing Muslims on the correct recitation of the Holy Quran. CAPT systems streamline the educational process, offering flexibility and accessibility at any time and place.

Numerous researchers have dedicated considerable efforts to attain this goal, and their endeavors can be classified into two main categories. The first relies on constructing a complete speech recognition system, commencing with statistical models such as hidden Markov models [4], [5] and progressing to more advanced models like deep neural networks (DNN) [6], [7], [8], [9], However, this approach necessitates manually segmenting hundreds of hours of Quranic recitations by specialists to generate a database suitable for constructing these systems. Typically, these research initiatives are affiliated with commercial entities which provide the necessary funding for their work. Often, these researchers do not disclose or publish the databases they create.

Another group of researchers, lacking the financial resources and capabilities, has pursued simpler methods to provide a simple speech recognition system to detect Tajweed rules in Quranic recitations through template-matching techniques or simple DNN models.

Template-matching in speech involves comparing the feature set of a target speech pattern with the feature set of an audio file, where the goal is to locate the target pattern within the file. This method has proven effective in detecting Tajweed rule patterns within Quranic recitations, as demonstrated by several researchers. For instance, Saber *et al.* [10] addressed the elevated and the lowered (التفخيم والترقيق) rules, studying eight common learners' mistakes in their recitation with a 95% verification accuracy. The researchers implemented the system on isolated Quranic words, making it unsuitable for complete sentences. The approach relied on mel-frequency cepstral coefficients (MFCCs) as a comparison feature between the audio reference and rule pattern. MFCCs are key speech features in the frequency domain, scaling frequencies to match what the human ear can hear closely [11]. In another study, Awaid *et al.* [12] utilized MFCCs, linear predictive coefficient (LPC), and other speech features to identify the position of 15 Quranic words in a complete Quranic recitation, achieving the highest detection accuracy of 94.2% using the MFCC features. Similarly, other speech recognition studies [13], [14] have also utilized MFCCs as a key feature in both Arabic and Quranic speech processing.

Several studies have explored using simple DNN models to detect Tajweed rules in Quranic recitations. One study focused on one of the Noon Sakinah rules (merging rule), using an artificial neural network with one hidden layer and training the model using a dataset of 300 recitations, but the system achieved a relatively low accuracy of 77.7% due to the small dataset size [15]. Harere *et al.* [16] used a long short-term memory (LSTM) model to detect the disconnected Madd (المد المنفصل) and some Noon Sakinah rules, training on 1500 utterances of verse 109 from Surah Al-Maeda (سورة المائدة). The model achieved a detection accuracy of about 96%. While these models provide promising results, their performance is limited by the small size of the datasets used.

There is a significant lack of publicly available datasets for Tajweed speech patterns, which are essential for developing classification techniques, such as DNN, and for assisting developers in building CAPT systems. This study addresses this gap by proposing an improved template-matching approach to create an automated method for detecting Tajweed rules in Quranic recitations with high accuracy. The ultimate goal is to establish a comprehensive database of Tajweed speech patterns to facilitate further research and the development of Tajweed learning applications.

Traditional Template matching involves comparing a single reference pattern with the audio file. However, this method is unsuitable for our study, as Tajweed rule features typically vary from one position to another and across different recitation sessions. Thus, comparing the audio file with only one pattern is insufficient. To resolve this issue, the proposed template-matching technique compares the Quranic verse audio file with multiple speech patterns of a specific rule. The closest matched pattern is then deemed the most reliable and selected. The proposed hypothesis states that increasing the number of Tajweed rule patterns in the database will improve the rule's detection accuracy.

This study focuses on the rules of Madd (المد), Noon Sakinah and Tanween (النون الساكنة والتنوين), and Meem Sakinah (الميم الساكنة) rules, which are essential for every beginner. It employs a reference database of Tajweed speech patterns from prior research [17], containing 487 speech segments aligned with the Tajweed rules identified in the first 40 verses of Surah Al-Anfal (سورة الانفال) as recited by Sheikh Al-Hosary, who is widely recognized as a leading scholar of Tajweed and recitations in the Islamic community [18].

Two additional efforts have been incorporated into this study to support its primary objective: assisting new learners in accurately reciting the Quran. the first one addresses a common pronunciation mistake of the Tajweed rules under the scope of this work which is the precise time duration of reciting them. Studying the variation of the detected rules' time duration has been done through a consistency analysis study. This analysis aids in establishing a reliable threshold for such variations, providing a valuable tool for evaluating the pronunciation correctness of new learners. The second effort involves developing a Tajweed learning application that utilizes this research's constructed Tajweed rules database to facilitate effective learning for new learners.

This paper is structured as follows: Section 2 introduces the key principles of Tajweed. Section 3 outlines the proposed method. Section 4 presents the results and discussion. Finally, section 5 concludes the paper and outlines directions for future work.

2. PRINCIPLES OF TAJWEED

The definition of Tajweed by linguistic definition is the Betterment. Its applied definition is Articulating each letter from its precisely articulating point with its specific characteristics [2]. The interest of the Tajweed science is the Quranic Words. It aims to prevent Quran readers from pronunciation mistakes during their recitations.

Madd is the lengthening of the pronunciation time of a vowel. There are three short and three long vowels in the Arabic language. The short vowels are Fathah (َ◌), Dhammah (ُ◌), and Kasrah (ِ◌). The time of pronouncing a short vowel represents a unit reference time for other rules. This unit of time is commonly called a movement (حركة). According to Tajweed instructors, a movement is the time required to either extend or contract a finger of the hand. The long vowels are found in the three Madd letters which are Alif Sakinah preceded by Fathah (َا), Yaa Sakinah preceded by Kasrah (ِي), and Waow Sakinah preceded by Dhammah (ُو).

The Madd rules in the Holy Quran can be classified into two main categories; the natural Madd and the sub-Madd [2]. The natural Madd is founded in the presence of one of the Madd letters that is not followed by Hamza (ء) or Sokon (ْ◌). The time of pronouncing a natural Madd is 2 movements.

The sub-Madd rules are based on the presence of one of the Madd letters followed by Hamza (ء) or Sokon (ْ◌). The sub-Madd has different forms with different times of pronunciation ranging from 2 to 6 movements [2]. Table 1 concludes the studied Madd rules in this work with their pronunciation times. As observed from the Table 1, a professional reciter should consistently pronounce each Madd type with its precise number of movements.

Table 1. Madd rules

Madd rule	Pronunciation duration (Number of movements)
Natural Madd (المد الطبيعي)	2
Connected Madd (المد المتصل)	4 or 5
Disconnected Madd (المد المنفصل)	2, 4, or 5
Accidental Madd (المد العارض للسكون)	2, 4, or 6
Necessary Madd (المد اللازم)	6

The Noon Sakinah (النون الساكنة), Tanween (التنوين), and Meem Sakinah (الميم الساكنة) rules have been addressed in previous work [17]. A common feature of these rules is the Ghunnah sound (صوت الغنة) which also takes a precise time duration to pronounce. The time duration of the Ghunnah sound can be simplified to approximately equal to two movements. Ghunnah sound is a nasal sound and characteristic of the Noon, Tanween, and Meem letters. Table 2 represents a set of Noon Sakinah, Tanween, and Meem Sakinah rules studied in this work.

Table 2. Tajweed rules that contain Ghunnah sound

Classification	Rule
Noon Sakinah and Tanween	Merging with Ghunnah (الإدغام بغنة)
	Change (الإقلاب)
	Hiding (الإخفاء)
Meem Sakinah	Merging (إدغام)
	Hiding (إخفاء)

3. METHOD

3.1. Segmentation process

The audio file is typically segmented into smaller utterances before the template-matching process [19]. The segmentation process relies on tracking the variation of one or more speech features throughout the audio file. A sudden variation of the same feature with time indicates a transition between speech syllables. Time and frequency domain features can be employed to detect the boundaries between speech syllables [20], [21] and segment them into smaller utterances.

The MFCC feature is chosen as a tracking feature to identify approximate boundaries within the recitation audio file, dividing each verse (Aya) into smaller components to facilitate the detection of Tajweed rules. Figure 1 illustrates the speech pattern of the Quranic text "يَسْأَلُونَكَ عَنِ الْأَنفَالِ", which is part of verse number 1 of Surah El-Anfal. The speech pattern is displayed in both time and frequency domains. Visual boundaries within the verse speech file indicate transitions from one utterance pattern to another. A new method has been implemented to automate the detection of these boundaries, which relies on finding the correlation between the adjacent audio frames. At the transition boundaries, the correlation decreases and can be effectively detected.

The proposed segmentation method involves two main steps: (1) MFCC calculation – the audio file is divided into overlapping 80-millisecond frames with a 5-millisecond step size, and MFCCs are computed and stored sequentially; (2) correlation function application – the pearson correlation coefficient (1) [22] is applied between the MFCC coefficients of each pair of adjacent frames.

$$corr = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}} \quad (1)$$

Here, x_i and y_i represent the MFCC coefficient at index i for two adjacent frames. $\bar{x}$ and $\bar{y}$ denote the mean of the MFCC coefficients of the two frames, and n is the total number of MFCC coefficients.

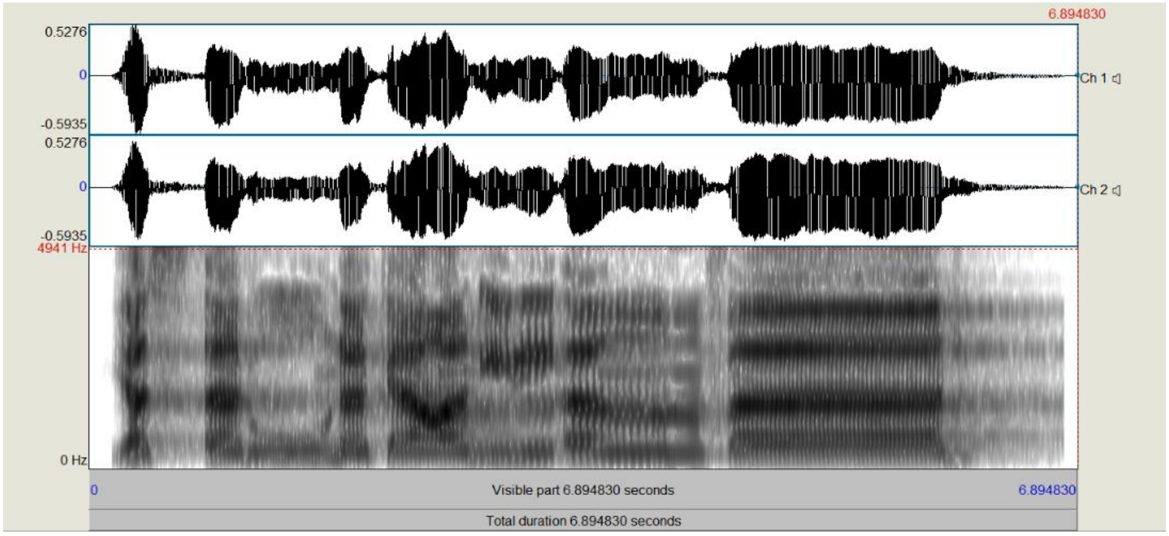


Figure 1. Time domain speech pattern and spectrogram of part of the first verse of Surah El-Anfal

Figure 2 illustrates the output of the correlation function applied to verse 1 of Surah El-Anfal, where Figure 2(a) initially exhibits fluctuations due to minor variations in speech patterns. Given the syllabic nature of the Arabic language, the typical syllable duration for normal-speed readers is approximately 177 milliseconds. While this duration may be shorter for high-speed readers, we assume a minimum syllable time of 80 milliseconds to accommodate a broader range of reading speeds. Since the distance between syllables is always greater than 80 milliseconds, rapid fluctuations that do not correspond to actual syllabic boundaries can be filtered out. As a result, the correlation graph is simplified, retaining only the prominent dips that indicate syllable boundaries, as shown in Figure 2(b).

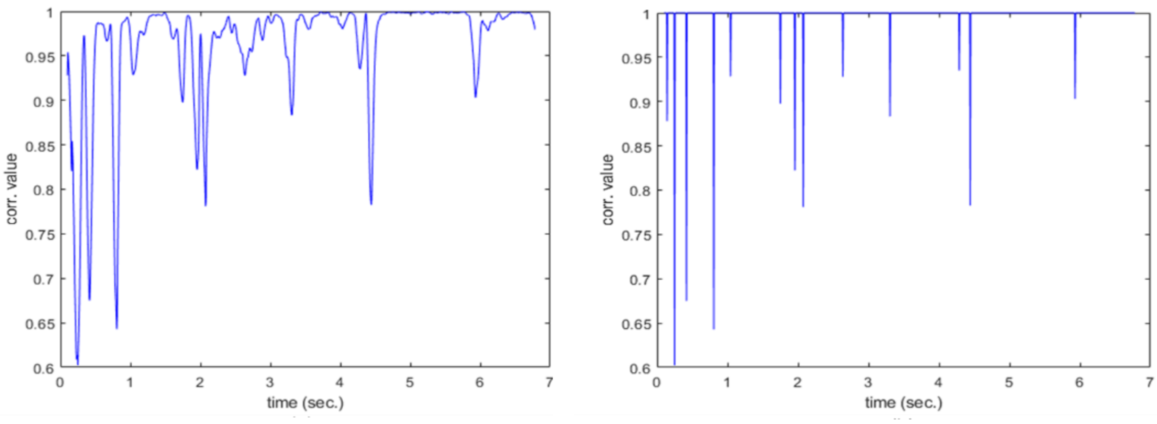


Figure 2. Application of the correlation function to verse 1 of Surah El-Anfal, showing (a) the original Correlation graph and (b) its simplified version

