

Interactive Arabic CAPTCHAs with Enhanced Usability and Accessibility

Hanan A. Al-Mutawa*, Suliman Aladhadh*, Mohammad T. Parvez**

*Department of Information Technology, College of Computer, Qassim University, Saudi Arabia

**Department of Computer Engineering, College of Computer, Qassim University, Saudi Arabia

441212465@qu.edu.sa; s.aladhadh@qu.edu.sa; m.parvez@qu.edu.sa

Abstract—CAPTCHAs are widely used as a security measure to protect authorized users and prevent unauthorized automated access. Their effectiveness is typically assessed by robustness against robotic programs. Nevertheless, user perspectives on the interactive experience and efficacy of CAPTCHAs have received limited attention, particularly in relation to comparative analyses of CAPTCHA variations across different locations and languages worldwide. This study enhances interactive Arabic CAPTCHA systems by incorporating features to support users with disabilities. To improve accessibility, we introduce audio-based CAPTCHA challenges that present spoken questions instead of text, with background noise added for security while ensuring usability for visually impaired individuals. Additionally, the system integrates localized Arabic accents rather than standard Arabic, making it more intuitive and user-friendly across regions. Extensive experiments with users with disabilities demonstrated improved usability and high success rates. These enhancements foster greater digital accessibility without compromising security, contributing to a more inclusive CAPTCHA system.

Index Terms—CAPTCHA, Accessibility, Usability, Variability

I. INTRODUCTION

The Internet has been an essential approach and play a significant role into both personal and professional life by transforming how individuals' access and share information (17). Recently, services that Internet offers and delivers are increasingly expanding across critical fields and sectors (such as health, education, finance, and government services in addition to others) and assist in reshaping such industries (1; 14). In light of the extensive reliance on Internet services, comprehending and addressing the related security challenges, including the development of efficient protective measures, has emerged as an important concern.

As the Internet becomes indispensable in our lives, the threats of vulnerability to cyber-attacks become not only real but also imminent (19). These include the threat of a personal impersonation, eavesdropping on personal communication or on communication between the individual and the Internet site (i.e., *e-services*), and exposure to attacks directed by robots. To prevent, or at least minimize these threats, researchers have used various methods in order to differentiate robots from humans, including human interactive proofs (HIPs). Research in this area has led to the adoption of a method called CAPTCHA, which is an acronym for Completely Automated Public Turing Test to Tell Computers and Humans Apart, to fend off these attacks and raise the system degree of resilience.

The idea behind a CAPTCHA is to enhance service security by ensuring that client requests originate from actual individuals rather than malicious malware through online operations. This software generates and categorizes assessments that are easy for individuals to solve but difficult for computer programs and automated bots (4).

To address various needs, several CAPTCHA types have been developed, including auditory, visual, and text-based CAPTCHA types. Among such types, text-based CAPTCHAs are the most common due to their distinct advantages over others. Furthermore, two main factors (usability and security) should be taken into consideration when designing any type of CAPTCHA. Usability is an element that ensures people can solve CAPTCHA challenges easily, whereas security aims to make CAPTCHAs more difficult in order to protect websites from robot hackers or automated bots. Achieving a balance between security and usability remains a significant challenge in CAPTCHA design (3).

To serve a broad spectrum of users, solving CAPTCHA tests must be feasible for all users including individuals with visual impairments. In fact, based on the World Health Organization, approximately 36 million people worldwide are completely blind, while around 217 million others experience varying degrees of visual impairment. Solving CAPTCHA tests is often consider to be annoying for users without disabilities, making it even more challenging (or even impossible) for users with disabilities. Therefore, CAPTCHA systems must be designed to be effective and accommodate all users, regardless of disability, as this technology primarily seeks to protect electronic services against automated bots (3).

Previous usability studies (4; 7; 8; 16) have investigated factors that influence CAPTCHA solving success, such as participant age, distortion mechanisms, and cognitive abilities. However, such studies often overlooked or were oblivious crucial usability aspects, such as accurately measuring user accuracy, performance, and standardizing testing devices.

Further research has explored methods to enhance the robustness of text-based CAPTCHA. For instance, (11) and (18) have proposed increasing distortion levels through mechanisms such as character length, slant, and word randomization. While these approaches improve security against bots, they may often negatively impact user experience. Thus, finding a balance between security and usability remains an essential challenge

for effective CAPTCHA design.

Accessibility research, particularly by (2; 3; 10) have highlighted challenges faced by users with disabilities, such as the difficulties visually impaired individuals encounter when solving text and image CAPTCHAs. Despite the proposal audio clips have been proposed to improve accessibility, providing an alternative approach for users to replay audio clips for better comprehension has frequently been overlooked.

CAPTCHA technology, while effective in differentiating humans from bots, still faces challenges in achieving optimal security and usability. Several existing solutions often lack the robustness required to counter sophisticated bots and fail to provide full accessibility for all users. This research aims to address these limitations by developing an improved framework for interactive CAPTCHAs.

The primary objective of this work is to design a system that generates and presents challenging yet user-friendly CAPTCHA tests, addressing the needs of both able-bodied and disabled users. To achieve this, we propose an enhanced CAPTCHA framework focused on accessibility, analyze its effectiveness, and evaluate the results. The contributions of this work are as follows:

- Improved accessibility through novel audio-based CAPTCHA challenges with background noise, tailored for visually impaired users while enhancing security.
- Increased usability for all users across different regions by incorporating local Arabic accents instead of standard Arabic in CAPTCHA questions.

The remainder of this paper is organized as follows: Section II presents related works. Section III describes the proposed CAPTCHA system. Section IV outlines the evaluation methodology. Section V presents the experimental results. Section VI discusses the findings obtained. Finally, Section VII concludes the paper and highlights future research directions.

II. RELATED WORKS

This section presents a review of research that has employed CAPTCHA as a user verification method, along with an evaluation of the features proposed by each CAPTCHA in the respective research articles. The following sub-sections examine studies on usability, variability, and accessibility as three important features of CAPTCHA in relevant research work.

A. Usability in CAPTCHAs

(16) explored improving CAPTCHA usability on smartphones by modifying input methods. They compared five existing CAPTCHA schemes with four novel prototypes that replaced traditional keyboard input with gestures, image displacement, or menu choices. Their study found that text-based schemes using a traditional keyboard had the highest success rates and fastest completion times. Interestingly, participants preferred traditional keyboard input over the novel methods. However, the study has some limitations their prototypes lacked error correction mechanisms, potentially influencing

user preference and impacting usability. Essentially, the study revealed that simply changing the input method doesn't necessarily improve mobile CAPTCHA usability and highlighted the need for robust error correction in alternative input designs.

(4) tackled the tension between CAPTCHA security and usability by proposing an optimization algorithm for text-based CAPTCHAs. The algorithm identifies and replaces confusing characters with more distinguishable ones, aiming to improve readability. A user study showed the optimization significantly increased accuracy (from 57.54% to 95.7%) and decreased response time (from 8.72 to 5.20 minutes). While demonstrating improved usability, the study's limitation is the lack of information about the devices used by participants, making device standardization a recommendation for future research. Essentially, the novelty is the character-level optimization algorithm to enhance CAPTCHA readability without sacrificing security.

(7) explored how text presentation in CAPTCHAs and user demographics affect usability, specifically response times. They tested four CAPTCHA types, including text-based ones, and analyzed the impact of age, gender, internet experience, and education on user performance. Using 100 participants and employing frequent pattern growth and k-medians clustering, they found significant response time variations across demographics. Younger users were faster with text-based CAPTCHAs, and age was a stronger predictor than education. While insightful regarding response times, the study's key limitation is its lack of accuracy measurements, preventing a complete picture of CAPTCHA usability. The novelty lies in their demographic analysis of CAPTCHA response times using specific data mining techniques.

(8) examined how user demographics and experience affect text-based CAPTCHA usability on tablets. They analyzed response times and success rates for 125 participants across different age, gender, education, and internet experience groups. Using K-medians clustering and the apriori algorithm, they found that users with intermediate internet experience performed best, while users over 35 struggled, regardless of other characteristics. The study's main limitation is the lack of a numerical accuracy metric, hindering a complete usability assessment. The novelty lies in its focus on tablet-based CAPTCHA usability, specifically considering the interplay of demographics and experience.

B. Variability in CAPTCHAs

(11) introduced a new text-based CAPTCHA designed to resist OCR attacks. Their system combines an alphanumeric phrase with a randomly generated algebraic equation, adding complexity through varied font styles, randomized character alignment, and non-lexical words. This multi-layered approach proved effective against OCR in their tests, achieving a 95% success rate across 100 samples. The study acknowledges a potential usability issue, as the mathematical equation might be difficult for some users. The novelty is the integration of a mathematical equation into the CAPTCHA design to bolster security against OCR.

A prior study by (18) have examined how text rotation, string length, and letter format (handwritten vs. typed) affect text-based CAPTCHA robustness against OCR attacks. Using OCR software to test CAPTCHA weakness, the study found all three variables significantly impacted robustness. The most robust CAPTCHAs were 10 characters long, rotated at +45 degrees, and handwritten, with only 39.7% of samples successfully deciphered. While noting longer strings improve security but potentially harm usability, the study's novelty lies in its experimental analysis of these specific design variables on CAPTCHA robustness using OCR.

C. Multi-feature CAPTCHAs

Here we investigate studies that target to balance between the usability and the variability of the challenges through different approaches. The researchers used different techniques in CAPTCHA tests, including the use of distortion, three-dimensional (3D), and visual cryptography (VC) technology, etc.

A prior study by (12) examined how different CAPTCHA distortion methods affect user performance, specifically comparing younger and older adults. They tested six distortion types: normal text, blot mask, line mask, thread noise, global warp, and geometry noise. Results showed older adults were slower and made more errors. Blot and line masks were the most challenging distortions, while thread noise, global warp, and geometry noise were similar in difficulty to normal text. Normal text CAPTCHAs had the best usability (fastest times, lowest errors), while line mask CAPTCHAs were the most robust (slowest times, highest errors). The study highlights the difficulty of balancing CAPTCHA usability and security across age groups. A limitation was potential language ambiguity for Chinese-speaking participants. The novelty lies in its direct comparison of various distortion techniques' impact on different age groups' CAPTCHA solving performance.

Further research introduced a new CAPTCHA interface designed specifically for mobile devices, addressing the need for security and user convenience research (13). Traditional CAPTCHAs, while effective for desktop environments, often fail to provide a seamless experience on mobile devices. The researchers developed a new image-based CAPTCHA called "CAPTCHA Zoo," which leverages the touch-screen capabilities of mobile devices. This CAPTCHA requires users to identify and select specific animals from a cluttered background, making it both user-friendly and secure. The novelty of their approach lies in the use of 3D models and natural backgrounds to create a more challenging and engaging CAPTCHA experience. Experimental results demonstrated that CAPTCHA Zoo is effective in distinguishing human users from bots while maintaining a high level of usability. This study highlights the importance of designing CAPTCHAs that are not only secure but also tailored to the unique interaction patterns of mobile device users.

A previous study by (15) analyzed the security of three 3D CAPTCHA systems (super CAPTCHA, 3dcaptcha, and teabag 3D) by attempting to extract information to enable OCR-based

attacks. Despite the difficulty of directly applying OCR to 3D CAPTCHAs, they found exploitable patterns. Their results showed varying attack success rates, with teabag 3D being the most vulnerable (76% success within 7 seconds) and super CAPTCHA the most resilient (27% success within 3 seconds). According to the study, increasing line distortion parameters or crowding and overlapping of characters will almost surely make the CAPTCHA harder to solve, which in turn increases its robustness. Therefore, such a technique negatively affects the ease of use of the CAPTCHA. The novelty lies in their approach of circumventing the 3D nature of the CAPTCHAs to facilitate OCR-based attacks, demonstrating vulnerabilities even in 3D designs.

(5) explored the usability and security of a new Arabic text-based CAPTCHA using visual cryptography (VC). The VC method encrypts Arabic text into two parts that users must overlay to decipher it. A user study with 169 participants showed that handwritten text from the same writer was easiest to solve, while random handwritten text was the hardest. Furthermore, the security was tested against various machine learning algorithms, with Naïve Bayes performing the worst (0.06% accuracy) and CNN the best (56% accuracy), though still relatively low. The study demonstrates the potential of VC-based Arabic CAPTCHAs, but notes that familiarity with VC and similar Arabic letters (e.g.) might influence results. Future work should address the similar letter issue.

Earlier research by (9) examined user performance with two common types of CAPTCHAs, specifically focusing on text- and image-based challenges. The researchers found that participants resolved image challenges faster and preferred them over text challenges. Despite the high error rates for both types, there was no significant difference between them. The study highlights the need for new methods to distinguish humans from robots without relying on user intervention. The novelty of their research lies in the low-fidelity methodology used to simulate real CAPTCHA experiences, ensuring replicability and providing insights into user preferences and performance. This approach underscores the importance of balancing security, usability, and accessibility in CAPTCHA design. Usability was assessed through response time, error rate, and user preferences. Response time for each challenge was recorded, and error rates were calculated for each type of CAPTCHA. User preferences were gathered through simple post-experiment questionnaires. Security was evaluated by analyzing how difficult it was for bots to bypass the CAPTCHAs. The study did not specifically focus on accessibility for individuals with disabilities.

D. Accessibility in CAPTCHAs

The study of (10) investigates the development of an accessible CAPTCHA for visually impaired users. Recognizing the inherent limitations of traditional text-based CAPTCHAs for this population, a novel bimodal approach incorporating both visual and auditory information was designed and implemented. Preliminary evaluation suggests the potential efficacy of this new CAPTCHA for both blind and sighted individuals. A key design challenge lies in optimizing the balance between usability and security, particularly for users with disabilities. To address this, the developed CAPTCHA tool utilizes a combination of pictorial and auditory stimuli, aiming to enhance both accessibility and security. Evaluation involving both blind and sighted participants demonstrated general user satisfaction and efficient task completion among blind users. The primary contribution of this research lies in its innovative integration of visual and auditory elements to create a more accessible and secure CAPTCHA, while also revealing vulnerabilities in existing 3D CAPTCHA designs. This approach offers potential improvements to security design applicable to a broader user base.

During the study by (2), which aimed to address the limitations of traditional text and image CAPTCHAs for users with disabilities, particularly the deaf and blind users. The study proposes two new gesture-based CAPTCHA systems: SeeAct (image-based questions) and HearAct (sound-based object identification) to improving accessibility for users with visual and auditory impairments. Both systems use click/swipe gestures for responses. As user study with blind, deaf, and sighted participants showed SeeAct and HearAct had a much higher success rate (90%) and faster completion times (27 seconds) than ReCAPTCHA (60% success, 1 minute 29 seconds). HearAct was more accessible for blind users (80% success), while SeeAct was preferred by sighted and deaf users (95% success). The study suggests gesture-based CAPTCHAs improve accessibility. A recommended improvement is adding a replay feature for audio and text instructions.

(3) recognized the limitations of traditional CAPTCHA systems for visually impaired users, particularly the difficulty in memorizing and solving audio-based CAPTCHAs. To address these challenges, they proposed a novel audio-visual CAPTCHA system called "IRemember". IRemember presents users with a common English phrase (e.g., "How are you?") both visually and aurally. Users are then tasked with identifying and entering the first letter of each word in the phrase (e.g., "hay"). To enhance accessibility, the system utilizes distinct sounds for the CAPTCHA question and the presented phrase. Furthermore, the system incorporates a curated list of 250 frequently used English phrases and adds background noise to enhance security against automated attacks. An experimental study involving 11 blind participants compared IRemember with traditional audio CAPTCHA. Results demonstrated that IRemember significantly outperformed traditional audio CAPTCHA in terms of both success rate (82.72% for IRemember vs. 48.18% for audio CAPTCHA)

and solving time (54.96 seconds for IRemember vs. 138.75 seconds for audio CAPTCHA). Limitations of the study include the limited number of participants and the exclusive use of the English language. Future research should address these limitations by expanding the participant pool and incorporating multilingual support to enhance the global accessibility of the IRemember CAPTCHA system.

III. PROPOSED CAPTCHA SYSTEM

The proposed CAPTCHA system that we developed is to enhance accessibility and usability of the prior studies, to be incorporated into the CAPTCHA tests.

To fully grasp the accessibility features our proposed work, we first analyze the previous studies in terms of ease of use, variability, and accessibility. The method applied in (6) is an interactive Arabic CAPTCHA system designed to enhance security against automated attacks. The system incorporates two key components: a dynamic challenge pool and an attack filtering mechanism. In addition, it presents users with a set of CAPTCHA challenges with varying difficulty levels. Raw CAPTCHA samples are subjected to a series of distortions (i.e., *word displacement*, *contour redrawing*, *noise addition*) at two different levels (Figure 1 provides an example). These distorted samples are then evaluated by Google Cloud OCR. CAPTCHA samples that successfully resist OCR attacks are classified and added to the challenge pool based on their difficulty level.

An experimental study involving 144 participants assessed the usability of the proposed CAPTCHA. Participants solved ten CAPTCHA challenges, which included binary, numerical, and image-based tasks. Results demonstrated a high overall accuracy (93.1%) and an average solving time of 7.01 seconds. To evaluate the security of the proposed scheme, a custom recognizer combining deep network-based noise removal and holistic word recognition achieved an extremely low accuracy of 0.11% in breaking the CAPTCHA. This high level of security is attributed to the effective distortion techniques applied to the CAPTCHA images.

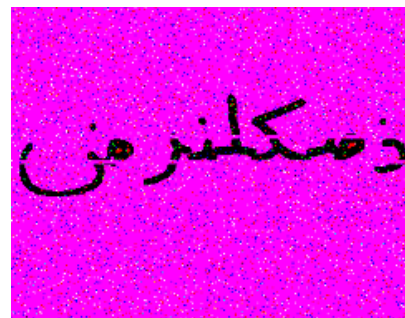


Fig. 1: Illustrations from CAPTCHAs as used in prior study(6).

The approach adopted by (6) is limited by its failure to consider the accessibility of the CAPTCHA system for users with disabilities. In order to improve the algorithm's efficacy and usability, particularly for users who are visually impaired,

we propose implementing an audio question with added background noise to increase the level of security, rather than displaying the question as written text for each CAPTCHA image, as has been previously done. Since the language used in the CAPTCHA images, including the questions, is Arabic, we propose incorporating the local accent of each country instead of standard Arabic to enhance the usability of the CAPTCHA tests. Therefore, the contributions of the present study can be summarized as follows:

- **Enhanced Accessibility:** The proposed adaptation improves accessibility by introducing audio-based CAPTCHA challenges with background noise, catering to visually impaired users while enhancing security.
- **Localized Usability:** By incorporating local Arabic accents instead of standard Arabic in CAPTCHA questions, the system increases ease of use and familiarity for users across different regions.

IV. EVALUATION METHODOLOGY

During this section, we aim to cover the methodology we applied to validate and evaluated the CAPTCHA idea presented in the previous section. Figure 2 provides a summary of our methodology. Our objective in this work is to design an interactive CAPTCHA system that protects electronic services from automated bots while also addressing the requirements of both able-bodied users and users with disabilities. Developing a framework for optimal interactive CAPTCHA challenges involves establishing a system that efficiently presents challenges to users while collecting sufficient data for offline analysis in order to determine the ideal parameters for an interactive CAPTCHA system.



Fig. 2: Methodology used in this work.

The experiment involved a collection of Arabic CAPTCHA samples carefully selected to be solved by 89 participants, including individuals with and without disabilities. The procedure consisted of three phases: selecting CAPTCHA samples, designing the user interface, and recruiting participants for the experiment. At the conclusion of the experiment, participants were required to complete a closed-ended questionnaire with predefined options or scales to select from.

A. Building CAPTCHAs

The Hypertext Preprocessor (PHP) programming language was used to develop a CAPTCHA generator that selects 15 CAPTCHAs from a sample containing 919 Arabic CAPTCHAs. The sample size was determined based on the findings of (5). These CAPTCHAs presented various interactive challenges to be solved, allowing the collection of sufficient data to analyze the proposed CAPTCHAs. Additionally, native Arabic speakers recorded Arabic audio files for

all specified CAPTCHA challenges, incorporating a variety of Arabic accents, including Saudi Arabian, Egyptian, and Tunisian. The total number of Arabic audio files amounted to 136, comprising eight audio files for each challenge in the Saudi Arabian accent, eight for each challenge in the Egyptian accent, eight for each challenge in the Tunisian accent, and 28 files for each letter of the Arabic language.

B. User Interface

Visual Studio (VS) Code with HTML was used to create a graphical user interface (GUI). Users were required to listen to the Arabic audio CAPTCHA challenges generated by the system, view the corresponding CAPTCHA image, identify the correct answer, and enter it in the designated input field before submitting their responses (Figure 3 is the GUI proposed in this study).



Fig. 3: GUI

C. Evaluation Process

To measure time and accuracy in this study, both were assessed by recording the time for each CAPTCHA, starting from when participants began solving it until they submitted their responses. Response time was defined as the duration between completing the audio playback and pressing the submission button. This duration could vary depending on when participants began solving the CAPTCHA and submitted their answers. Accuracy was determined based on whether the submitted answer was correct or incorrect. Both response time and accuracy were recorded for each proposed CAPTCHA for subsequent analysis.

Upon completion of the experiment, quantitative measures were used to analyze the data (response time and accuracy) collected by the system for each CAPTCHA sample included in this study. These measures included the count, minimum (Min), maximum (Max), summation (Sum), and

average (AVG). Subsequently, the findings were evaluated and compared with relevant literature to assess whether the targeted features, in terms of ease of use, variability, and accessibility for participants, including those with disabilities, were achieved.

V. EXPERIMENTATION

During this study, various devices were used, along with the specifications of the computer employed for conducting the experiment and the development tools utilized, including their respective versions, as detailed in Table I.

TABLE I: Environment setup.

Item	Specifications
OS	Windows 10
CPU	Intel Core i7-5500U @ 2.40 GHz
RAM	8 GB
Visual Studio Code	v1.87.2
PHP	v8.2.12
phpMyAdmin	v5.2.1
cPanel	v110.0 (build 15)

To implement and evaluate the proposed system, we investigated the selected CAPTCHA samples, the CAPTCHA challenges, and the user workflows. The CAPTCHA system was designed and tested by realusers to assess its usability and security.

A. CAPTCHA Samples

The CAPTCHA images used in this investigation are identical to those employed in (5)'s study. These samples consist of images of random handwritten Arabic words created by various writers, with two levels of distortion applied to enhance their security. Furthermore, these images are accompanied by audio recordings. Figure 4 provides examples of CAPTCHA samples used in this research.

B. CAPTCHA Challenges

There is an opportunity to use multiple challenges per CAPTCHA. In this study, we aimed to use eight challenges (listed in Table II) within the proposed Arabic handwritten text-based CAPTCHA framework. Based on their resolution methods, the challenges are divided into two distinct categories: some require participants to provide a binary response (yes or no), while others require an integer response.



(a) CAPTCHA with distortion of level one



(b) CAPTCHA with distortion of level two

Fig. 4: Illustration of CAPTCHAs with two distortion levels as used in this work.

TABLE II: The CAPTCHA challenges utilized in the proposed system.

Approach of CAPTCHA Solving	Challenge
Binary answers	Is this character present in the image?
	Are any of these characters present in the image?
	How many characters are there in this word?
	How many characters with dots are there in this image?
Reading and counting	How many characters without dots are there in this image?
	How many isolated (not connected) characters are there?
	How many characters are connected to the left?
	How many characters are connected to the right?

All of the aforementioned challenges, as illustrated in Table II, were audio-recorded and developed in three regional Arabic accents: Saudi Arabian, Egyptian, and Tunisian. Additionally, these challenges were recorded by individuals of both genders (male and female) with background noise included to enhance

security. Furthermore, each letter of the Arabic language was recorded separately to be tested with the first and second challenges. It is worth noting that the selection of challenges is not fixed or limited to a specific solving approach, allowing for the addition or modification of challenges for future applications. Nevertheless, the current challenges were selected based on their popularity among users to ensure the system's ease of use.

C. User Workflows

To implement the enhancements proposed in this study, we utilized the PHP programming language along with VS Code to develop a responsive website compatible with various devices, ensuring accessibility for anonymous participants from Kingdom of Saudi Arabia, