Arabic Speech Recognition for Quranic Words: Building an Audio Dictionary

Karim A. Mohammad, Sahar A. Fawzi, Ahmed M. El-Bialy, Ahmed H. Kandil

***Abstract*—This paper presents the development and implementation of an Automatic Speech Recognition (ASR)-based Audio Dictionary Database for Quranic words, integrating modern speech processing and recognition techniques. The system utilizes speech feature extraction, pattern recognition, and audio matching to detect and log Quranic words from recitations. A computational approach extracts all 19,058 unique dictionary words from the Quranic text, and a dedicated speech processing application has been developed to detect these words in Quranic recitations. The application employs enhanced speech pattern-matching techniques, successfully identifying all dictionary words in Chapter 30 (1,272 words) and spotting 22 Quranic words across the entire Quran in 373 occurrences, achieving 95% accuracy. The Audio Dictionary of Quranic Words is a novel contribution that addresses the shortage of publicly available datasets for researchers. This dataset can be used to support Tajweed instruction, advance Arabic speech technology, and serve as a valuable resource for linguistic research, ASR development, and speech-based AI applications.**

***Keyword*—Arabic Audio Processing, Quranic Audio Dictionary, Speech Keyword Spotting, Digital Quranic Studies, Keyword Detection.**

I. INTRODUCTION

APPROXIMATELY 422 million people speak Arabic, making it one of the five most spoken languages in the world [1]; however, Arabic ASR research has not received the same level of attention as other languages, such as English. Given that Arabic is also the language of the holy Quran, it holds significant importance, particularly for non-Arabic-speaking Muslims who rely on Quranic recitations for religious and educational purposes. This underscores the need for advancing ASR technologies tailored to Arabic, especially in Quranic studies and applications.

The Holy Quran is the central religious text of the Islamic community, followed by approximately 2 billion Muslims worldwide [2]. It is traditionally recited with precise phonetic rules known as Tajweed rules [3]. Accurate recitation is essential for religious observance and the preservation of the text's integrity. It consists of 30 chapters, each containing a collection of Surahs. Each Surah comprises a specific number of verses (Ayat), forming the structure of the sacred text.

With the increasing number of non-Arabic-speaking Muslims, there has been a growing demand for technologies that support Quranic learning, particularly Computer-Aided Pronunciation Training (CAPT) systems [4]. These systems enhance the learning process by providing accessible, flexible tools for mastering Quranic recitation.

Over the past decades, numerous overlapping efforts have been made to explore modern methods for supporting Quranic learning, generally adopting one of two main approaches. The first approach involves constructing a full speech recognition system. This began with statistical models such as Hidden Markov Models (HMM) [5], [6] and later evolved to incorporate advanced techniques such as Deep Neural Networks (DNN) [7], [8], [9], [10]. While these systems demonstrate high accuracy, their development necessitates skilled professionals' manual segmentation of extensive Quranic recitations, resulting in significant time and resource requirements. Consequently, these initiatives often rely on commercial funding, with the created databases typically being proprietary.

Another group of researchers, constrained by limited resources, has explored more simple methods for detecting Tajweed rules and Quranic words in Quranic recitations. These approaches typically involve template-matching techniques and basic DNN models, providing a simpler alternative to more complex systems.

The template-matching method compares the feature set of a specific pattern with that of an audio file to spot the location of the desired pattern within the audio. This technique is highly effective in identifying Tajweed rule patterns in Quranic recitations, as evidenced by various studies. For instance, Saber et al. [11] focused on the elevated and lowered pronunciation rules (التفخيم والترقيق) and analyzed eight common mistakes made by learners during recitation, achieving a verification accuracy of 95%. However, their approach was restricted to isolated Quranic words, making it unsuitable for analyzing entire verses. This system employed Mel-Frequency Cepstral Coefficients (MFCCs), a critical feature in speech processing that adjusts frequency scales to align with human auditory perception [12]. In a different study, Awaid et al. [13]

Manuscript received [1 may 2025]; revised [23 June 2025]; accepted [17 July 2025]; Date of publication [21 July 2025].

Karim A. Mohammad is with Systems and Biomedical Engineering Department, Faculty of Engineering, Cairo University, Giza, Egypt.(e-mail: karimaly08@gmail.com). (corresponding author)

Sahar A. Fawzi is with information Technology & Computer Science, Nile University, Giza, Egypt (email: sfawzi@nu.edu.eg).

