Journal of Theoretical and Applied Information Technology

31st August 2023. Vol.101. No 16
© 2023 Little Lion Scientific

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

QURAN RECITER IDENTIFICATION: TECHNIQUES AND CHALLENGES

MOHAMMED ALATIYYAH[1]

[1]Assistant Professor, Department of Computer Science, College of Sciences and Humanities in Aflaj,

Prince Sattam Bin Abdulaziz University, Saudi Arabia

E-mail: [1]m.alatiyyah@psau.edu.sa

ABSTRACT

Speech-based intelligent systems are gaining an increasing popularity and importance due to their wide range of applications in our daily life. Most of the research efforts within speaker identification target the English language. However, efforts that target the Arabic language and the holy Quran are still limited. For Muslims, the Holy Quran is the main religious book of Islam. The Holy Quran verses must be read according to very restricted rules known as "Tajweed' to guarantee the correct pronunciation of verses. The task of identifying the holy Quran reciter or reader based on many features in the corresponding acoustic wave is known as Quranic reciter identification process. It is considered a more challenging task than other speaker identification tasks as it depends on "Tajweed". As a result, this paper provides a survey of the Holy Quran reciter identification problem, describing the proposed techniques, models and challenges in this area, focusing on the advances made during the last decade to help future researchers who aim to enhance previous results.

Keywords: *Reciter Identification, Quran, Arabic Language, Speaker Identification, Machine Learning.*

1. INTRODUCTION

Arabic is a widespread language that is already used in more than 22 countries worldwide by 300 million speakers [1]. The Holy Quran is the main reference for religious instructions for 1.8 billion Muslim around the world, and it's written in the Quranic Arabic (classical Arabic), the standardized literary form of Arabic. Also, the holy Quran consisits of 114 chapters or "Surahs", each chapter consists of a specific number of verses. It's an essential activity for Muslims to recite the holy Quran and listen to Quranic recitations. Since many reciters perform Quranic recitations, listeners may need to recognize the reciter's identity through intelligent systems to find more recitations for the same reciter. These systems accept a human voice as an input and process it through different processing phases in order to recognize the identity.

1) *The rules determining Quranic recitation's proper pronunciation and melodic intonation are known as "Tajweed". In other words, the set of accurate rules that govern how the Holy Quran should be read [2]. Voice is a physiological and behavioral biometric that conveys information which is related to a speaker's traits, such as age,* feelings, gender, and ethnicity. According to Arabic language phonetics, the standard Arabic language has 28 consonants, three short vowels and their three long counterparts [3]. It is worth mentioning that, Tajweed rules with their melodic rhythm require handling additional sounds. As a result, the identification systems that target identifying the holy Quran reciters are significantly affected by Tajweed rules which are followed by reciters while reading Quran. These rules introduce an additional challenge and more difficulties to speaker identification systems. The following are some examples of these Tajweed rules.

- Emphasis which is defined as a heaviness that enters the letter's body and the reciter's mouth is filled with its reverberation.
- Prolongation which is the prolongation of vowel sound.
- "Ghunnah" is considered a sound which resonates within the nasal cavity.

Also the Holy Quran that is recited by different reciters will probably differ from one person to another in terms of how the sentence will be delivered. This is because of the difference in their larynx sizes, vocal tract shapes, and also other parts of their voice production organs, in addition to the reciter's characteristic manner of speaking and expressing the emotional side in recitations.

Generally, Speaker identification is the task of identifying an unknown speaker's identity using speaker's speech signal [4]. The usefulness of developing intelligent systems that target automatic speaker identification problems is increasing with the information processing growing importance in a lot of fields such as crime detection [5],[6], and surveillance [7],[8].Since Arabic is such a complex language morphologically, many speaker identification systems are mainly designed for the English language, but still there is a lag in the research that focuses on developing speaker identification systems for Arabic language and the Holy Quran compared to English language.

During the last decade, there was an effort to target the Holy Quran to intensify the automatic understanding of Quranic verses and processing of the voice signals for Quranic speech. Several studies attempted to verify the correctness of Quranic recitation to help the Holy Quran's learners to recite the holy Quran correctly, detect recitation errors and correct pronunciation [9-11]. Also, some studies targeted recognizing the recitation type (style) which is the way by which the holy Quran is being recited, and there are seven types such as Warsh and Hafs [12]. Other studies focused on identifying the holy Quran reciters in order to help listeners recognizing the reciter's identity and selecting their preferences [13-17]. The process of identifying the Holy Quran accurate reciter based on many features in the corresponding acoustic wave can be defined as the Holy Quran reciter identification [18]. However, the progress in the research targets this problem is still slow. Therefore , this research aims to provide an overview of the research efforts that focused on this problem by conducting a survey reviewing the relevant studies and mobile applications as the number of papers conducted in this regard is relatively small, and it is a significantly challenging identification task that needs more research efforts, so we wanted to shed light on what has been done in this aspect and clarify the challenges that hinder progress in it, with a statement of the results obtained. We also focus on showing how the relatively small size datasets used by the proposed models of relevant studies affect our interpretation of obtained results.

It's worth mentioning that no significant review article targeted "specifically" the Quranic reciter identification problem before, which motivated our research to pay attention for this problem and shed the light on what has been done in this regard by discussing and summarizing the current state of the research based on previous efforts. we tend to present future researchers who aim to enhance the current results to models, techniques that have been used, drawbacks, results, and challenges. We are not concerned with proposing new technique or approach. Therefore, the main significant contributions of this research are listed as following:

- Describing the main structure for reciter identification problem.
- Covering all Quran reciter identification efforts from 2012 until 2023.
- Listing the used and available datasets and discussing problems related to them.
- Determining the pre-processing techniques to deal with the problem.
- Determining common features extraction and classification techniques capable of dealing with the holy Quran reciter identification problem and their limitations.
- Identifying the existing Quranic reciter identification mobile applications.
- Shed the light on the limitations of validation techniques and how they can affect the interpretation of the results of proposed models.
- Developing a new simple criterion to describe datasets and show the effect of datasets volume on interpreting accuracies achieved by the studies covered in this survey.
- Listing the main challenges that affect the progress in the research in this field .

The rest of this paper is organized as: Section 2 provides a background for the holy Quran reciter identification problem , Section 3 presents the process of collecting the relevant material and the research methodology, Section 4 includes data acquisition and pre-processing phase, Section 5 presents the features extraction techniques that were utilized for reciter identification task, Section 6 includes the commonly used classifiers in reciter identification task followed by a discussion on validation techniques, Section 7 analyzes relevant mobile applications, Section 8 introduces an overall discussion followed by the research challenges introduced in Section 9 , and Section 10 provides the conclusion.

2. BACKGROUND

Unlike other speaker identification tasks, The Holy Quran reciter identification is a very challenging task since Quranic recitations are performed with regard to the rules of "Tajweed' which results in a relatively more nonstationary signal than the ordinary speech [19]. Additionally, the emotional features that are added by the reciter and the transition from one tone or pitch to another while reciting the Holy Quran distinguish Quranic recitations from ordinary speech [20], and result in more important features should be taken in consideration while processing input signals . The structure of the holy Quran reciter identification task consists of three phases: data acquisition and pre-processing, feature extraction, and classification [21].

Acquiring data and pre-processing is considered the initial phase in any speaker identification system. It aims to obtain Quranic audio files and store them then apply pre-processing steps to prepare and organize data for further analysis. Accordingly, Quranic audio samples for different reciters are the input for the identification system. The main purpose of preprocessing steps is to make the identification system more efficient computationally [22]. Pre-processing steps aim to modify the speech signal to be more acceptable for further phases. These steps contain, Endpoint detection to specify the starting and ending points of the speech [23] and cleaning undesirable noises to decrease noise effect in order to get better accuracy [24]. It's worth mentioning that Quranic recitations may have different types of noise and several silence intervals that should be eliminated and filtered out before processing. Accordingly, preprocessing is considered a critical phase for reciter identification problems in order to simplify further phases and gain reliable results.

Extracting features of speaker's voice from speech signal is a pivotal stage. Human speech signals are considered a powerful medium of communication that involves rich information about the speaker, such as gender, accent, emotional traits, etc. These unique and important characteristics enable researchers to determine speakers by voiceprint recognition [4],[26]. Many feature extraction techniques have been used by researchers for speaker identification problems, such as Mel-Frequency Cepstral Coefficient technique [27], linear-Frequency Cepstral Coefficient technique [28], Linear Prediction Coding technique [29], and Linear Predictive Cepstral Coefficients technique [30], and Discrete Wavelet Transform (DWT) [31]. In terms of the challenging task of reciter identification, MFCC technique is commonly used by researchers for extracting information related to the reciter [16] [20], [32]. This is because it performs effectively and it has a high recognition accuracy and a low computational cost compared to other techniques. Also, MFCC is considered to be less Complex and more to imitate the auditory system of humans.

On the other hand, LPC technique is rarely used for extracting features for reciter identification [15], and that could be as a result of LPC drawbacks. LPC is considered a spectrum feature extraction technique that can introduce good interpretation for the frequency-domain and time-domain. Generally, LPC technique is a popular feature extraction technique in other speaker identification studies since it is easy and fast to apply, and it can also extract time-varying formant data and store it [33-35]. However, it is not the suitable technique to represent speech since it assumes that signal is stationary within a given frame, so it doesn't accurately analyze the localized events. In addition, the unvoiced and nasalized sounds can't be captured properly using LPC .Some other drawbacks could be addressed related to LPC. For instance, reducing bit rate introduces information loss. Accordingly, compared with MFCC technique, LPC isn't a common technique in extracting features for reciter identification task.

The extracted features are fed into classifiers to build trained models. Different types of classifiers can be employed to identify speakers such as Hidden Markov Model (HMM) [36], Gaussian Mixture Model (GMM) [37],[38], Support Vector Machines (SVM) [39], Vector Quantization (VQ) [40], K-Nearest Neighbor (KNN) [41],[42], Artificial Neural Network (ANN) [43] and Deep Neural Networks (DNNs) [4]. It's worth mentioning that, SVM classifier is commonly used for reciter identification since it performs effectively with small datsets [17],[18]. Also, VQ and GMM classifiers are used for reciter identification as their high recognition ability [16],[20]. Moreover, Recurrent Neural Networks (RNNs) which are a type of (DNNs) are considered more suitable for reciter identification task because of their ability to model time-variant data effectively [19].

The Quranic reciter identification task is completely different from Quranic recitation verification and correction task that is based on Automatic Speech Recognition techniques (ASR).

© 2023 Little Lion Scientific
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Quranic recitation verification is the process of verifying the correctness of recitations and reporting recitation errors to assist learners reciting the holy Quran correctly. Also, Quranic recitation recognition is the process of identifying the type of recitation ("Qira'ah") among the authorized recitation styles based on ASR techniques. ASR targets identifying spoken words without focusing on who is the speaker.on the other hand, Automatic Speaker Identification targets extracting and characterizing the information in the speech signal conveying the identity of speaker .it's worth mentioning that, there are significant pararrels between speaker identification and Quran reciter identification. However, identifiying the holy Quran reciter is a challenging task as:

- Each reciter has special phonetic characteristics varying over time. These characteristics affect word pronunciation and the way of delivering verses.
- Tajweed rules that lead to deal with additional sounds and considering more features.
- The highly emotional side in reciting the holy Quran is controlled by reciter's manner of expressing it.

The exisiting review papers only dicussed Quranic recitation veification [11]. No review paper discussed identifying the Quranic reciter. As a result, in this research we focus on the holy Quran reciter identification problem , reviewing efforts in this regard dusring the last decade from 2012 to 2023, and describe limitations and research challenges .

3. MATERIALS AND METHODS

In this section, we discuss the process of collecting the relevant material that achieved our goal of research and the research methodology.

3.1 Search Strategy For Publications

The relevant publications were chosen after extensive and accurate searches through Scholarly databases and web search engines. Scopus , Web of Science and Google Scholar were the main sources of journal articles and conference proceedings. The search strings were built based on four major keywords, "Quran," "Reciter," "Identification," and "Recognition." However, some synonyms were used in order to gain suitable search results and ensure a complete coverage. The search strings were built using keywords and the logical operators (AND, OR). We conducted many search attempts with different keywords combinations aiming to find the most suitable search strings that lead to suitable results. We retrieved the publications that matched the search strings and were published from 2012 until 2023.The total number of retrieved results was 62.

3.2 Publications Selection

To ensure the relevance of retrieved results to our research, we put some criteria for filtering those results and classify them into two groups: relevant studies and irrelevant studies. Our main focus in this research was on publications that targeted the Holy Quran reciter identification problem. Thus, papers that discuss recitation verification or recitation-type recognition were excluded from our study. The full set of our inclusion and exclusion criteria are listed in Table 1. Publications were screened based on title, abstract and the main content to ensure that they met our criteria.

We ended up with a number 17 paper achieved our goal of the research after filtering out a number of 62 publications that matched search strings. Those 17 publications are considered as the total number of research efforts targeted reciter identification task during the period from 2012 to 2023 as shown in Figure 1. The statistical distribution of the selected material based on the type of publications is shown in Figure 2.

Table 1: Inclusion And Exclusion Criteria For filtering Publications.

Criteria	Description
Inclusion	Studies focus on identifying Quran reciter using speaker identification techniques.
	Studies that are written in English.
Exclusion	Studies focus on Quranic recitation verification.
	Studies focus on Quranic recitation type identification.
	Studies focus on the Holy Quran in general.
	Studies focus on speaker recognition in general.
	Studies that are not written in English.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

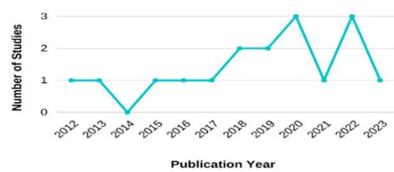


Figure 1: Statistical Distribution Of The Selected Material Based On The Publication Year

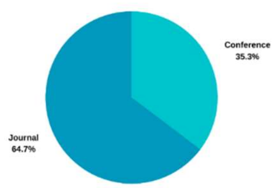


Figure 2: Statistical Distribution Of The Selected Material Based On The Publication Type.

3.3 Data Extraction:

In order to extract important data from the relevant papers ,Analyzing relevant publications is conducted in terms of these points :

- The way by which datasets were collected for training classifiers.
- The pre-processing steps that were applied on datasets.
- Feature extraction techniques that were used for extracting the useful features for model training .
- Classification algorithms that were used and their performance.
- Research challenges were faced by researchers.

3.4 Search Strategy For Mobile Applications

The relevant mobile applications were chosen after conducting an extensive search through Google play and App store . For constructing search strings, we used specific relevant keywords. These keywords were "Quran" , "Qari"," Reader", "Reciter" "Identification" and "recitation" . we also used the Arabic keywords that have the same meaning as English ones. In addition, we also used some synonyms to gain the relevant results. We conducted several search attempts using different combinations of these of keywords. We retrieved 11 search results related for Quran mobile applications.

3.5 Mobile Applications Selection

After applying search strings in Google play and App storeand retrieved results, we applied some criteria to filter results and keep relevant mobile applications. We focused on mobile applications that enable identifying the quranic reciter (Qari), whether identifying reciter is the only function of the application or just a sub function. We excluded Quran mobile applications that taget only recitation correctness or verification .we also excluded quran reading mobile application .Table 2 shows inclusion and exclusion criteria for filtering and selecting relevant mobile application. After checking the retrieved mobile applications specifications and applying inclusion and exclusion criteria, we ended up with two mobile applications "Rateel" [44] and "Nadyy" [45].

Table 2: Inclusion And Exclusion Criteria For filtering Mobile Applications

Criteria	Description
Inclusion	Quran mobile application focusing on identifying Quran reciter (Qari) or offering a Quran reciter identification option.
	Free mobile applications.
Exclusion	Quran reading mobile applications.
	Quran mobile applications that don't offer identifying Qari option.
	Applications that use a human-based identification approach.
	Paid mobile applications.

4. DATA ACQUISITION AND PRE-PROCESSING

The first phase of the reciter identification system starts with constructing a dataset and preparing it for further analysis by pre-processing steps. Thus, this section discusses the datasets and pre-processing steps that were utilized through the previous proposed reciter identification studies.

4.1 Datasets

Data is the most essential component for the reciter identification system. Using a relevant, large and high quality dataset helps develop successful automatic identification systems. Also publicly available datasets are important to intensify the research that targets Arabic speaker recognition problems in order to develop reliable classification systems and allow researchers to compare their proposed systems on the same benchmark.

Unfortunately, the research area of identifying the Holy Quran reciters lacks common and public datasets that can be used as large, well-designed balanced, and less biased resources of recitation samples and their corresponding reciters. However, in this regard we can address three different Quranic public available datasets," AR-DAD" [46], "EveryAyah" [47], and "Tarteel" [48]:

- The AR-DAD dataset is a publicly available dataset by various reciters from several countries [47],[46]. AR-DAD consists of almost 15810 Arabic audio clips that were collected from 37 chapters from the holy Quran for 30 popular reciters. Moreover, it contains 397 audio clips for 12 competent imitators of those reciters.
- Tarteel is a labelled and crowdsourced Quranic dataset contains 25000 Quranic audio files from different people worldwide forming 67.39 hours of audio. Also, the dataset covers a various recitation styles, speeds and proficiencies. It's worth mentioning that, recitations were collected from over 1200 individuals of different genders, ethnicities and ages, and this variety is considered as one of the main advantages of the dataset as it helps reducing bias .
- EveryAyah is a public dataset contains Quranic audio files of the whole Quran for 52 professional reciters. Only 26 reciters have recorded every verse of the holy Quran. Also , the dataset contains Arabic and translated Quranic recitaions. Two recitation styles were included in the dataset (Warsh and Hafs). It is worth mentioning that, the expert audio files in AR-DAD dataset were extracted from the dataset of EveryAyah.