The dips shown in Figure 2(b) indicate a low correlation between adjacent audio frames. These low-correlation points are assumed to serve as segmentation boundaries, dividing the speech file into smaller patterns. The resulting segmentation of the speech file is visualized in Figure 3.

The Tajweed utterance segments targeted in this research exhibit a high correlation between their speech frames, particularly in sounds like Madd and Ghunnah. As a result, they can be effectively isolated as separate segments during the segmentation process, making them easier to identify. This approach allows us to compare the feature set of the Tajweed rule pattern within the core part of each segment rather than analyzing the entire speech file. Consequently, this technique enhances processing speed and improves detection accuracy.

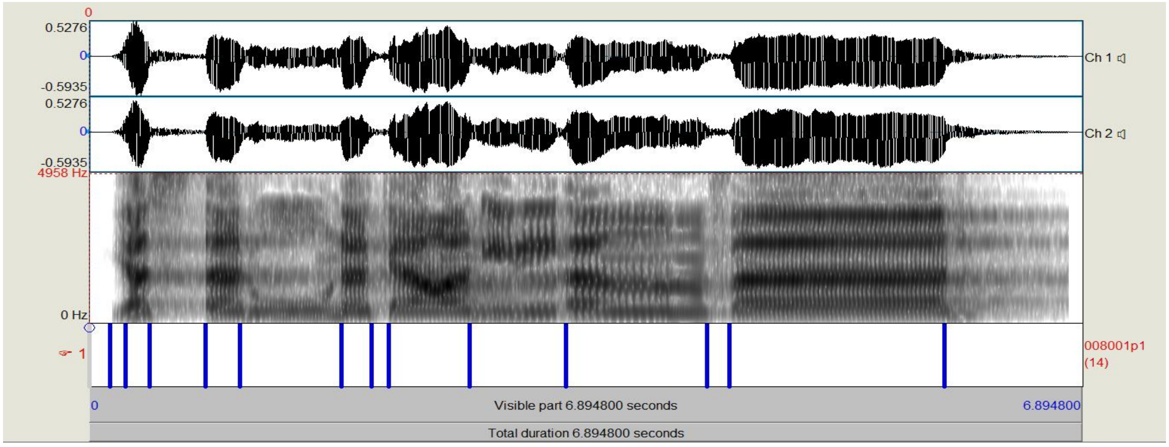


Figure 3. Segmentation result of part of the first verse of Surah El-Anfal

3.2. Processing of the system input files

Before initiating the rules detection procedure, several files need to be prepared, as outlined in Figure 4 and described below.

a) Dataset of verses (Ayat) texts: The verses under investigation are stored in a database. Each row in this database file includes the verse text, Sura name, Sura number, and verse number.
b) Dataset of verses sound files: This dataset comprises recorded sound files of Sheikh Al-Hosary, with each verse corresponding to a distinct sound file. In cases where the verse includes reading stops (silent patterns within the speech), it is stored in parts based on the number of stops. This step is essential for Tajweed rules that depend on the presence of these stops, such as the Accidental Madd rule, which must be defined before analysis.
c) Tajweed Rules: Each Tajweed rule under investigation is represented by a collection of text patterns that indicate its presence within Quranic verses. These text patterns, such as (اء - اَئ – ئِ) for the Connected Madd rule, are stored in a computerized form as a dataset. This structured dataset enables efficient text-based searches to detect and analyze Tajweed rules within the verses.
d) Audio Tajweed rules database: pre-detected Tajweed rules patterns from previous work [17] are stored for use in the pattern-matching step.
e) Initial Script files: Each verse is linked to a script file, a .txt file that initially contains the verse text and the total duration of the verse. The results of the segmentation process and the identified Tajweed rules within the verse audio file will be documented in this file. The script structure follows the Praat script format, compatible with the Praat toolkit – an accessible computer software package for speech analysis and phonetics [23].

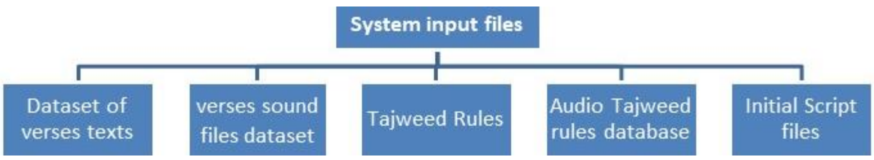


Figure 4. Input files of the rule's detection approach

3.3. System block diagram

The rules under investigation have been assigned unique abbreviated codes [17] for labeling the detected rules. The algorithm detects these rules sequentially, beginning with sub-Madd rules, followed by Noon/Meem Sakina rules, and finally, the natural Madd rule. The Tajweed rule detection process can be simplified as shown in Figure 5. It begins with a text-based search for the rule, followed by a speech-based search, and concludes with logging the detection results.

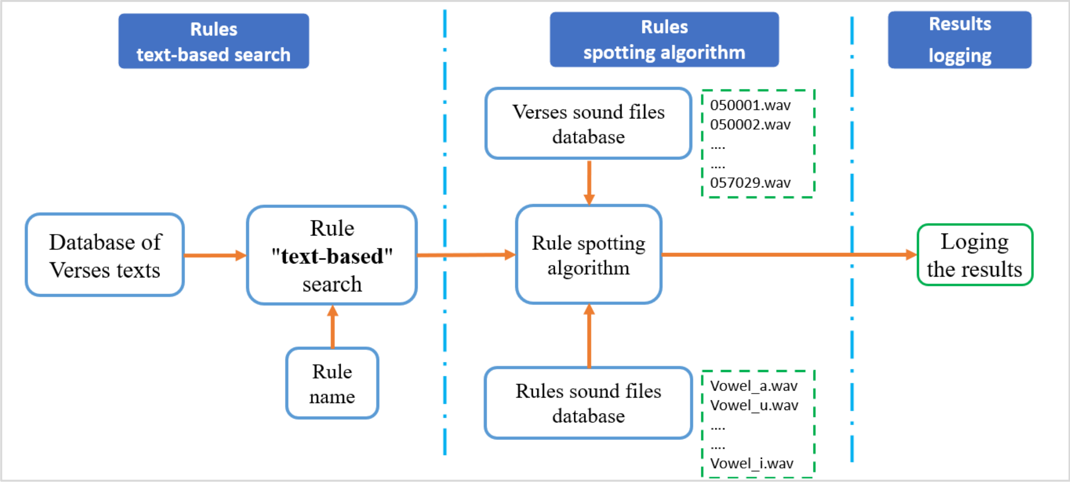


Figure 5. Tajweed rules detection flow using the pattern matching technique

3.3.1. Applied rules text-based search

Each Tajweed rule has a distinct textual representation in the written Quranic text. Since a single rule can have multiple textual forms, the search process identifies all possible variations of the rule within the verse. As shown in Figure 6, the text-based search identifies the occurrences of the Tajweed rule within the selected verse, providing its repetition count, positions in the verse text, and additional supporting information such as the rule code and the number of movements.

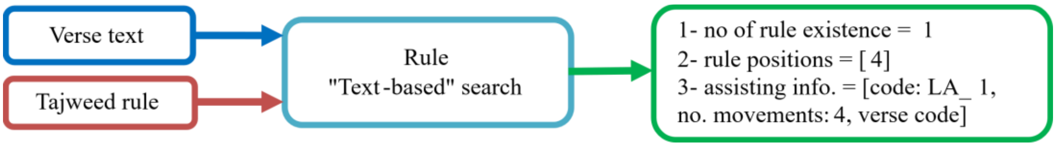


Figure 6. text-based search for Tajweed rules

3.3.2. Rules spotting algorithm

An initial step is performed to load relevant files and extract necessary information. The text-based search results provide the type of Tajweed rule being identified, its number of occurrences in the verse, and the number of movements associated with it based on the recitation of the Sheikh. The corresponding audio file of the verse is loaded from the verses' audio dataset, along with the verse script file where the detection results will be recorded. Additionally, the Tajweed reference audio patterns for the rule are retrieved from the Tajweed patterns database.

The spotting algorithm is illustrated in Figure 7, It begins by segmenting the verse's audio file into smaller segments. As previously mentioned, the audio frames of the Tajweed rules under investigation have a high correlation with each other, which allows them to separate from the preceding and following audio segments during the segmentation process.

Next, the audio features, specifically 14 MFCCs, are extracted from each segment of the verse's audio and the reference Tajweed patterns. The audio files are first divided into frames, each 20 milliseconds long, with a 10-millisecond overlap. A Hamming window is applied to these frames before extracting the 14 MFCC features.

Once the features are extracted, a comparison is performed between the features of one reference Tajweed pattern and the core part of each audio segment of the verse. The similarity is measured using the mean square error (MSE) function, where the segment with the lowest error score represents the best matching pattern. This process is repeated for each reference Tajweed pattern, identifying the best match in each iteration. Finally, using the election method, the segment selected the most times across all iterations is considered the best matching pattern. The spotting algorithm outputs the start and end times for each detected Tajweed rule, providing precise time boundaries for its occurrence within the verse.

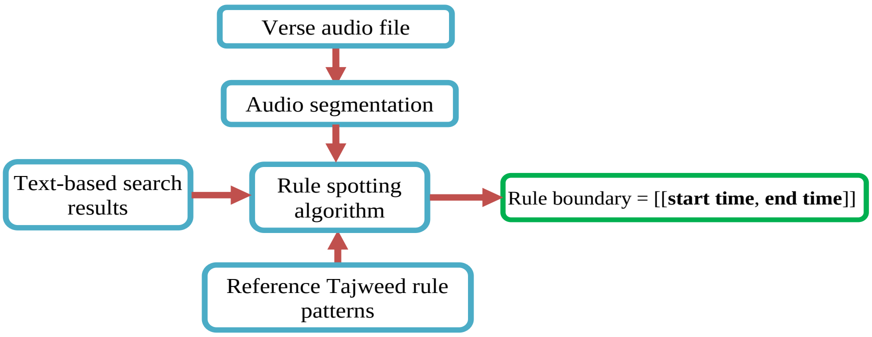


Figure 7. Speech-based search for Tajweed rules

3.3.3. Results logging

A Quranic verse can contain multiple Tajweed rules. For each detected rule, the verse script file is updated with the detection positions within the verse's audio file, specifying the start and end times of the detected patterns along with the corresponding rule code. To facilitate result validation, the verse script file is formatted for compatibility with the Praat interface, which provides a powerful visualization of both the speech signal and its annotations. As an example, Figure 8 illustrates the detection results of the Natural Madd rule (u:), the Noon Sakinah rule (hiding N_4_1), and the Accidental Madd rule (LA_4) in verse 1 of Surah Al-Anfal. The upper part of the figure represents the verse’s speech file in the time and frequency (Spectrogram) domains, while the lower part visualizes the script file information, highlighting the detected rules and allowing for isolated playback of each detected segment.

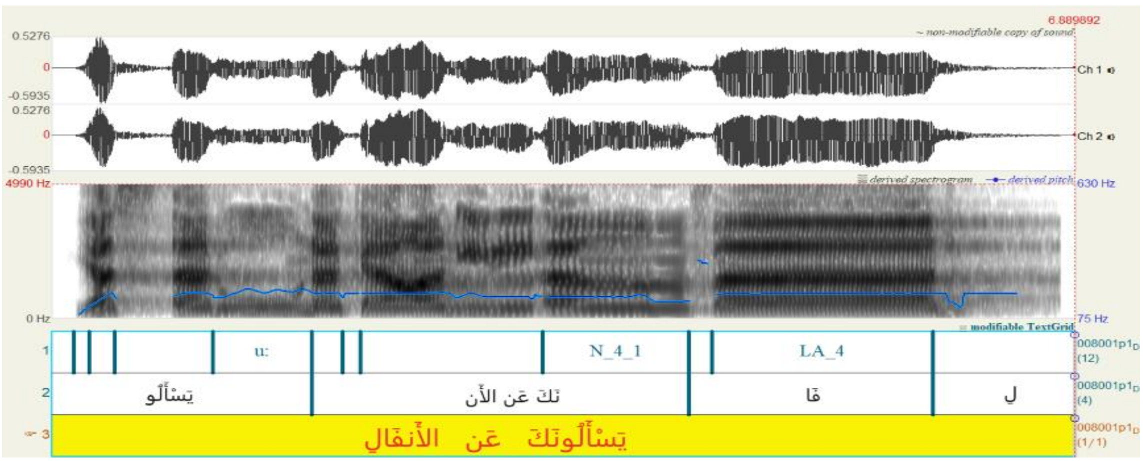


Figure 8. Detection results in part of the first verse of Surah El-Anfal

4. RESULTS AND DISCUSSION

4.1. Rules detection

The pronunciation of a Tajweed rule can vary across different recitation contexts, making its accurate detection a challenging task. This proposed pattern-matching approach is based on the hypothesis that having a diverse database of previously detected Tajweed rules across various recitation contexts enhances the accuracy of detecting new occurrences. To test this hypothesis and evaluate the effectiveness of the proposed pattern-matching approach, we utilize previously detected Tajweed rules to identify the same rules in a different Quranic dataset. The newly detected rules are then added to the Tajweed rule database, enriching it with more variations. This enriched database is subsequently used to search for the same rules in another dataset, allowing us to assess the impact of database expansion on the accuracy and efficiency of Tajweed rule detection.

The Audio Tajweed Rules database used in this study was imported from a previous work [17], where Tajweed rules were detected using the traditional pattern-matching approach. That study focused on the first 40 verses of Surah Al-Anfal from Sheikh Al-Hosary's recordings, achieving an accuracy of 93.8%. In this work, the proposed pattern-matching approach was applied to the remaining 35 verses of Surah Al-Anfal, resulting in a higher detection accuracy of 95.9%, as shown in Table 3. This improvement demonstrates the effectiveness of the proposed method and supports the research hypothesis.

Table 4 provides a detailed breakdown of the detected Tajweed rules within the 35 verses of Surah Al-Anfal. Each row in the table lists a detected Tajweed rule along with its frequency of occurrence, the number of movements of each Tajweed rule, the average duration of the rule's pronunciation, and the mean duration of a single movement.

Table 3. Detection accuracy of the detected rules in the latest 35 verses of Surah El-Anfal

Rule type	Total number of rules	The new proposed detection algorithm	
		Number of detection errors	detection accuracy
Madd rules (duration >= 4 movements)	86	0	100%
Natural Madd	322	17	94.7%
Noon /Meem Sakinah and Tanween	158	6	96.2%
All detected rules	566	23	95.9%

Table 4. Detection results of the latest 35 verses of Surah El-Anfal of Sheik Al-Hosary records

Detected rule name		Number of detections	Number of movements	The mean time of rule duration (seconds)	The mean time of one movement (second)
Connected Madd (المد المتصل)		13	4	1.55	0.39
Disconnected Madd (المد المنفصل)		31	4	1.54	0.38
Necessary Madd (المد اللازم)		1	6	2.72	0.45
Accidental Madd (المد العارض للسكون)		41	4	1.48	0.37
Natural Madd (المد الطبيعي)		322	2	0.58	0.29
"Meem Sakinah" (الميم الساكنة)	Hiding (الإخفاء)	6	2*	0.828	0.414
	Insertion (الإدغام)	18	2*	0.822	0.411
"Noon Sakinah and Tanween" (النون الساكنة والتنوين)	Merging with Ghunnah (الإدغام بغنة)	72	2*	0.851	0.425
	Change (الإقلاب)	7	2*	0.787	0.394
	Hiding (الإخفاء)	55	2*	0.759	0.3798

* Approximately equal 2 movements – two-time durations of one vowel

The audio Tajweed rules database has been expanded to include the detection results from the 35 verses of Surah Al-Anfal, now totaling 1,053 rules. This enhancement enriches the database with a broader variety of rule patterns. Following this update, the proposed detection approach was applied to a different dataset, specifically Chapter 30 (جزء عم) of the Holy Quran.

Tables 5 and 6 present the detection results for Chapter 30, where the overall detection accuracy increased to 97.1%, surpassing the accuracy achieved on the last 35 verses of Surah Al-Anfal. This improvement highlights the impact of an enriched Tajweed rules database, demonstrating that the accuracy of rule detection in a different Quranic dataset improves as the dataset grows with newly detected rules, aligning with the research hypothesis.

The results highlight the significance of this research in accurately detecting Tajweed rules within Quranic recitation. However, a key limitation of this method is its reliance on existing patterns in the database. If a detected rule pattern does not have a similar match in the database, the pattern-matching approach fails to identify it. This limitation can be mitigated by continuously enriching the database with diverse Tajweed patterns from different recitations. Expanding the dataset in this way will enhance the system's ability to detect rules more effectively, and addressing this limitation remains a key direction for future work.

Table 5. Detection accuracy of the detected rules in Chapter 30 (جزء عم)

Rule type	Total number of rules	The new proposed detection algorithm	
		Number of detection errors	detection accuracy
Madd rules (duration >= 4 movements)	281	1	99.6%
Natural Madd	861	26	97.0%
Noon/Meem Sakinah and Tanween	388	17	95.6%
All detected rules	1530	44	97.1%

Table 6. Detection results of Chapter 30 (جزء عم) of Sheik Al-Hosary records

Detected rule name		Number of detections	number of movements	The mean time of rule duration (seconds)	The mean time of one movement (second)
Connected Madd (المد المتصل)		57	4	1.684	0.413
Disconnected Madd (المد المنفصل)		70	4	1.627	0.407
Necessary Madd (المد اللازم)		3	6	2.943	0.49
Accidental Madd (المد العارض للسكون)		151	4	1.471	0.367
Natural Madd (المد الطبيعي)		861	2	0.694	0.347
"Meem Sakinah" (الميم الساكنة)	Hiding (الإخفاء)	9	2	0.914	0.457
	Insertion (الإدغام)	49	2	0.894	0.447
"Noon Sakinah and Tanween" (النون الساكنة والتنوين)	Merging with Ghunnah (الإدغام بغنة)	198	2	0.895	0.448
	Change (الإقلاب)	11	2	0.857	0.428
	Hiding (الإخفاء)	121	2	0.878	0.439

4.2. Consistency analysis

One of the contributions of this research to Tajweed education is the analysis of pronunciation time variations for the Tajweed rules examined in this study. The consistency analysis measures how consistently the reciter applies the same Tajweed rule in different positions within the same recitation. This analysis establishes a reference benchmark, allowing new learners to evaluate their proficiency by measuring their consistency against it.

It is normal and acceptable to find variations in the pronunciation time of a Tajweed rule across different recitation sessions. However, such variations should not occur within the same recitation session. Therefore, Surah Al-Anfal was selected for this analysis, as it was recited in a single continuous session. The detection results for the entire Surah were aggregated and are presented in Table 7, showing a total of 1,053 detected rules.

Figure 9 presents the consistency histograms of the natural Madd and Noon Sakinah & Tanween rules, highlighting variations in movement time duration. These variations follow a normal distribution curve with mean durations of 0.295 seconds for the natural Madd rule as shown in Figure 9(a) and 0.4 seconds for the Noon Sakinah & Tanween rules as shown in Figure 9(b).

The benchmark for Tajweed rule pronunciation consistency can be defined in two ways. The first approach involves comparing the variation distribution of a learner's recitation with the reference distribution obtained from the certified reciter, using histograms to visualize and assess similarity. The second, more simplified approach, calculates the percentage of variations that fall within ± one standard deviation of the normal distribution curve, representing approximately 68% of the data [24]. This method provides a straightforward measure of consistency, indicating how closely a learner's recitation aligns with the reference benchmark.

Table 7. Detection results of Surah El-Anfal of Sheik Al-Hosary records

Detected rule name		Number of detections	number of movements	The mean time of rule duration (second)	The mean time of one movement (second)
Connected Madd (المد المتصل)		28	4	1.56	0.39
Disconnected Madd (المد المنفصل)		58	4	1.52	0.38
Necessary Madd (المد اللازم)		4	6	2.69	0.45
Accidental Madd (المد العارض للسكون)		88	4	1.44	0.36
Natural Madd (المد الطبيعي)		603	2	0.59	0.295
"Meem Sakinah" (الميم الساكنة)	Hiding (الإخفاء)	11	2*	0.78	0.39
	Insertion (الإدغام)	28	2*	0.80	0.4
"Noon Sakinah and Tanween" (النون الساكنة والتنوين)	Merging with Ghunnah (الإدغام بغنة)	129	2*	0.82	0.41
	Change (الإقلاب)	8	2*	0.78	0.39
	Hiding (الإخفاء)	96	2*	0.76	0.38

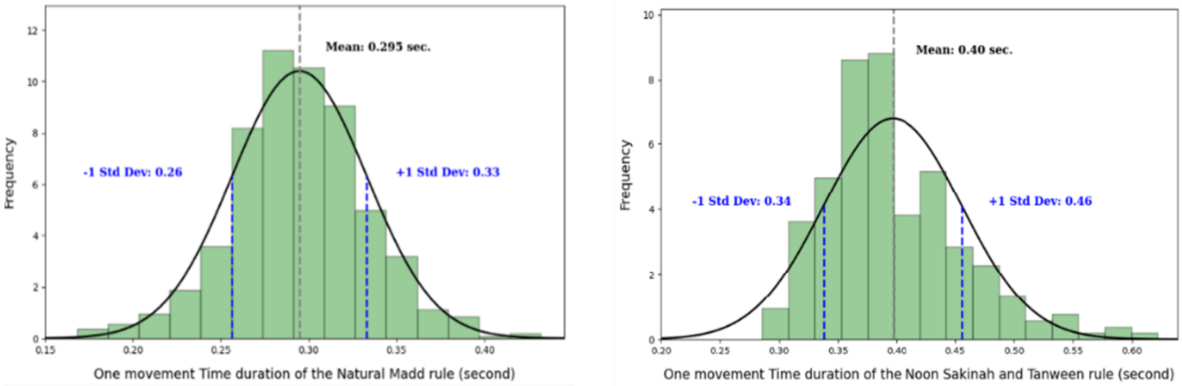
* Approximately equal 2 movements – two-time durations of one vowel

The percentage variation in movement time is determined by normalizing these variations to the mean movement time, as expressed in (2).

$$\%\text{movement } time\ variation = \frac{measured\ movement\ time - movement\ meantime}{movement\ meantime} * 100 \qquad (2)$$

For the natural Madd rule, the movement time varies between 0.26 seconds and 0.33 seconds within the boundaries of ± one standard deviation. Applying (2), the calculated percentage variation for the natural Madd rule is approximately 12.7%.

For the Noon Sakinah & Tanween rules, the movement time varies between 0.34 seconds and 0.46 seconds within the boundaries of ± one standard deviation. By applying (2), the calculated percentage variation for the Noon Sakinah & Tanween rules is approximately 13.5%.



(a) (b)

Figure 9. Consistency histogram of (a) Natural Madd and (b) Noon Sakinah and Tanween rules

The consistency study can be extended to more Tajweed rules in future work to establish a comprehensive benchmark for evaluating new learners. These results can be integrated into Tajweed learning applications, offering a structured validation method that differs from traditional speech recognition as it directly assesses pronunciation consistency.

4.3. Tajweed learning application

The proposed approach effectively generates a script file for each verse in Surah Al-Anfal and Chapter 30. It contains details of the detected Tajweed rules and their positions within the verse's audio file. These script files were then used to develop a Tajweed application designed to teach the identified rules to new learners. The application includes a user-friendly graphical user interface (GUI) developed in Python using the Tkinter package [25], [26]. Figure 10 displays the Tajweed App’s GUI, with numbered markers highlighting its various functions.

a) Verse Selection: The verse audio file can be chosen using field (1).
b) Script File Loading: After selecting a verse, its script file is loaded, and the verse text appears in the field (3).
c) Rule Selection: The user can choose the target Tajweed rule from field (2).
d) Audio Interaction:
   – Play Audio: The user can play the selected verse's audio using a button (4).
   – Visualize Speech: Pressing the button (5) draws the speech waveform in the time domain in the field (7).

Rule Detection: Pressing button (6) initiates the Tajweed rule detection. The application parses the script file to locate instances of the selected rule within the verse's audio and marks them in red in the field (7). The total number of detected rules is displayed in the field (8).

e) Rule Playback and Visualization: The learner can listen to the detected rule separately using the button (9). As the rule's sound plays, a moving cursor highlights its position in the field (7), synchronizing the visual marker with the audio. This feature helps learners recognize the rule's pronunciation and placement within the speech file.

The implemented application demonstrates the effectiveness of the generated Tajweed rules pattern database in teaching these rules to new learners.

Figure 10.1 Tajweed learning application

5. CONCLUSION

The proposed Tajweed rule detection method has demonstrated superior accuracy compared to the traditional pattern-matching approach. It achieved a detection accuracy of 95.9% on Surah Al-Anfal, surpassing the 93.8% accuracy of the traditional method. Moreover, when the Tajweed rules dataset was enriched with new detection results from Surah Al-Anfal, the accuracy further increased to 97.1% in Chapter 30 (جزء عم), highlighting the effectiveness of dataset expansion in improving detection performance.

The generated script files of the detected rules have proven valuable in logging their precise positions within the Quranic speech dataset. These files also facilitate the segmentation of Quranic verses into smaller utterances, providing a structured foundation for future Tajweed rule analysis. Additionally, the script files played a crucial role in developing a Tajweed learning application designed to assist new learners in properly reciting the Holy Quran by allowing them to visualize and listen to detected rules interactively.

The variation in rule pronunciation durations was analyzed using a reference reciter, Sheikh Al-Hosary, to establish a benchmark for pronunciation consistency. The computed consistency boundaries serve as a reference standard, enabling learners to assess their proficiency by comparing their recitation consistency against an accredited reciter. Furthermore, using the proposed detection method, the Tajweed rule database generated from this study can be leveraged to detect additional rules in new Quranic datasets. As this database expands, it can be integrated with advanced classification techniques, such as DNN, to further enhance the automated detection of Tajweed rules in Quranic recitations.

FUNDING INFORMATION

Authors state no funding is involved.

AUTHOR CONTRIBUTIONS STATEMENT

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Karim Aly Mohammad	✓	✓	✓	✓	✓	✓		✓	✓		✓			
Ahmed Hisham Kandil	✓	✓		✓	✓	✓	✓			✓		✓	✓	
Ahmed Mohamed EL-Bialy	✓	✓			✓		✓			✓		✓		
Sahar Ali Fawzy	✓	✓			✓		✓			✓		✓		

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author, [KAM], upon reasonable request.

REFERENCES

[1] T. M. Johnson and P. F. Crossing, "Projecting global religious populations, 2020–50," *Journal of Religion and Demography*, vol. 8, no. 1–2, pp. 124–153, 2021, doi: 10.1163/2589742X-12347114.

[2] K. Czerepinski, "Introduction to tajweed of the Qur'an," in *Tajweed Rules of the Quran, Part One*, Bangladesh: Darussalam Publication, 2000, pp. 23–24.

[3] C. Agarwal and P. Chakraborty, "A review of tools and techniques for computer aided pronunciation training (CAPT) in English," *Education and Information Technologies*, vol. 24, no. 6, pp. 3731–3743, Nov. 2019, doi: 10.1007/s10639-019-09955-7.

[4] S. Abdou *et al.*, "Computer aided pronunciation learning system using speech recognition techniques," *Proc. INTERSPEECH 2006 - ICSLP, Ninth International Conference on Spoken Language Processing*, Sep. 2006, pp. 849–852.

[5] S. E. Hamid, O. Abdel-Hamid, and M. Rashwan, "Performance tuning and system evaluation for computer aided pronunciation learning," *Proc. NEMLAR International Conference on Arabic Language Resources and Tools*, 2004.

[6] N. O. Balula, M. Rashwan, and S. Abdou, "Automatic speech recognition (ASR) systems for learning Arabic language and Al-Quran recitation: A review," *International Journal of Computer Science and Mobile Computing*, vol. 10, no. 7, pp. 91–100, 2021, doi: 10.47760/ijcsmc.2021.v10i07.013.

[7] M. Al-Marri, H. Raafat, M. Abdallah, S. Abdou, and M. Rashwan, "Computer aided qur'an pronunciation using dnn," *Journal of Intelligent and Fuzzy Systems*, vol. 34, no. 5, pp. 3257–3271, 2018.

[8] M. S. Elaraby, M. Abdallah, S. Abdou, and M. Rashwan, "A deep neural networks (DNN) based models for a computer aided pronunciation learning system," *Proc. Speech and Computer: 18th International Conference (SPECOM 2016)*, Aug. 2016, pp. 51–58.

[9] M. Abdallah, M. Al-Marri, S. Abdou, H. Raafat, M. Rashwan, and M. El-Gamal, "Improving Holy Qur'an recitation system using hybrid deep neural network-hidden markov model approach," *3rd International Conference on Islamic Applications in Computer Science And Technology*, Dec. 2015, pp. 1–8.

[10] M. S. Abdo, A. H. Kandil, A. M. El-Bialy, and S. A. Fawzy, "Automatic detection for some common pronunciation mistakes applied to chosen Quran sounds," *2010 5th Cairo International Biomedical Engineering Conference (CIBEC)*, Dec. 2010, pp. 219–222.

[11] L. R. Rabiner and R. W. Schafer, "The cepstrum and homomorphic speech processing," in *Theory and Applications of Digital Speech Processing*, 1st ed., Englewood Cliffs, NJ, USA: Prentice-Hall Inc., 1975.

[12] M. Awaid, A. H. Kandil, and S. A. Fawzy, "Audio search based on keyword spotting in Arabic Language," *International Journal of Advanced Computer Science and Applications (IJACSA)*, vol. 5, no. 2, pp. 128–133, 2014, doi: 10.14569/IJACSA.2014.050219.

[13] A. Quteishat, M. Younis, A. Qtaishat, and A. Abuhamdah, "Intelligent Arabic letters speech recognition system based on mel frequency cepstral coefficients," *International Journal of Electrical and Computer Engineering (IJECE).*, vol. 13, no. 3, pp. 3348–3358, Jun. 2023, doi: 10.11591/ijece.v13i3.pp3348-3358.

[14] T. M. H. Asda, T. S. Gunawan, M. Kartiwi, and H. Mansor, "Development of Quran reciter identification system using MFCC and neural network," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 1, no. 1, pp. 168–175, Jan. 2016, doi: 10.11591/ijeecs.v1.i1.pp168-175.

[15] F. Ahmad, S. Z. Yahya, Z. Saad, and A. R. Ahmad, "Tajweed classification using artificial neural network," *Proceedings of the 2018 International Conference on Smart Communications and Networking (SmartNets)*, Nov. 2018, pp. 1–4.

[16] A. A. Harere and K. A. Jallad, "Mispronunciation detection of basic quranic recitation rules using deep learning," May 10, 2023, *arXiv*: arXiv:2305.06429. doi: 10.48550/arXiv.2305.06429.

[17] K. A. Mohammad, A. H. Kandil, A. M. EL-Bialy, and S. A. Fawzi, "Automated detection and consistency analysis of tajweed recitation rules in the holy Quran," *Proceedings of the 2022 4th Novel Intelligent and Leading Emerging Sciences Conference (NILES)*, Oct. 2022, pp. 364–367.

[18] "Biography of Sheikh Mahmoud Khalil Al-Hosary," Al-Azhar Global Fatwa Center. Accessed: Dec. 01, 2023. [Online]. Available: https://www.azhar.eg/fatwacenter/fatwa/qadaya/PgrID/13003/PageID/4/ArtMID/7988/ArticleID/56402

[19] A. E. Sakran, S. M. Abdou, S. E. Hamid, and M. Rashwan, "A review: automatic speech segmentation," *International Journal of Computer Science and Mobile Computing (IJCSMC)*, vol. 6, no. 4, pp. 308–315, Apr. 2017.

[20] R. A. Ulfattah, S. N. Endah, R. Kusumaningrum, and S. Adhy, "Continuous speech segmentation using local adaptive thresholding technique in the blocking block area method," *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, vol. 18, no. 1, pp. 407–418, Feb. 2020, doi: 10.12928/telkomnika.v18i1.13958.

[21] M. Kalamani, D. S. Valarmathy, and S. Anitha, "Hybrid speech segmentation algorithm for continuous speech recognition," *International Journal of Applied Information and Communication Engineering*, vol. 1, no. 1, pp. 39–46, 2015.

[22] J. L. Rodgers and W. A. Nicewander, "Thirteen ways to look at the correlation coefficient," *The American Statistician*, vol. 42, no. 1, pp. 59–66, Feb. 1988, doi: 10.1080/00031305.1988.10475524.

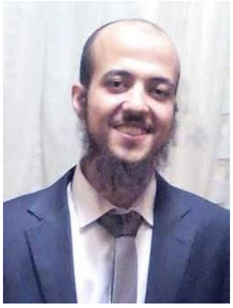
[23] "Praat: doing phonetics by computer," Accessed: Mar. 21, 2025. [Online]. Available: https://www.fon.hum.uva.nl/praat/

[24] D. S. Moore, G. P. McCabe, and B. A. Craig, "Density curves and normal distributions," in *Introduction to the Practice of Statistics*, 9th ed., New York, NY, USA: W. H. Freeman, 2016, p. 118.

[25] "tkinter — Python interface to Tcl/Tk," Python documentation. Accessed: Dec. 01, 2023. [Online]. Available: https://docs.python.org/3/library/tkinter.html

[26] "Welcome to Python.org," Python.org. Accessed: Dec. 01, 2023. [Online]. Available: https://www.python.org/

BIOGRAPHIES OF AUTHORS



Karim Aly Mohammad his B.Sc. and M.S. degrees from Ain-Shams University's Electronics and Communication Department in 2008 and 2015, respectively. From 2017 to 2021, he served as the research and development manager in the Medical Engineering Group, focusing on the development and manufacturing of medical devices. Following that, he joined Si-Vision Company, where he currently works in the research and development department, specializing in embedded systems development and integrated circuits (IC) testing. Mohammad is also pursuing a Ph.D. in Systems and Biomedical Engineering at Cairo University. His ongoing research focuses on pattern recognition in Arabic speech recordings, detecting Tajweed rules in Quranic audio files recited by renowned reciters, and applying modern AI classification techniques. He can be contacted at email: karimaly08@gmail.com.

Prof. Dr. Ahmed Hisham Kandil received his B.Sc. and M.S from the Systems and Biomedical Engineering, Cairo University (1980 and 1984). Dr. Kandil got his PhD from Systems Engineering - Case Western Reserve University, Cleveland, Ohio, USA (1991). Dr. Kandil was the General Manager of the SysNet Lab in Systems and Biomedical Engineering at Cairo University, for four years starting in 2004. In 2008, he was the General Manager of the Center of Advanced Software and Biomedical Consultation (CASBEC), in the Faculty of Engineering Cairo University. In 2012, he was the Chairman of the Biomedical and System Engineering department at the Shorouk Higher Engineering Institute for two years. Currently, he is a professor in Systems and Biomedical Engineering at Cairo University. His current research topics include the application of pattern recognition methodologies to medical image processing, optical character recognition for cursive arabic text, and arabic speech recognition. other research interests include robotics and nonlinear control for dynamical systems, arabic text to speech, computer assisted learning using multimedia, cephalometric, modeling, and stress analysis applications in Dentistry. He can be contacted at email: ahkandil@eng1.cu.edu.eg.

Prof. Dr. Ahmed Mohamed El-Bialy received his B.Sc. (1980) and M.S. (1984) in Systems and Biomedical Engineering from Cairo University and his Ph.D. in Systems Engineering from Case Western Reserve University, USA (1990). He served as Executive Manager of the Multimedia Lab at Cairo University's Center for Advanced Software Development and Applications, and in 2008, became Executive Manager of the Center of Advanced Software and Biomedical Consultation (CASBEC), Faculty of Engineering, Cairo University. From 2011 to 2013, he headed the Biomedical and Systems Engineering Department at the Shorouk Higher Engineering Institute. He then served as General Manager of CASBEC (2012–2014) and Dean of Cairo Higher Institute for Engineering, Computer Sciences, and Management (2014–2016). Currently, he is a full Professor at the Faculty of Engineering, Cairo University, and Head of Biomedical and Systems Engineering at Shorouk Institute. His research spans pattern recognition, nonlinear system analysis, neuromuscular systems, multimedia, 3D animation, and educational technology. His recent focus is on automating orthodontic clinics. Dr. El-Bialy has supervised over 150 theses and 200 senior projects and evaluated more than 320 theses and final projects. He can be contacted at email: Abialy@eng1.cu.edu.eg.

Prof. Dr. Sahar Ali Fawzy graduated from the Systems and Biomedical Engineering Department, Faculty of Engineering, Cairo University in 1985. She received her Master's and Ph.D. degrees in 1989 and 2002 respectively, from the Systems and Biomedical Engineering Department, Faculty of Engineering, Cairo University. She participated in different research groups and addressed diverse fields such as Deep Learning, Modeling and Finite Element Analysis, Telemedicine, Hospital Information Systems, Image processing, Arabic text-to-speech systems, and Speech and Speaker Recognition. She has several peer-reviewed research papers published in international journals and conferences. She supervised 8 Ph.D. and 20 Master's theses in the fields mentioned above. Graduation projects under her supervision were funded by ITIDA, STDF, and NTRA. 2019th graduation project titled "Aiding System for elderly, visually impaired and illiteracy" won the second prize in the Ministry of Communication, IBTIECAR competition, augmented reality section. She can be contacted at email: sfawzi@nu.edu.eg.