Ahmed M. EL-Bialy, is with Systems and Biomedical Engineering Department, Faculty of Engineering, El-Shorouk Academy, Cairo, Egypt. (email: abialy@eng1.cu.edu.eg)

Ahmed H. Kandil is with Systems and Biomedical Engineering Department, Faculty of Engineering, Cairo University, Giza, Egypt. (e-mail ahkandil@eng1.cu.edu.eg)

also used MFCCs alongside Linear Predictive Coefficients (LPC) and other speech features to detect the locations of 15 Quranic words in full recitations. Their work resulted in a maximum detection accuracy of 94.2%, demonstrating the effectiveness of MFCCs as a distinguishing feature. These studies, along with others [14], [15], [16], [17], have demonstrated that MFCCs are among the most effective speech features for use in template-matching tasks.

Many studies have investigated the use of simple DNN models for detecting Tajweed rules in Quranic recitations. One such study examined the Noon Sakinah rule (حكم الإدغام), utilizing a dataset of 300 recitations. However, the model's accuracy was relatively low at 77.7%, primarily due to the limited dataset size [18]. In another effort, Omran et al. [19] applied MFCC features and a two-stage Convolutional Neural Network (CNN) to detect the Qalqalah rule (حكم القلقلة) using a dataset of 589 utterances. Their approach achieved a 92% accuracy in letter classification and over 90% accuracy in rule detection. Similarly, Harere et al. [20] employed a Long Short-Term Memory (LSTM) model to identify the disconnected Madd (المد المنفصل) and some Noon Sakinah rules. This model, trained on 1500 utterances from verse 109 of Surat Al-Maeda (سورة المائدة), achieved approximately 96% detection accuracy. Although these models show significant potential, their performance is often constrained by the small sizes of the datasets available for training.

There is a notable gap in publicly available datasets for Quranic speech processing. This research addresses this gap by creating an audio dictionary of Quranic words. This study utilized Quranic recitations by Sheikh El-Hosary, a highly esteemed scholar and one of the foremost authorities on Tajweed and Quranic recitation in the Islamic world [21]. The chosen narration is Hafs An Asim, the most widely used recitation style.

By "audio dictionary," we mean finding a phonetic equivalent for each word in the Holy Quran based on the recitations of an accredited Quran reciter. To clarify the concept of this research, Table I provides an example of the proposed structure of the Audio dictionary for the word Ibrahim "إبراهيم". The table shows the start and end times of the word "إبراهيم" in the audio file for each verse in which it appears in Sheikh El-Hosary's recitation, in addition to the Sura name, Sura number, verse number, verse text, and the corresponding recitation.

The main objective of this research is to build a comprehensive Audio Dictionary Database of Quranic words. This database is intended to:

1. Support computer-aided pronunciation learning (CAPL) applications for Tajweed teaching.
2. Provide a robust dataset for modern classification and recognition techniques.
3. Facilitate acoustic studies of Tajweed rules.

This paper is organized as follows: Section II outlines the system materials and methods. Section III presents the results and discussion. Finally, Section IV provides the conclusion and outlines future directions.

Table I
The proposed database structure of the Audio dictionary of the Quranic words.

Sura name	Sura number	Verse number	Verse text	Dictionary word	Start (sec.)	End (sec.)	Reciter
الصافات	037	109	سَلَٰمٌ عَلَىٰٓ إِبۡرَٰهِيمَ	إِبۡرَٰهِيم	4.496	8.006	El-Hosary
الأعلى	087	019	صُحُفِ إِبۡرَٰهِيمَ وَمُوسَىٰ	إِبۡرَٰهِيم	1.293	3.643	El-Hosary

II. Materials and methods

As stated in the introduction, the primary objective of this work is to develop an audio dictionary of Quranic words. This goal is accomplished through three main phases. The first phase involves constructing a text-based dictionary of Quranic words, carefully considering the unique characteristics of Quranic recitation. The second phase focuses on identifying and extracting these Quranic words from the corresponding audio recordings of verses. Finally, the third phase involves checking the spotting results and logging them in an appropriate format. These phases are illustrated in Fig 1 and described in the sections below.

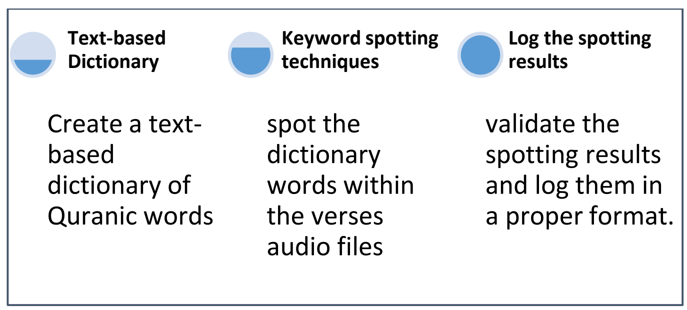

Fig 1 Proposed phases of constructing the Quranic Words audio dictionary.

A. Building a Text-Based Dictionary for the Quranic Words:

The initial phase of this work involved creating a text-based dictionary of Quranic words. While many Islamic researchers have developed text-based dictionaries of Quranic words, our approach is distinct, incorporating the phonetic equivalents of the words pronounced in Quranic recitation. A challenge arises from applying Tajweed rules, which sometimes result in phonetic overlaps between two or more Quranic words during recitation.

Hamzat al-wasl (همزة الوصل) is a primary cause of sound overlap between two consecutive Quranic words. For example, the word (العالمين) begins with hamzat al-wasl. When recited at the start of a verse, it is pronounced clearly as the Alif sound (أ). However, when preceded by another word, such as in (رب

العالمين), hamzat al-wasl is dropped in pronunciation, causing the two words to merge smoothly in speech. In the phrase (رب العالمين) there is no distinct sound corresponding to each individual word in this context. The words "رب" and "العالمين" cannot be phonetically isolated, as their pronunciation merges due to the presence of hamzat al-wasl. Consequently, it is necessary to represent "رب العالمين" as a single, isolated entity.

A text-based analysis algorithm was developed to identify dictionary words in all Quranic verses. The algorithm checks for the presence of Hamzat al-Wasl; if detected, it merges the affected word with the preceding word to form a single dictionary entry. Otherwise, each word is treated independently. This process yielded a dictionary containing 66,889 isolated words, including repeated occurrences, and 19,058 unique words. This comprehensive text-based dictionary serves as the foundation for developing the audio dictionary, ensuring complete coverage of the Quranic text.

Table II provides a sample of isolated words from Surat Al-Kawthar (الكوثر).

Table II
Example of the isolated words within Surat Al-Kawthar (سورة الكوثر)

Sura name	Sura number	Verse number	Verse text	No. of isolated words	Isolated word #1	Isolated word #2	Isolated word #3
الكوثر	108	001	إِنَّآ أَعۡطَيۡنَٰكَ ٱلۡكَوۡثَرَ	2	إِنَّآ	أَعۡطَيۡنَٰكَ ٱلۡكَوۡثَرَ	
الكوثر	108	002	فَصَلِّ لِرَبِّكَ وَٱنۡحَرۡ	3	فَصَلِّ	لِرَبِّكَ	وَٱنۡحَرۡ
الكوثر	108	003	إِنَّ شَانِئَكَ هُوَ ٱلۡأَبۡتَرُ	3	إِنَّ	شَانِئَكَ	هُوَ ٱلۡأَبۡتَرُ

Each row in Table II contains information about the sura name, sura number, verse number, verse text, the number of isolated words, and additional cells listing the isolated words. As shown in the first verse, the word (الكوثر) begins with a Hamzat al-Wasl and is preceded by another word (أعطيناك), so it merges with it to form a single dictionary word: (أعطيناك الكوثر). The same applies in the third verse with the phrase (هو الأبتر). Conversely, words that do not begin with Hamzat al-Wasl are treated as independent units and can be directly matched against dictionary entries without merging.

B. Keyword Spotting Techniques

Keyword spotting is the crucial component of building the audio dictionary of Quranic words. A new application has been developed to simplify the process of spotting Quranic words within the verses' audio files, as shown in Fig 2.

The application was developed in Python using the Tkinter package for the GUI and was tested on a standard PC with an Intel Core i7 processor and 16GB RAM running Windows 11.

The application utilizes the pattern-matching approach with many enhanced detection techniques. The Mel-Frequency Cepstral Coefficients (MFCCs) are used for feature extraction of the audio files, a well-established method in speech processing due to its effectiveness in capturing the phonetic properties of speech [12].

The main App features are:

1. Providing the necessary keyword detection techniques to accurately identify dictionary words.
2. Validate the detected results.
3. Refinement of the spotted dictionary words.
4. Log the detected results in an appropriate format.
5. User-friendly GUI.

Fig 2 The graphical user interface of the Keyword spotting application

The application's directory is organized into four primary subdirectories, each serving a specific purpose. These subdirectories are described as follows:

1. Quranic Speech Database: This subdirectory contains audio files of Sheikh El-Hosary's Quranic recitation, with one audio file corresponding to each verse of the Quran.
2. Quranic Words Spotting Application: This subdirectory contains the source code for the application created to identify Quranic words in audio recordings.
3. Quranic Words Spotting Results: This section contains the results generated by the application.
4. Quranic Words Spotting Resources: This subdirectory is divided into two categories:
   - Text Resources: Includes the dictionary words under study (which can be changed as needed), a database of Quranic verse texts, and their corresponding isolated words.
   - Speech Resources: Includes reference audio files for the words being studied.

The application workflow consists of four main modules, as illustrated in Fig 3 and described in the following sections.

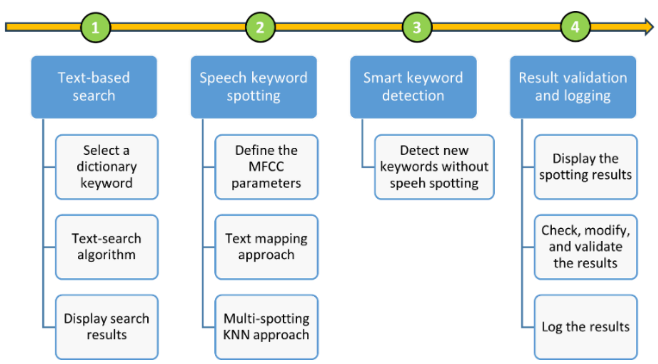

Fig 3 Workflow of the keyword spotting application

Text-Based Search:

Fig 4 illustrates the application's text-based search process, which operates as follows:

1. The app loads the Quranic dictionary words from the text resources subdirectory and displays them on the application GUI.
2. The user selects a word from the displayed list and initiates the search by pressing the "Start search" button. This action triggers a text-based search within the 'Isolated Words' columns, as shown in Table II. This approach allows the search process to occur not directly within the raw verse text, as some words may not appear as phonetically isolated units but instead within the predefined dictionary words associated with each verse.
3. The results of the text-based search are displayed. Subsequently, the user can select one of the search results to proceed with a speech-based search.

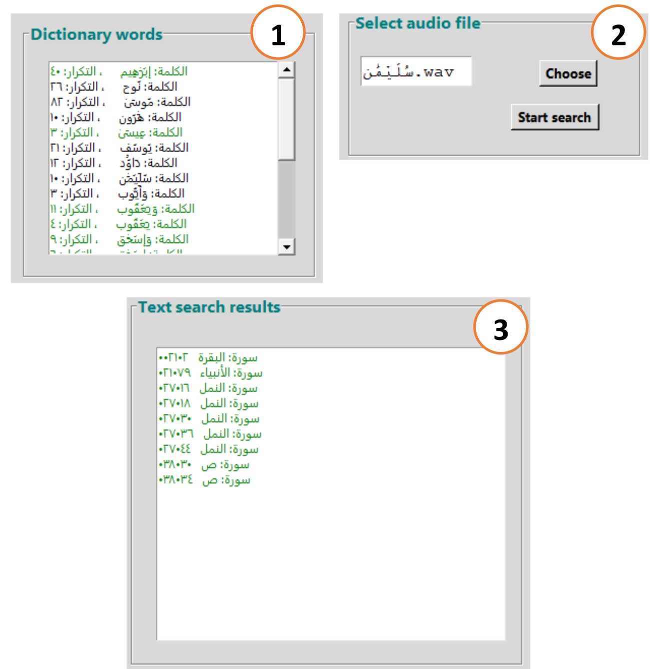

Fig 4 The application interface of the text-based search module

Speech Keyword Spotting:

This section focuses on the primary functionality of the application: identifying the precise location of a specified Quranic word within verses that contain it. The application is designed to offer users flexibility by allowing them to configure various parameters and apply enhancement techniques to optimize the spotting process, as shown in Fig 5.