It is noticeable that, most of the proposed studies in the reciter identification area depended on locally designed small datasets. These datasets vary in the number of recitation samples which are the Quranic verses recited by each reciter, and in the number of reciters. However, the study proposed by [49] used 1000 Quranic audio files from the AR-DAD dataset. Using small size datasets can be considered as a critical point in evaluating the success of models because the reliability of identification drops when using large datasets. Also there was a significant bais in most of datasets used in Quranic reciter identification studies towards male reciters. Table 2 summarizes most of the datasets that were used in the previous studies.

4.2 Pre-processing Steps

After constructing datasets, it is important to organize data before the feature extraction process. Thus, some pre-processing steps can be applied on the collected datasets to put data in a simplified, organized, and useful format [50]. These steps involve:

- End Point Detection: this step aims to specify the starting and ending points of recorded words. It is a required step to isolate the speech of interest from the background. Spectral energy or short-time energy is commonly used as the primary feature with other features in endpoint detection algorithms. It's worth mentioning that the endpoint detection features become not so reliable in non-stationary noisy environments (That has a time-varying statistical profile) [2]. The proposed reciter identification study by [20] utilized the method introduced by [51] for silence removal and audio segmentation. This utilized method is based on two simple audio features which are signal energy and spectral centroid. When the feature sequences are extracted, the thresholding approach is applied on those sequences to detect segments of the speech.

- Noise filtering: this step aims to decrease the effects of noise in the speech in order to get better identification results. Authors in the proposed reciter identification study of [20] addressed utilizing the algorithm of Hourri and Kharroubi to eliminate the defect areas and produce suitable voice segments, this algorithm depended on transforming the MFCC features into Deep speaker features (DeepSFs) using DNN in order to increase MFCC noise robust nature [52]. The filtration process was employed by changing the frames' length and also the step's length.

5. FEATURE EXTRACTION TECHNIQUES

The Feature extraction is the second essential phase for the reciter identification system in which the important features of a reciter's voice can be extracted to be used for further processing. The audio signals are the input files to the identification system, so it's essential to extract important audio features into a format that's described as machine-understandable using feature extraction techniques

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

to be used for training classifiers to identify reciters [16].

As we mentioned in section 2, several feature extraction techniques were used in speaker identification systems in general, such as MFCC, LFCC, LDC, DWT and LDCC, etc. The following subsections discuss feature extraction techniques that were used for reciter identification .

Table 3:The Holy Quran Reciters Identification Datasets Description In Previous Studies.

Paper	Date	Number of Reciters	Gender	Description
[13]	2012	20	Unknown	Contains 20 sample recordings from 20 reciters for one verse for training and 200 recordings for ten different verses for every reciter for testing.
[15]	2015	4	Males	Conrains 44 Quranic audio samples.
[16]	2017	5	Males	Contains 15 Quranic audio samples.
[18]	2019	15	Males	1650 audio files from 110 verses from chapter 18
[20]	2022	14	Males	Each reciter recited 4 verses of Surah Al-Kawthar.
[53]	2019	12	Unknown	Contains 120 Quranic audio samples.
[54]	2020	10	Males	1100 wav files for chapter 18 and 830 wav files for chapter 36.
[55]	2020	30	Unknown	15810 audio files for reciters. 397 audio files for top imitators.
[56]	2020	7	Males	Contains 2134 Quranic audio samples.
[49]	2021	10	Unkown	Contains 1000 audio files from AR-DAD
[32]	2022	10	Males	2060 wav files for chapter 7 and 300 wav files for chapter 32.
[57]	2022	60	Unknown	10 hours of recitations forming 960 wav files
[58]	2023	7	Males	Each reciter recited 80 minutes of Quran chapters on an audio file

5.1 Mel-Frequency Cepstral Coefficients (MFCCs)

MFCCs are low-level information that is principally based on spectral information which is derived from a short time-window segment of speech (about 20 milliseconds) [28]. The parameters of the spectrum that have a dependency on the speaker's physical characteristics are the coefficients.

The MFCC technique includes various steps as shown in Figure 1, starting from dividing the input speech signal into frames, followed by applying the windowing to decrease the discontinuous effect on the signal after dividing it into frames [59]. Then the given signal is represented in frequency domain using the Fast Fourier Transform (FFT) [60]. In the next step, the frequency domain signal is converted to Mel frequency scale. Then employing Discrete Cosine Transform (DCT) to convert the log Mel scale spectrum to time domain. Mel Frequency Cepstrum Coefficient is the result of this conversion [61].

5.1.1 MFCCs in relevant studies

It is noticeable that the MFCC technique was the most common one used by most reciter identification studies as it shown in Table 3. That could be explained as MFCC has less complexity and can imitate the auditory system of humans. According to [16], MFCC is an easy feature extraction technique to moderate and can handle multiple languages and multiple speakers. According to the comparative study introduced by [62], the MFCC technique was observed to be more efficient to work with Arabic language since it achieved a higher accuracy compared with other methods that were used in the mentioned study. Also according to the study by [54], applying the MFCC technique for extracting features is due of the variation in verses length in each chapter of the holy Quran. Therefore, most of the reciter identification studies used MFCC for extracting features.

MFCC technique is mainly affected by two components, the first one is the number of filters used, and the second one is window type.The proposed study by [63] introduced comparison experiments to reach the best implementation of MFCC as a feature extraction technique for Quranic speaker recognition. To achieve that, the authors choose a Quranic verse (بسم الله الرحمن الرحيم) that was spoken by five different reciters for determining the suitable number of filters and the perfect window type. The number of filters was determined after varying it till reach the best recognition performance. Too many filters or too few ones didn’t lead to high accuracy for recognition. Considering the chosen number of filters, Rectangular window and Hamming window were used and the results showed

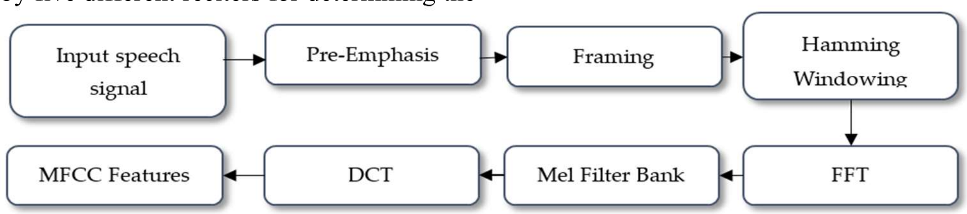


Figure 3: MFCC Technique Structure

that the maximum efficiency was achieved when Hamming window was used.

The proposed reciter identification system by [18] used MFCC for extracting features for the 1650 acoustic waves. With the use of a Hamming window, the speech signal was divided into frames of 15 ms. The authors explained choosing a 15 ms frame size as they experimentally determined that better identification performance was generated by 15 ms frames because the Holy Quran reciters have a slow recitation rate in general. The authors mentioned that their identification system’s ability to learn signal characteristics was not enhanced by using a large frame.

The study by [20] utilized the MFCC technique for extracting the sound features of every reciter for the Holy Quranic verses. The log-energy coefficients (the zeroth coefficients) were excluded as they carry a small amount of infor-mation about the speaker.

The proposed study by [53] used two different approaches to represent audio Quranic files: The first approach analyzed audio in the frequency domain and used (MFCCs and pitch) as features for classifiers, and the second one used spectrum to analyze the audio file in the image domain and used Auto-Correlogram approach [64] for feature extraction. The results showed that the classifiers that were trained using MFCCs and pitch can learn effectively the holy Quran reciter than using Auto-Correlograms. Also the proposed reciter identification systems by [13], [19] , [32] utilized the MFCC technique for extracting features for Quranic reciters to feed them as inputs to classifiers.

5.2 Linear Predictive Coding (LPC)

Linear predictive coding (LPC) is a feature extraction technique for analog signal encoding. In LPC technique, a particular value is predicted using a linear function of the past values of the signal. The linear predictive filter allows the next sample’s value to be determined by a linear combination of previous samples. LPC is considered a fast technique which is suitable to work with systems in which the audio is transmitted over a large range [65].Some disadvantages could be considered for LPC [66] as follows:

- A noticeable quality loss may happen as bit rate may be reduced by LDC significantly. However, the speech can still be audible and understood.
- The characteristics of the vocal tract can’t be represented by LPC from the glottal dynamics.
- LPC can’t capture the nasalized and unvoiced sounds correctly.
- Creating each speaker’s model using LPC takes more time and increases the compu-tational cost.

5.3 Discrete Wavelet Transform (DWT)

The Wavelet Transform (WT) is considered a suitable tool for analyzing nonstationary signals like speech. WT decomposes the signals in a set of basic functions known as wavelet. The Discrete wavelet transform (DWT) is a specific type of Wavelet Transform (WT) and it represents signals in frequency and time in a compact manner for efficient computation [67]. The sound signal can be processed efficiently with DWT because of its efficient time-frequency localization, multi resolution and multi scale analyzing features. The reciter identification study proposed by [15] utilized a combination of discrete wavelet transform (DWT) and LPC features to enhance identification accuracy rather than using every one of them singly Since the proposed system achieved the best accuracy when a random forest (RF) classifier was trained with the combination of LPC and DWT features.

5.4 Deep Learning-Based Feature Extraction Models

Deep learning-based feature extraction models have recently become highly preferred in recognition problems as they can achieve high-level contextual representation. Also, they can learn effectively the basic units for less labeled data and capture various aspects without a clear segmentation for words. For instance, Wav2vec2.0 [10] and HuBERT [21], which were utilized as audio representation learning tools in the proposed study by [49] to identify Quranic reciter.

5.4.1 Wav2vec2.0 model

Wav2vec2.0 model uses a multilayer convolutional neural network (CNN) in processing the raw audio data to obtain latent audio representations of 25ms each. Wav2vec2.0 performs joint learning between quantization of the latent representation which is considered as the main strength point [49]. For feature extraction, the representations are encapsulated into a transformer and a quantizer.

5.4.2 HuBERT model

The Hidden-Unit BERT (HuBERT) is a self-supervised speech representation learning approach that can utilize a step of offline clustering for providing aligned target labels for a BERT-like prediction loss [68]. HuBERT relies on the unsupervised clustering step's consistency for producing a better representation.

In Table 4, we summarize features extraction techniques used in reciter identification studies in terms of techniques most important advantages, critical drawbacks and suitability to work with the problem.

6. CLASSIFICATION TECHNIQUES

After successfully extracting the reciter's voice features, the third phase of the reciter identification system is the classification phase. The Reciter identification task is considered a Multi-class classification problem in which the classifier classifies n-vectors into *M* groups. Several classifiers can be used and perform effectively for this task. In this section, we shed light on classification algorithms that were used for the holy Quran reciters identification problem based on previous studies, followed by a discussion on validation techniques limitations and the relationship with the reliability of the research presented in this regard.

6.1 Random Forest (RF) Classifier

Random Forest (RF) is a powerful statistical classifier in which many classification trees are fit to data then the predictions that are resulted from these trees are combined. RF combines two resampling methods: features random selection and bagging. Compared to other statistical classifiers, RF classifiers have a very high accuracy for classification [15]. The most important advantages of RF are as follows [13]:

- Less sensitivity for outlier data.
- An importance score for every feature that contributes to the classifier can be generated automatically using RF.
- Handling highly non-linear data.
- It overcomes the overfitting problem.

The proposed study by [15] introduced a speaker identification system for Quran reciters using a random forest (RF) classifier and it achieved high performance with a combination of LPC and DWT features rather than using it with LPC features or DWT features singly. The experiment was applied on a small dataset of 44 samples for only four reciters which can be considered as a limitation for their results. Also, the RF classifier was chosen and compared with other classifiers (Naïve Bayes and J48) to recognize the holy Quran reciter by [53] and RF showed good results when it was trained using MFCC and Pitch features rather than using Auto-Correlograms. The same study showed that RF and Naïve Bayes could learn to identify reciter with a good recognition model and coinciding results that outperforms J48. The low performance of J48 compared with RF was explained by the authors as it could be because of the sensitivity of J48 to noise in data. It's worth mentioning that, there was no mention of any noise filtering steps were performed on the dataset before training classifiers in order to avoid this.

6.2 Artificial Neural Network (ANN) Classifier

AN Artificial neural network is a biologically-inspired information processing system that imitates the human brain. It consists of multiple layers of simple processing elements known as neurons that are highly interconnected [69]. ANN can be used successfully in several areas such as image analysis and biochemical analysis [70].Although ANNs have been used in speech recognition systems successfully, we can address some limitations of using ANNs in voice recognition as follows [32]:

- ANN is not a solution to the problems of voice recognition as there is no structured methodology and it has an empirical nature.
- The problem-solving methodology of ANN can't be described accurately.
- The quality of the output achieved by ANN sometimes is unpredictable.

Table 4: Features Extraction Techniques Utilized In Reciter Identification Problem

Features Extraction Technique	Studies	Advantages	Critical Drawbacks	Suitability
MFCC	[13],[16],[18], [19],[20],[32], [53],[54],[57], [58],[63]	High Accuracy. Low computational coast. Less complexity. Humans auditory system imitation.	Sensitivity for noise.	Very suitable to work with reciter identification problem and achieve excellent results.
LPC	[16]	Efficient and powerful technique for representing speech signal in a mathematical form.	Quality loss. Noisy environment poor performance.	LPC features Can be used for reciter identification combined with DWT features for enhancing results.
DWT	[15]	Powerful technique for signal processing . Reducing noise in input speech signal .	Shift sensitivity lack of phase information which describes the behavior of non-stationary signal accurately. poor directionality.	Better results can be achieved for reciter identification when using DWT and LPC features together than using DWT singely
Deep learning-based (Wav2vec2.0, HuBERT)	[49]	High-level contextual representation. High accuracy . low validation loss. Learning basic units for less labeled data. Capturing aspects without words clear segmentation.	Requiring an expensive computational cost. Slowing down the inference speed.	Suitable and Works effectively .

The proposed reciter identification system by [54] showed an efficient performance for the ANN classifier when it was trained using MFCC features to identify reciters. The authors explained the misclassification as it was due to the way of reciting different verses, the emotion that the reciter used with every verse and the tone variations. All of these reasons are due to the restricted rules of "Tajweed" that should be followed.

6.3 Support Vector Machine (SVM) Classifier

SVM is a supervised machine learning algorithm that aims to find the hyperplane that classifies or separates the data points. The samples that lie near the hyperplane are called support vectors. The margin is defined as the maximum distance between those support vectors and the hyperplane [71]. SVM classifier can perform effectively with relatively small-size datasets and unstructured data. Moreover, SVM has a scalability to high dimensional spaces which are considered one of its important advantages without having a risk of falling into overfitting.

On the other hand, for large datasets, the training process for SVM is usually time-consuming for large datasets .

SVM performed with promising results in previous reciter identification studies. It showed a superiority over ANN when both of them were trained and tested using MFCCS features to identify holy Quran reciters by [19]. Moreover, SVM was used with different combinations of perceptual features and achieved promising results in identifying Quranic reciter by [17]. The study introduced by [56], showed that the performance of SVM degraded by replacing the reciters with imitators. According to authors, the degradation was expected to be as a result of the noisy dataset of imitators and also because some imitators imitated the professional reciters poorly.

6.4 Vector Quantization (VQ) Classifier

VQ is an efficient data compression method that squeezes data points into smaller datasets. These data points are remapped to a finite even number of clusters represented by their centroid in order to achieve data compression [72]. The VQ technique is a widely used method in speaker identification problems [73]. However, the complexity and increasing codebook memory can be addressed as disadvantages of this technique [74].

An improved Linde–Buzo–Gray (LBG) of VQ algorithm was proposed by [20] to identify Quranic reciter. The improved (LBG-VQ) technique was based on iterating till codebooks ‘centroids reach the best optimal values. Also, the technique proposed by [20] outperformed the models presented by [19], [75] ,[76] utilizing 16 centroids.

6.5 Gaussian Mixture Model (GMM) Classifier

Gaussian Mixture Model (GMM) is described as a parametric Probability Density Function (PDF) that is represented by the combination of Gaussian component densities. GMM technique has a high recognition ability, so that it is commonly used in speaker identification problems. Moreover, compared with other modeling approaches, GMM can be trained so fast. Also with relative ease, GMM can be scaled and updated to add new speakers [25]. So generally, GMM technique can provide a simple effective speaker representation and it is computationally inexpensive.

GMM technique was utilized to learn reciter with MFCC features in the proposed study by [16].It is worth mention that, under the small volume of dataset used by the model introduced by [16], the high accuracy achieved by the proposed model could be misleading.

6.6 Deep Neural Networks (DNN) Classifiers

A Deep Neural Network (DNN) is an artificial neural network (ANN) with many hidden layers between the input and output layers [77]. The role of these additional layers is to extract features from lower layers which enables the efficient modeling of complex data. Deep Neural Networks DNNs have been commonly used in speaker identification in recent years, since they have proven to be more effective than traditional techniques [78]. Also, some studies investigated DNN in a noisy environment, and it performed well [79]. Several types of DNNs were used in reciter identification proposed studies, such as MLP, CNN, and RNN, that were utilized by [19],[49] and showed successful performance.

6.7 Validation

The goal of the classification model is typically not just to achieve a good classifica-tion on the given dataset, but rather to achieve a good classification on new data that we have not yet seen. There is no perfect way to get such a model without some considerations and assumptions about future data. One very common assumption is that the data is described by a formal probability model. With that assumption, techniques from statistics probability can be used to predict how well a model will work on new (unseen) data. The basic assumption here is that the future data will look like the test data. This approach has been very successful in many applications. If a classification model predicts the outcomes for new unseen data values as well, or nearly as well, as it predicts the outcomes on the data used to form the model, it is said to have a good-generalization ability. In the opposite case, when the model makes predictions on new unseen data that are much worse than the predictions on the data used to form the model, the model is said to have poor generalization ability. The simple but effective process that can help to assess the general-ization ability of the classification model is known as the validation, and it is the key point of the classification model.

In the out-of-sample validation technique, the data is divided into two sets, a training set and a test set (validation set). This is often done randomly, with 80% of the data put into the training set and 20% put into the test set. A common way to describe this is to say that '20% of the data were reserved for validation'. Another common choice for the split ratio between the training and the test sets is 70%-30%. The classifier uses only the data in the training set. The classifier model that we come up with is based only on the data in the training set; the data in the test set is never been seen by the model. Then we judge the model by its error fit on the test set. Since the model was trained without any knowledge of data in the test set, the test set is considered as the data that is new and unseen, and the performance of our model on this data gives us at least an idea of how our model will perform on new, unseen data. If the prediction error on the test set is much larger than the prediction error on the training set, we conclude that our model has poor generalization ability.

Another sophisticated validation technique, is known as cross-validation, can be considered. Cross-validation divides the data set into ten sets, called folds (5-fold cross-validation is also commonly used).
Then fit the model using folds 1, 2. . . 9 as training data, and fold 10 as test data. (So far, this is the same as out-of-sample validation.). Then the model uses folds 1, 2. . . 8, 10 as training data and fold 9 as the test data. We continue, using a model for each choice of one of the folds as the test set. We end up with ten different models, and ten assessments of these models.

This introduction is important to understand the limitations of out-of-sample and cross-validation techniques and their impact on the accuracy of the proposed Quran reciter identification models. The first limitation is that the assumption that the test data and future data are similar can fail in some cases and applications. The second limitation arises when the dataset is small, which makes it harder to interpret out-of-sample or cross-validation results. In this case, the test results of the cross-validation can vary considerably, due to the luck of which data points fall into the different folds. We can expect to see more variation in test prediction error than with larger datasets. Therefore, one of the main challenges affecting the robustness of the model is the size of the data. The accuracy of these models is highly dependent on the sizes of their datasets. We noticed that most of the proposed models for Quranic reciter identification adopted a relatively small data size, which can't be considered sufficient to converge with the real pattern, as this may lead to a misleading pattern.

According to studies covered in this survey, the models proposed by [13],[15] ,[16],[18],[49],[53],[57] used relatively small datasets as compared to [32],[31],[55]. we will now define a new simple criterion that could help get reasonable interpretations for accuracies achieved by the studies covered in this survey . In this part, we excluded studies with unkown audio files number from our calculations.
Let:

$$R = \frac{\text{The volume of data used in a specific paper}}{\text{The highest volume of data used among the whole survey}} \quad (1)$$

where the R values range from 0 to 1. Let us further define the Relative-volume-Degree (RVD), as another representation of R scaled from relatively very high, to relatively very low as follows:

$$RVD \begin{cases} \text{Relatively very high} & if \quad \text{RVD} \geq 0.75 \\ \text{Relatively high} & if \quad 0.75 > \text{RVD} \geq 0.5 \\ \text{Relatively low} & if \quad 0.5 > \text{RVD} \geq 0.25 \\ \text{Relatively very low} & if \quad RVD \leq 0.25 \end{cases} \quad (2)$$

The following Table 5 shows RVD in terms of R, the corresponding number of samples and the number of reciters considered for different Quran reciter identification models. As shown in Table 5, the RVD conveys a relatively very high volume of data for only the study introduced by [55] and a relatively very low volume of data for rest of the studies .This means that the reliability of most of the proposed models may require consideration.

7. MOBILE APPLICATIONS

Quran mobile applications have a increasing popularity between muslims worldwide. These applications offer several options for users and can be classified into three categories: Quran reading applications that intoduce a digital copy of the holy Quran for users such as "Khatmah" [80], Quran learning applications by which user can find an audio version of the holy Quran and this type of applications may be supported by Quran memorization option by which users can memorize Quran by playing audio parts of Quran for expert reciters and allowing user to repeat them . Quran learning applications can also present memorization testing service by which user can users can test the correctness of their memorization, and this testing service can be human or AI -based . another service can be included in Quran learning applications is about verifying the correctness of recitation be enabling users to recite specific parts of Quran then the system can analyze input speech and detect any recitation errors such as "Tarteel" [81] .

The last category of Quran applications and the one we concerned with is Quran reciter (Qari) identification applications by which the user can record a specific part of quranic recitation for unkown reciter then the system analyzes input signal and provides reciter details. It's worth mentioning that, number of applications that target identifying the reciter is still limited and in this regard we analyze two reciter identification mobile applications which are "Naddy" and "Rateel". Tables 6,7 show the results of analyzing these two applications in terms of specification, method of functioning and limitations.

8. DISCUSSION

After analyzing the proposed Quranic reciter identification studies in terms of the pre-defined points presented in section 3, it's worth mentioning that, the Holy Quran reciter identification problem is considered a challenging task that needs more future efforts to work on limitations and challenges.

Despite the difficulty of this task, significant research efforts targeted developing systems for identifying Quranic reciters. We noticed that Most of reciter identification proposed studies depended on small and local datasets. Also these datasets were built with a significant bias towards male reciters which motivates future researchers

to explore creating more qualified, large and less biased datasets to build more professional and reliable identification models and explore including more reciters from different regions. Also, Pre-processing steps need to gain more attention from researchers in terms of reducing noise effect which greatly influences classification results.

Features extraction is an important step in any identification system. It was noticed that the MFCC technique was an effective feature extraction technique to work with the problem and it was employed in most proposed studies. It also achieves high accuracy and low computational cost but it is still sensitive to noise. LPC also has poor performance in noisy environments. As a result, alternative techniques need to be estimated and utilized to deal with this challenge. Several enhancement techniques can be estimated to enhance the performance in the presence of noise such as spectral subtraction (SS) which can reduce noise effect and it has good performance for enhancing ASR systems [86]. Also techniques like relative spectra [84] and perceptual linear prediction [85] can be estimated with reciter identification.

In terms of classification phase, it was noticed that using RNNs could be more suitable to deal with reciter identification task due to their ability to model time-variant data effectively [19].We can't ignore the importance of depending on large and high-quality datasets to get the advantage of DNNs .other classifiers such as RF and GMM showed good performance alongside with their low computational cost.

Generally, it's difficult to compare the performance of models that were used in the previous studies since these models were trained using different datasets and also used different criteria but we can summarize these studies as shown in Table 8. From Table 8 we notice that, the use of MFCC with GMM [16], SVM [18],[57], BLSTM [19], KNN [54], ANN [54], and LBG-VQ [20] classifiers lead to the highest accuracies among all the studies presented in this survey. However, relatively few studies with other classifiers such as Naïve Bayes, J48, and RF [53], used MFCC, and they achieved relatively low accuracies as the studies presented by [13]. As for using the J48 classifier with MFCC, the study presented by [53] indicated that the J84 is sensitive for noise, and this explains the relatively low accuracy resulting from such a model.

The Quranic reciter identification model presented by [56] showed that the perceptual features with SVM, SVM-RBF, SVM-Linear, Decision Tree, Logistic Regression, and Random Forest classifiers delivered comparable accuracies among them (≈90%) and the XGBoosting classifier showed the best results. Regarding the perceptual features with SVM, the same result was achieved by the study presented by [17] using the same dataset. In addition, the study by [55] used a combination of perceptual and acoustic features with several classifiers and achieved high results with a large dataset. According to the study by [49], the two deep learning-based feature extraction techniques, HuBERT and wav2vec2.0 with CNN, RNN, and MLP deep neural networks classifiers, introduced high performance, and in this respect wav2vec2.0 outperformed HuBERT. The RF classifier gave higher accuracy when it was used with LPC+DWT features but the accuracy was significantly reduced when LPC and DWT features were used separately.

Besides the feature extraction techniques and the classification methods, the volume of data and the noise, influence the performance of models. Thus, the mentioned accuracies in Table 8 could be misleading if we don't take into consideration the corresponding volume of data and the presence or absence of noise.

We believe that our research is a significant contribution to the topic of the Holy Quran reciter identification. However, the lack of research studies on this topic is considered our major research limitation, also most of the models used in the

© 2023 Little Lion Scientific
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

relevant studies were designed empirically, which doesn't enable us to generalize our findings. We hope that our research will encourage more research efforts in this regard.

Table 5: RVD Corresponding To Quranic Reciter Identification Models.

Paper	Number of Samples	Number of Reciters	RVD
[13]	200	20	Relatively low
[15]	44	4	Relatively low
[16]	15	5	Relatively low
[18]	1650	15	Relatively Low
[32]	2060/300	10	Relatively low
[49]	1000	10	Relatively low
[53]	120	12	Relatively low
[54]	1100/830	10	Relatively low
[56]	2134	7	Relatively low
[57]	960	60	Relatively low
[55]	15810	7	Relatively high

Table 6: Reciter Identification Mobile applications

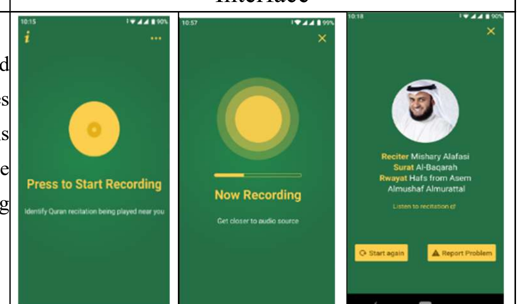
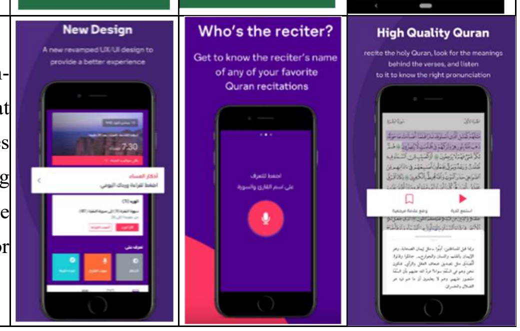
Mobile Application	Platform	Specifications	Interface
Nadyy	Android	A speaker recognition-based Mobile application that provides the details of Quranic recitations including reciter name, surah name , and rwaya (narration) by allowing user to record quranic recitations.	
Rateel	Android - IOS	A speech and speaker recognition-based mobile application that provides different services encludes the option of identifying the reciter of the holy Quran and the surah's name from any audio or video input.	

Table 7: Reciter Identification Mobile Applications: Method of Functioning And Limitations

Mobile Application	Method of Functioning	Limitations
Nadyy	It allows user to record a specific part of Quranic recitation. It uses signal processing for analyzing input signal by extracting information and matching against previously extracted information stored in system's huge database. System's huge database consists of recitations from diverse sources such as "Islamwe", "Assabile".	Supporting only old versions of android.
Rateel	It uses signal processing for analyzng recorded audio files from users .	The application shows a critical sensitivity for background noise and refuses most of recorded audios . These limitations make it not suitable enough for easy usage.

Table 8:Table 8: A Summary Of Covered Rciter Identification Publications.

Paper	Dataset Type	Feature Extraction Techniques	Classifiers	Accuracy
[13]	local	MFCC	LBG-VQ	Recognition rate=86.5% for clean samples. For noisy samples with SNR ranged from 25 dB to 0 dB.
[15]	local	DWT LPC DWT+LPC	RF	accuracy = 77.27% accuracy = 88.63% accuracy = 100%
[16]	local	MFCC	GMM	accuracy = 100%
[17]	local	perceptual features	SVM	accuracy = 90%
[18]	local	MFCC	SVM ANN	accuracy = 96.59% accuracy = 86.1%
[19]	local	MFCC	BLSTM	accuracy = 99.89%
[20]	local	MFCC	LBG-VQ	accuracy = 98.21%
[32]	local	MFCC	KNN ANN SVM	average recognition rate = 97.62% , 96.7% average recognition rate = 97.03 % 96.08% NA
[49]	AR-DAD	wav2vec2.0 HuBERT	MLP RNN CNN	accuracy = 96.23% for wav2vec2.0
[53]	local	MFCCs and Pitch Auto-correlograms	Naïve Bayes J48 RF	MFCCs and Pitch: accuracy = 88%, accuracy = 78%, accuracy = 88% Auto-correlograms: accuracy = 81%, accuracy = 78%, accuracy = 78%
[54]	local	MFCC	ANN KNN	average recognition rate = 97.6% , 96.7% average recognition rate = 97.03 % ,96.08%

[56]	local	perceptual features, short time energy	XGBoost SVM SVM-RBF DT LR RF	accuracy = 93.44% accuracy = around 90% accuracy = around 90% accuracy = around 90% accuracy = around 90% accuracy= 70.26%
[55]	Local	perceptual and acoustic features	DT RF LR SVM-LINEAR SVM-RBF XG GMM-UBM BLSTM CNN	accuracy = 73.33% accuracy =93.67% accuracy = 98% accuracy = 99% accuracy = 99% accuracy= 97.67% accuracy = 97.33% accuracy = 91% accuracy = around 94.33%
[57]	Local	MFCC	CNN SVM	accuracy = 99.66% accuracy = 99%

9 RESEARCH CHALLENGES

In this section, we focus on the most critical challenges faced by researchers in the Holy Quran reciter identification.

9.1 Datasets

The research in Quranic reciter identification lacks publicly available, large, balanced, and high-quality datasets that can help researcher get reliable results and allow them to compare their proposed systems on the same benchmark. It's important to ensure that the dataset is balanced enough and doesn't have classes with fewer recitations than others and reduce bias towards specific reciters or specific region. This bias also can be towards male reciters with a significant ignorance to female reciters.

9.2 Noise

For Quranic reciter identification problem, noise is considered a great challenge. The impact of background noises on the classification results is a critical problem to address. Also, echo sound can be generated during Quranic recitation [82]. Noise is considered the most significant factor that can affect identification accuracy. Although different noise elimination algorithms can be applied, these algorithms still can't eliminate noise effectively. Also, anti-noise algorithms still have limitations and specificity [83]. As a result, extracting robust features and building anti-noise models that can be adapted to several noise environments is still an open research challenge in the field.

9.3 Number of Reciters

The number of reciters impacts the efficiency of the identification system. As the number of reciters increases in our dataset, we may have to compromise with the correct identification of the reciter.

10. CONCLUSION

The Holy Quran reciter identification is considered one of the most challenging tasks in Arabic speaker recognition as it depends on "Tajweed" rules, emotional reciter's features and tone transitions. In this research, we conducted a survey on the holy Quran reciter identification task, focusing on the proposed models, techniques, drawbacks and challenges to help future researchers to understand the current state of the research in this area and motivate them to intensify efforts to enhance the research in this regard aiming to obtain more efficient results and work on the addressed limitations and challenges. we analyzed the existing relevant mobile applications and the results showed the necessity of developing the practical aspect of this topic to handle drawbacks and limitations. we also shed light on the impact of the limitations of evaluation techniques on interpreting the results of proposed models. Our research was conducted using a small sample size of relevant studies due to the lack of research on this topic which motivate our future work to be expanded to include more studies in this regard and also include efforts targeting unknown reciters and other types of Islamic emotional speech, such as spiritual songs (Ibtihal).

Acknowledgments: This study is supported via funding from Prince Sattam bin Abdulaziz University project number (PSAU/2023/R/1444).

REFERENCES:

[1] I. Al-Huri, "Arabic Language: Historic and Sociolinguistic Characteristics", *English Lit. Lang. Rev.*,Vol. 1, 2015, pp. 28-36.

[2] Z. Razak, N. J. Ibrahim, M.Y.I. Idris, and E.M. Tamil, "Quranic Verse Recitation Recognition Module for Support in J-QAF Learning: A Review ", *IJCSNS.*, Vol. 8, 2008.

[3] I. Sabir and N. Alsaeed, "A Brief Description of Consonants in Modern Standard Arabic", *Linguistics and Literature Studies.*,Vol. 2, No. 7 , 2014, pp. 185-189.

[4] F. Ye and J. Yang, " A Deep Neural Network Model for Speaker Identification", *Appl. Sci*, Vol. 11, No. 8, 2021.

[5] M. Algabri, H. Mathkour, M.A. Bencherif, M. Alsulaiman, and M.A. Mekhtiche, "Automatic Speaker Recognition for Mobile Forensic Applications" *Mob. Inf, Syst.*, Vol. 2017, 2017, pp. 1-6. https://doi.org/10.1155/2017/6986391.

[6] M. Fabien and P. Motlicek, "Open-Set Speaker Identification Pipeline in Live Criminal Investigations", *Proceedings of 2021 ISCA Symposium on Security and Privacy in Speech Communication,* November 10-12, 2021.

[7] P.J. Deng Barger and S. Sridharan, "On the Performance and Use of Speaker Recognition Systems for Surveillance. *Proceedings of 2006 IEEE International Conference on Video and Signal Based Surveillance*, Sydney, NSW, Australia, November 22-24 , 2006.

[8] F. Alegre, G. Soldi, N. Evans, B. Fauve, and J . Liu, "Evasion and Obfuscation in Speaker Recognition Surveillance and Forensics", *Proceedings of the 2nd International Workshop on Biometrics and Forensics*, Valletta, Malta, March 27-28, 2014.

[9] A. Mohammed, M.S. Sunar, and M.S. Hj Salam, "Quranic Verses Verification Using Speech Recognition Techniques", *J. Teknol.*, Vol. 73, 2015, pp. 99-106.

[10] A. Hafzalla and S. Mahdi, "Verification System for Quran Recitation Recordings", *IJCA.*, Vol. 163, 2017, pp. 6-11.

[11] S.S. Alrumiah and A.A. Al-Shargabi, "Intelligent Quran Recitation Recognition and Verification: Research Trends and Open Issues", *Arab J Sci Eng*., 2022.

[12] N. Khalid *et a*l., "AN EFFICIENT HOLY QURAN RECITATION RECOGNIZER BASED ON SVM LEARNING MODEL", *JJCIT.,* Vol. 06, No. 04, 2020.

[13] M.A Hussaini and R.W Aldhaheri, "An Automatic Qari Recognition System", *Proceedings of 2012 International Conference on Advanced Computer Science Applications and Technologies (ACSAT)*, Kuala Lumpur, Malaysia, November 26-28 , 2012.

[14] J. Ali, R. Khan, N. Ahmad, and I. Maqsood,
" Random Forests and Decision Trees", *IJCSI.*, Vol .9, No. 3, 2012.

[15] S.M. Shah and S.N. Ahsan, "Arabic Speaker Identification System Using Combination of DWT and LPC Features", *Proceedings of 2014 International Conference on Open Source Systems & Technologies*, Lahore, Pakistan, December 18-20 , 2014, pp. 176-181.

[16] T.S. Gunawan, N.A. Muhamat Saleh, and M. Kartiwi, "Development of Quranic Reciter Identification System Using MFCC and GMM Classifier", *IJECE.*, Vol. 8, No. 1, 2018, pp. 372-378.

[17] A. Elnagar, R. Ismail, B. Alattas, and A. Alfalasi, " Automatic Classification of Reciters of Quranic Audio Clips", *Proceedings of 2018 IEEE/ACS 15th International Conference on Computer Systems and Applications (AICCSA),* Aqaba, Jordan, 2018, Pp. 1-6.

[18] K.M.O Nahar, M. Al-Shannaq, A. Manasrah, R. Alshorman, and I. Alazzam, "A Holy Quran Reader/Reciter Identification System Using Support Vector Machine", *IJMLC.*, Vol. 9, No. 4, 2019, pp. 458-464.

[19] A. Qayyum, S. Latif, and J. Qadir, "Quran Reciter Identification: A Deep Learning Approach", *Proceedings of 2018 7th International Conference on Computer and Communication Engineering (ICCCE)*, Kuala Lumpur, Malaysia, September 19-20 , 2018, pp. 492-497.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[20] M.A. Al-Jarrah *et al.*, "Accurate Reader Identification for the Arabic Holy Quran Recitations Based on an Enhanced VQ Algorithm", *Rev. Intell. Artif.*, Vol. 36, 2022, pp. 815-823.

[21] H. Beigi, Fundamentals of Speaker Recognition , Springer: New York, NY, USA, 2011.

[22] M. Labied, A. Belangour, M. Banane, and A. Erraissi, "An Overview of Automatic Speech Recognition Preprocessing Techniques", *Proceedings of 2022 International Conference on Decision Aid Sciences and Applications (DASA)*, Chiangrai, Thailand, March 23-25, 2022. pp. 23-25 .

[23] N.N. Lokhande, N.S. Nehe, and P.S. Vikhe, "Voice Activity Detection Algorithm for Speech Recognition Applications", *Proceedings of International Conference in Computational Intelligence (ICCIA) 2011*, January, 2011.

[24] T.S. Mohammed, K.M. Aljebory, M.A. Abdul Rasheed, M.S. Al-Ani, and A.M. Sagheer, "Analysis of Methods and Techniques Used for Speaker Identification, Recognition, and Verification: A Study on Quarter-Century Research Outcomes", *Iraqi J. Sci.*, Vol. 62, 2021, pp. 3256-3281.

[25] A. Fazel and S. Chakrabartty, "An Overview of Statistical Pattern Recognition Techniques for Speaker Verification", *IEEE Circuits Syst. Mag.*, Vol. 11, 2011, pp. 62-81.

[26] B. Li, "On identity authentication technology of distance education system based on voiceprint recognition", *Proceedings of the 30th Chinese Control Conference*, Yantai, China, July 22–24 , 2011, pp. 5718–5721.

[27] B. Zhen *et a*l., "On the Importance of Components of the MFCC in Speech and Speaker Recognition", *Proceedings of 6th International Conference on Spoken Language Processing (ICSLP 2000)*, Beijing, China, October 16-20, 2000.

[28] P. Mishra and P.K. Mishra, "A Study of Various Speech Features and Classifiers Used in Speaker Identification", *IJERT*., Vol. 5, 2016.

[29] N. Chauhan, T. Isshiki, and D. Li, "Speaker Recognition Using LPC, MFCC, ZCR Features with ANN and SVM Classifier for Large Input Database", *Proceedings of 2019 IEEE 4th International Conference on Computer and Communication Systems (ICCCS)*, Singapore, February 23-25 , 2019, Pp. 130-133.

[30] Y. Kawakami, L. Wang, A. Kai, and S. Nakagawa, "Speaker Identification by Combining Various Vocal Tract and Vocal Source Features", *Proceedings of 17th International Conference, TSD 2014*, Brno, Czech Republic, Sept 8–12 , 2014, pp. 382-389.

[31] F.E. AbuAladas and V.W. Samawi, "Speaker Identification Using Selected Features from DWT", *Proceedings of 2013 Intern-tional Conference on Advanced Computer Science Applications and Technologies*, Kuching, Malaysia, 2013.

[32] M.M. Al Anazi and O.R. Shahin, "A Machine Learning Model for the Identification of the Holy Quran Reciter Utilizing K-Nearest Neighbor and Artificial Neural Networks*", Inf. Sci. Lett.*,Vol. 11, No. 4, 2022, pp. 1093-1102.

[33] S.S. Tirumala, S.R. Shahamiri, A.S. Garhwal, and R. Wang, "Speaker Identification Features Extraction Methods: A Systematic Review. Expert Syst", *Appl*, Vol. 90, 2017, pp. 250-271.

[34] F.Z. Chelali, K. Sadeddine, and A. Djeradi, "A. Speaker Identification System Using LPC - Application on Berber Language", *HDSKD*., Vol. 1, No. 2, 2015, pp. 29-46.

[35] M. Chandra, P. Nandi, A. Kumari, S. Mishra, '. Spectral-Subtraction Based Features for Speaker Identification", *Proceedings of the 3rd International Conference on Frontiers of Intelligent Computing: Theory and Applications (FICTA) 2014*, Bhuba-neswar, India, November 4-15 , 2014.

[36] S. Kalaivani, "Modified Hidden Markov Model for Speaker Identification System", *IJACEE*., Vol. 2, 2017, pp. 1-7.

[37] D.A. Reynolds, "Automatic Speaker Recognition Using Gaussian Mixture Speaker Models", *Linc. Lab. J.*, Vol. 8, No. 2, 1995, pp. 91-108.

[38] A. Maurya, A, D. Kumar, and R. Agarwal, "Speaker Recognition for Hindi Speech Signal Using MFCC-GMM Approach", *Procedia Comput. Sci.,* Vol. 125 , 2018, pp. 880-887.

[39] V.R. Apsingekar and P.L. De Leon, "Support Vector Machine Based Speaker Identification Systems Using GMM Parameters" *Proceedings of 2009 Conference Record of the Forty-Third Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, USA, Nov 1-4 , 2009.pp. 1766-1769.

[40] H.B. Kekre *et al*., "Speaker Recognition Using Vector Quantiza-tion by MFCC and KMCG Clustering Algorithm", *Proceedings of 2012 International Conference on Communication,*

In-formation & Computing Technology (ICCICT), Mumbai, India, Oct 19-20 , 2012.

[41] Ranny, "Voice Recognition Using K Nearest Neighbor and Double Distance Method" *Proceedings of 2016 International Conference on Industrial Engineering, Management Science and Application (ICIMSA)*, Jeju, South Korea, May 23-26 ,2016, pp. 1-5.

[42] J. Kacur, R. Vargic, and P. Mulinka, P, "Speaker Identification by K-Nearest Neighbors: Application of PCA and LDA Prior to KNN", *Proceedings of 2011 18th International Conference on Systems, Signals and Image Processing*, Sarajevo, Bosnia and Herzegovina, June 16-18 ,2011.

[43] M.M. Hossain, B. Ahmed, and M.A. Asrafi, "Real Time Speaker Identification Using Artificial Neural Network", *Proceedings of 2007 10th International Conference on Computer and Information Technology*, Dhaka, Bangladesh, Dec 27-29 , 2007, Pp. 1-5.

[44] Rateel - Apps on Google Play [Internet]. play.google.com.[cited 2023 May 15]. Available from:https://play.google.com/store/apps/details?id=sa.com.stc.Rateel&hl=en&gl=US

[45] Nadyy: Identify Famous Quran R - Apps on Google Play [Internet]. play.google.com. [cited 2023 May 15]. Available from: https://play.google.com/store/apps/details?id=com.nadyy&hl=en&gl=US

[46] M. Lataifeh and A. Elnagar, "Ar-DAD: Arabic Diversified Audio Dataset", *Data Br*, Vol. 33, 2020.

[47] EveryAyah Quran Files. Available online: http://www.everyayah.com.

[48] H. Khan, A.A. Allaban, and M. Moussa, "The Tarteel dataset: crowdsourced and labeled Quranic recitation", *Proceedings of 35th Conference on Neural Information Processing Systems (NeurIPS 2021)*, Sydney, Australia, Dec 6-14, 2021, pp. 1–12 .

[49] A. Moustafa and S.A. Aly, "Towards an Efficient Voice Identification Using Wav2Vec2.0 and HuBERT Based on the Quran Reciters Dataset", *ArXiv*. 2021.

[50] W.M. Muhammad, R. Muhammad, A. Muhammad, and A. Martinez-Enriquez, "A. Voice Content Matching System for Quran Readers", *Proceedings of 2010 Ninth Mexican International Conference on Artificial Intelligence*, Pachuca, Mexico , Nov 8-13 ,2010, pp. 148-153.

[51] T. Giannakopoulos, "A method for silence removal and segmentation of speech signals", MATLAB Central File Exchange, University of Athens, 2023.

[52] S. Hourri and J.A. Kharroubi, "Deep Learning Approach for Speaker Recognition", *Int. J. Speech Technol*., Vol. 23, 2019, pp. 123-131.

[53] R.U. Khan, A.M. Qamar, and M. Hadwan, "Quranic Reciter Recognition: A Machine Learning Approach", *Adv. Sci. Technol. Eng. Syst. J.* , Vol. 4, No. 6, 2019, pp. 173-176.

[54] J.H. Alkhateeb, "A Machine Learning Approach for Recognizing the Holy Quran Reciter", *IJACSA*., Vol. 11, No. 7, 2020, pp. 268-271.

[55] M. Lataifeh *et al*., "Arabic audio clips: Identification and discrimination of authentic Cantillations from imitations', *Neurosci. J.,*Vol. 418, 2020, pp. 162–77..

[56] A. Elnagar and M. Lataifeh, "Predicting Quranic Audio Clips Reciters Using Classical Machine Learning Algorithms: A Comparative Study", *SCI.* , Vol. 874, 2019, pp. 187-209.

[57] M. Hadwan, A. H. Alsayadi, and S. AL-Hagree , "An end-to-end transformer-based automatic speech recognition for qur'an reciters", *CMC.*, Vol. 74, No. 2, 2023..

[58] G. Samara *et al.*," The Recognition of Holy Qur'an Reciters Using the MFCCs' Technique and Deep Learning", *Advances in Multimedia*, vol. 2023, 2023.

[59] M.R. Firmansyah, R. Hidayat, and A. Bejo, "Comparison of Windowing Function on Feature Extraction Using MFCC for Speaker Identification", *Proceedings of 2021 International Conference on Intelligent Cybernetics Technology & Applications (ICICyTA)*, Bandung, Indonesia, Dec 1-2 ,2021

[60] S. Ali *et al.,* "Mel Frequency Cepstral Coefficient: A Review. In Proceedings of the 2nd Interna-tional Conference on ICT for Digital, Smart, and Sustainable Development", *ICIDSSD 2020*, New Delhi, India, Feb 27-28 ,2020.

[61] M.A. Mazumder and R.A. Salam,"Feature Extraction Techniques for Speech Processing: A Review", *IJATCSE*., Vol. 8, No. 1.3, 2019, pp. 285-292.

[62] B.N. Ishaq and B.W. Gawali, Comparative Analysis of MFCC, DTW&ANN for Arabic Speech Recognition. *IJIRAE*., Vol. 1, No. 11, 2014, pp. 56-61.

[63] M. Bezoui, A. Elmoutaouakkil, A. Benihssane, "Feature Extraction of Some Quranic Recitation

Using Mel-Frequency Cepstral Coeficients (MFCC)", *Proceedings of 2016 5th International Conference on Multimedia Computing and Systems (ICMCS)*, Marrakech, Morocco, September 29 – October 1, 2016.

[64] J. Huang *et al*., "Image Indexing Using Color Correlograms", *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, San Juan, Puerto Rico, USA, 1997.

[65] J. Joshy and K. Sambyo, "A Comparison and Contrast of the Various Feature Extraction Techniques in Speaker Recognition", *IJIPPR*. , Vol. 9, No. 11, 2016, pp. 99-108.

[66] V. Sharma and P.K. Bansal, "A Review On Speaker Recognition Approaches And Challenges", *IJERT*., Vol. 2, 2013, pp. 1581-1588.

[67] P. Sangwan, D. Sheoran, and S. Bhardwaj, "Speech Recognition Using Wavelet Based Feature Extraction Techniques", *Glob. J. Enterp. Inf.* Syst., Vol. 9, 2017.

[68] W.-N. Hsu *et al*., "HuBERT: Self-Supervised Speech Repre-sentation Learning by Masked Prediction of Hidden Units" *IEEE/ACM Trans. Audio Speech Lang.*, Vol. 29, 2021, pp. 3451-3460.

[69] R. Badgujar *et al*., "Speaker Identification Using Artificial Neural Network", *IJERT*., Vol. 3, 2015.

[70] J. Zupan, " Introduction to Artifcial Neural Network (ANN) Methods: What They Are and How to Use Them", *Acta Chim. Slov.* ,Vol. 41, 1994, pp. 327-352.

[71] W.S. Noble, "What Is a Support Vector Machine?", *Nat. Biotechnol., Vol.* 24, pp. 1565-1567.

[72] R. Vankayalapati *et al*., K-Means Algorithm for Clustering of Learners Per-formance Levels Using Machine Learning Techniques", *Rev. Intell. Artif.*, Vol. 35, 2021, pp. 99-104.

[73] Z. Lei, Y. Yang, and Z. Wu, ," Speaker Identification Using the VQ-Based Discriminative Kernels", *Proceedings of International Conference on Audio- and Video-Based Biometric Person Authentication*, NY, USA, Jul 20-22, 2005.

[74] M.A.U. Khan and M.J.T. Smith, "Fundamentals of Vector Quantization", In Handbook of Image and Video Processing", 2nd ed., A. Bovic, Eds, Academic Press: New York, 2005; pp. 673-688.

[75] M. Deng *et al*., "Heart Sound Classification Based on Improved MFCC Features and Convolutional Recurrent Neural Networks", *Neural Netw.,* Vol. 130, 2020, pp. 22-32.

[76] M. Fontaine, S.A. Love, and M. Latinus, "Familiarity and Voice Representation: From Acoustic-Based Representation to Voice Averages", *Front. Psychol*., Vol. 8, 2017.

[77] A.A. Mohamed, A. Eltokhy, and A.A. Zekry, "An Overview of the Development of Speaker Recognition Techniques for Vari-ous Applications", *J. Commun.* , Vol. 17, 2022, pp. 632-642.

[78] Y. Lecun *et al*., "Gradient-Based Learning Applied to Document Recognition*", Proceedings of the IEEE*, Nov 1998.

[79] S.S. Tirumala and S.R. Shahamiri, "A Review on Deep Learning Approaches in Speaker Identification", *Proceedings of the 8th International Conference on Signal Processing Systems 2016*, Auckland , New Zealand November 21-24, 2016, pp. 142-147.

[80] ختمة Khatmah - مصحف،أذان،أذكار - Apps on Google Play [Internet]. play.google.com. [cited 2023 May 15]. Available from: https://play.google.com/store/apps/details?id=com.khatmah.android&hl=en&gl=US

[81] Tarteel: Quran A.I. - Apps on Google Play [Internet]. play.google.com. [cited 2023 May 15]. Available from: https://play.google.com/store/apps/details?id=com.mmmoussa.iqra&hl=en&gl=US

[82] N. Kamarudin *et al*., "Acoustic Echo Cancellation Using Adaptive Filtering Algorithms for Quranic Accents (Qiraat) Identification", *Int. J. Speech Technol.*, Vol. 19, 2016, pp. 393-405.

[83] X. Yuan et al., "Overview of the Development of Speaker Recognition", *JPCS.* 2021.

[84] H. Hermansky and N. Morgan, , "RASTA Processing of Speech", *IEEE Trans Audio Speech Lang Process*, Vol. 2, No. 4, 1994, pp. 578-589.

[85] H. Hermansky, 'Perceptual Linear Predictive (PLP) Analysis of Speech', *J. Acoust. Soc. Am.*, Vol. 87, No. 4, 1990, pp. 1738-1752.

[86] S. Boll, 'Suppression of Acoustic Noise in Speech Using Spectral Subtraction', *IEEE Trans Acoust Speech Signal Process., Vol. 27,* 1979, pp. 113-120.