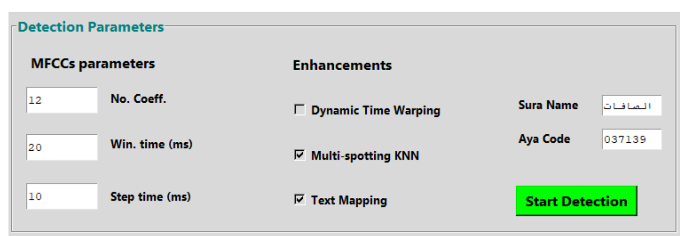

Fig 5 The application interface for setting the speech spotting parameters

Basic spotting technique: The pattern matching technique is used to identify Quranic words in the recorded audio files of the Quran. Here, the MFCC feature set of the desired Quranic word is compared with the MFCC feature set of the verse containing the word being searched. This process continues until the best matching location is found, thereby completing the spotting process. Through the application interface, users can configure the MFCC parameters by selecting the window duration, the overlap time between consecutive windows, and the desired number of coefficients, as illustrated in Fig 5.

The best matching location is determined by the Euclidean distance between the Quranic word MFCC and the verse audio file MFCC, as shown in equation (1). The lowest calculated distance indicates the highest similarity, representing the best matching position.

$$Euclidean\ distance = \sqrt{\sum_{i=1}^{n}(qi - pi)^2} \qquad (1)$$

Which qi and pi are elements in the MFCC feature matrices of the reference word and the Quranic verse, respectively. n represents the total number of elements in these feature matrices.

Text mapping approach: The algorithm attempts to estimate the target word's location within the audio file based on its position in the Quranic verse text while considering Tajweed rules.

This approach first involves identifying the text of the Quran before and after the target word, determining the total number of characters in the text before and after the word, and assigning a recitation duration of 1 unit to each character. It also considers Tajweed rules that affect the pronunciation duration of the letters, including different types of Madd rules, and incorporates the durations of these elongations into the calculations.

Subsequently, alignment is performed between the actual recitation time duration of the verse and the calculated phonetic units surrounding the Quranic word. This process allows for estimating the position of the word within the audio file of the verse. Finally, the algorithm adds a time boundary around the expected position to ensure the word is located within this boundary.

This approach is particularly useful for longer and medium-length verses, where the recitation time usually exceeds one minute. After applying this method, the app searches for the word within the boundary defined by the text-mapping approach, which is typically around 20 seconds. By doing so, the time required for detection is reduced by limiting the search area to a boundary smaller than the total duration of the verse, eliminating detection error for the same reason.

Find word boundary: This method leverages previous search results for words from the audio dictionary. If the target word is situated between two words whose locations have been previously identified, it implies that the target word is definitively located between these two words. Consequently, a smaller search boundary can be established for locating the word instead of searching through the entire audio file of the verse.

Multi-spotting using K-Nearest Neighbors (KNN): The standard spotting approach relies on comparing a reference word pattern with the verse audio file. However, the pronunciation of words can vary from one instance to another for the same word due to various factors, including differences in the recitation session and the emotional state of the reciter. This variation can alter the word's pronunciation and, consequently, the associated feature set, making the spotting process more challenging.

To address this phenomenon, a multi-spotting approach has been adopted, involving comparing multiple reference patterns of the same word from previously identified detections with the verse audio file during the spotting process. To enhance this method's efficiency, the previously identified locations of the word are clustered into K clusters using the KNN approach, a well-known clustering technique that can group a set of data into K clusters based on its feature set [22].

Fig 6 is an example of clustering 10 detected locations of the word "إبراهيم" in 10 verses into 5 clusters based on the MFFC features of these words. The approach employs a representative word from each cluster during the spotting process. This allows us to select a limited number from a previously identified word, considering the pronunciation variability between alternative locations. Consequently, this enhances the detection accuracy of the spotting process.

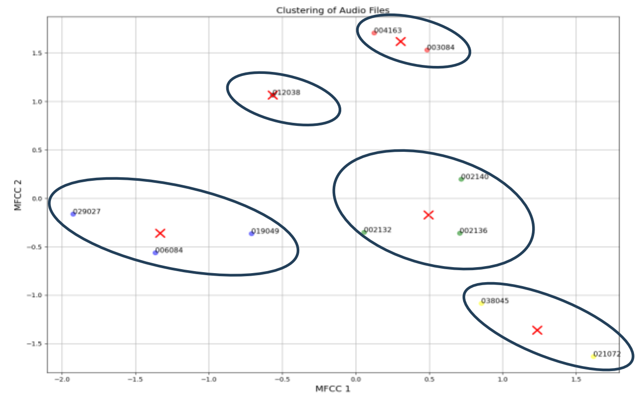

Fig 6 Clustering 10 detected locations of the word "إبراهيم" into 5 clusters

Smart keyword detection: This proposed method aims to detect dictionary words within the audio files of Quranic verses without relying on comprehensive speech spotting. Table III illustrates three expected scenarios of this smart keyword detection approach. The shaded cells in the table represent pre-detected words, while the unshaded cells indicate undetected words. The three scenarios are outlined as follows:

1. The word being detected is the first word in the verse, while the second word has been detected previously.
2. The word being detected is the last word in the verse, and the penultimate word has been previously detected.
3. The word being detected is situated between two previously identified words.

Although this approach seems relatively simple, it is very effective for identifying new words without needing complete speech detection.

Table III
Examples of the smart keyword detection approach

Sura name	Sura number	Verse number	Verse text	Isolated word #١	Isolated word #٢	Isolated word #٣	Isolated word #٤
الفاتحة	٠٠١	00٤	مَٰلِكِ يَوۡمِ ٱلدِّينِ	مَٰلِكِ	يَوۡمِ ٱلدِّينِ		
الفاتحة	٠٠١	00٥	إِيَّاكَ نَعۡبُدُ وَإِيَّاكَ نَسۡتَعِينُ	إِيَّاكَ	نَعۡبُدُ	وَإِيَّاكَ	نَسۡتَعِينُ

Result validation and logging: One of the application's key features is its built-in capability to validate results within the same interface. This feature enables users to review and refine the keyword spotting results as needed, ensuring a seamless and integrated experience. Fig 7 illustrates the application's interface designed for the validation and logging process.

Fig 7 The application interface of the validation and logging module.

The GUI provides a variety of features and options to enhance result validation and refinement, including:

1. Audio playback options: Users can listen to the complete audio file of the verse, the detected word, and the corresponding reference word to assess the accuracy of the spotting results.
2. Visual adjustment tools: Red cursors are displayed on the interface to indicate the start and end positions of the detected word. Users can manually modify these cursor positions to refine the result if necessary. A zoom-in feature is available to assist users in making precise adjustments.
3. Multiple match review: The application presents the top three matching results, allowing users to switch between them and validate the most accurate match. After the user validates the spotting results, the next stage involves logging these results. In this phase, the application records each detected word in a dedicated Excel sheet assigned to that word, ensuring an organized and systematic record of its respective spotting results. This intuitive interface provides a reliable and user-friendly way to validate, refine, and log keyword spotting results, enhancing the overall functionality and usability of the application.

III. RESULTS AND DISCUSSION

The application's proficiency was evaluated using two distinct methods.

1. Spotting words distributed across Quranic verses: This approach evaluated the application's accuracy in detecting specific words dispersed throughout various Quranic recitations. This test offers insight into the app's effectiveness in spotting words that occur in different contexts.
2. Spotting words from Chapter 30 of the Quran: This method assessed the application's reliability and efficiency in creating a comprehensive audio dictionary of Quranic words. By concentrating on a single chapter, this test emphasizes the app's capability to consistently identify and catalog words in a systematic manner.

The first evaluation approach utilized a set of 22 Quranic words, selected as proper nouns distributed across various Quranic verses. The speech reference words corresponding to these 22 words were manually extracted from specific Quranic verses using the application.

Table IV presents a summary of the results, detailing the detected proper nouns, the frequency of each word's occurrence within Quranic verses, the number of correct detections, the number of false detections, and the calculated accuracy for each word individually. This evaluation provides a comprehensive insight into the application's accuracy in detecting Quranic words dispersed throughout recitations.

Table IV
spotting results of 22 Quranic words

Quranic words	Total detections	True detections	False detections	Accuracy
مُحَمَّد	4	4	0	100%
إِبْرَٰهِيم	40	40	0	100%
نُوح	26	23	3	88%
مُوسَىٰ	82	81	1	99%
عِيسَىٰ	3	3	0	100%
هَٰرُون	10	9	1	90%
يُوسُف	19	19	0	100%
دَاوُۥد	12	10	2	83%
سُلَيْمَٰن	10	10	0	100%
ءَادَم	14	14	0	100%
وَيَعْقُوب	11	11	0	100%
وَإِسْحَٰق	9	9	0	100%
زَكَرِيَّا	3	3	0	100%
يَحْيَىٰ	4	4	0	100%
وَإِسْمَٰعِيل	9	6	3	67%
وَٱلْيَسَعَ	2	2	0	100%
يُونُس	2	2	0	100%
لُوطًا	4	3	1	75%
هُود	3	2	1	67%
شُعَيْبًا	6	5	1	83%
وَذَا ٱلْكِفْلِ	2	2	0	100%
فِرْعَوْن	62	58	4	94%

All words	**337**	**320**	**17**	**95%**

As illustrated in Table IV, the application successfully detected 320 occurrences of the words under study, achieving an overall spotting accuracy of 95%. Variations in pronunciation of the same word across different recitation sessions emerged as the primary source of false detections. The implementation of the multi-spotting KNN approach substantially mitigated this issue. Nevertheless, in some cases, specific positions lacked similarly pronounced words within the database's previously detected entries, leading to occasional false detections.

The relatively low accuracy in some words as (هود) does not, by itself, indicate a significant variation in the effectiveness of the proposed system, the false detection count for the word (هود) is only 1, the overall accuracy is 67% because the total number of detections is 3. Hence, a single error represents a large proportion of the total. In contrast, for the word (موسى), one false detection out of 82 results in a much higher accuracy of approximately 99%.

The second evaluation approach focused on the application's ability to detect all words within Chapter 30 (جزء عم) of the Holy Quran. This chapter holds unique significance in the context of Quranic learning, as it is traditionally used as the foundational chapter for novice learners. In this evaluation, all words within Chapter 30 were successfully detected, amounting to 1,991 occurrences and 1,272 unique words. The detected words represent 6.7% of the comprehensive audio dictionary, which consists of 19,058 words. These results highlight the application's effectiveness as a tool for constructing a detailed and accurate audio dictionary of the Quran.

IV. CONCLUSION

In this work, we applied promising speech processing approaches to develop an Arabic ASR-based Audio dictionary. A graphical user interface (GUI) application was implemented, allowing users to employ various speech recognition techniques to achieve high detection accuracy of Quranic words in audio recordings.

The system accurately detected all words in Chapter 30 (جزء عم) and successfully identified 22 Quranic words throughout the entire Quran in 373 occurrences, achieving 95% accuracy. This research provides a valuable dataset for speech recognition, Tajweed learning, and Quranic recitation studies, contributing to advancing Arabic ASR, keyword spotting, and digital Quranic studies.

Future work will focus on expanding the dataset to include multiple reciters, enhancing detection accuracy, and integrating the system into educational applications to improve the learning of Quranic recitation for non-Arabic speakers.

REFERENCES

[1] "List of countries and territories where Arabic is an official language," *Wikipedia*. Feb. 07, 2025. Accessed: Feb. 19, 2025. [Online]. Available: https://en.wikipedia.org/w/index.php?title=List_of_countries_and_territories_where_Arabic_is_an_official_language&oldid=1274450185

[2] T. M. Johnson and P. F. Crossing, "Projecting Global Religious Populations, 2020–50," *J. Relig. Demogr.*, vol. 8, no. 1–2, pp. 124–153, 2021, doi: 10.1163/2589742X-12347114.

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

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

[5] S. E. Hamid and M. Rashwan, "Automatic Generation of Hypotheses for Automatic Diagnosis of Pronunciation Errors," in *International Conference on Arabic Language Resources and Tools*, Cairo, Egypt: NEMLAR, Sep. 2004, pp. 140–143.

[6] S. Abdou, S. E. Hamid Morsi, M. Rashwan, A. Samir, O. Abdel-Hamid, M. Shahin, and W. Nazih, "Computer aided pronunciation learning system using speech recognition techniques," in *Proc. INTERSPEECH 2006 - ICSLP, Ninth International Conference on Spoken Language Processing*, Pittsburgh, PA, USA: International Speech Communication Association (ISCA), Sep. 2006, pp. 849–852. doi: 10.21437/Interspeech.2006-287.

[7] 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," in *3rd International Conference on Islamic Applications in Computer Science And Technology*, Turkey: International Journal on Islamic Applications in Computer Science & Technology, Dec. 2015, pp. 1–8. Accessed: Dec. 01, 2023. [Online]. Available: https://www.sign-ific-ance.co.uk/index.php/IJASAT/article/view/1337

[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," in *Proc. Speech and Computer: 18th International Conference (SPECOM 2016)*, A. Ronzhin, R. Potapova, and G. Németh, Eds., in Lecture Notes in Computer Science. Budapest, Hungary: Springer International Publishing, Aug. 2016, pp. 51–58. doi: 10.1007/978-3-319-43958-7_5.

[9] M. Al-Marri, H. Raafat, M. Abdallah, S. Abdou, and M. Rashwan, "Computer aided qur'an pronunciation using dnn," *J. Intell. Fuzzy Syst.*, vol. 34, no. 5, pp. 3257–3271, 2018.

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

[11] M. Saber and A. H. Kandil, "Semi-Automatic Segmentation System for Syllables Extraction from Continuous Arabic Audio Signal," *Int. J. Adv. Comput. Sci. Appl. (IJACSA)*, vol. 7, no. 1, pp. 535–540, 2016, doi: 10.14569/IJACSA.2016.070173.

[12] 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.

[13] M. Awaid, A. H. Kandil, and S. A. Fawzy, "Audio Search Based on Keyword Spotting in Arabic Language," *Int. J. Adv. Comput. Sci. Appl. (IJACSA)*, vol. 5, no. 2, pp. 128–133, 2014, doi: 10.14569/IJACSA.2014.050219.

[14] A. Quteishat, M. Younis, A. Qtaishat, and A. Abuhamdah, "Intelligent Arabic letters speech recognition system based on mel frequency cepstral coefficients," *Int. J. Electr. Comput. Eng. (IJECE).*, vol. 13, no. 3, pp. 3348–3358, Jun. 2023, doi: 10.11591/ijece.v13i3.pp3348-3358.

[15] 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," in *Proc. 2022 4th Novel Intell. Leading Emerg. Sci. Conf. (NILES)*, Giza, Egypt: IEEE, Oct. 2022, pp. 364–367. doi: 10.1109/NILES56402.2022.9942371.

[16] K. A. Mohammad, A. H. Kandil, A. M. EL-Bialy, and S. A. Fawzi, "Consistency Analysis of Madd rules in the Holy Quran," in *Proc. 2021 3rd Novel Intell. Leading Emerg. Sci. Conf. (NILES)*, Giza, Egypt: IEEE, 2021, pp. 291–294. doi: 10.1109/NILES53778.2021.9600505.

[17] T. M. H. Asda, T. S. Gunawan, M. Kartiwi, and H. Mansor, "Development of Quran Reciter Identification System Using MFCC and Neural Network," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 1, no. 1, pp. 168–175, Jan. 2016, doi: 10.11591/ijeecs.v1.i1.pp168-175.

[18] F. Ahmad, S. Z. Yahya, Z. Saad, and A. R. Ahmad, "Tajweed Classification Using Artificial Neural Network," in *Proc. 2018 Int. Conf. Smart Commun. Netw. (SmartNets)*, Yasmine Hammamet, Tunisia: IEEE, Nov. 2018, pp. 1–4. doi: 10.1109/SMARTNETS.2018.8707394.

[19]D. Omran, A. Kandil, S. Samy, A. ElBialy, and S. Fawzi, "Convolutional Neural Networks for speech recognition case study: Recitation Rules of the holy Quran," *MSA Eng. J.*, vol. 2, no. 4, pp. 1–12, Dec. 2023, doi: 10.21608/msaeng.2023.225120.1335.
[20]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.
[21]"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
[22]K. Taunk, S. De, S. Verma, and A. Swetapadma, "A Brief Review of Nearest Neighbor Algorithm for Learning and Classification," in *2019 International Conference on Intelligent Computing and Control Systems (ICCS)*, May 2019, pp. 1255–1260. doi: 10.1109/ICCS45141.2019.9065747.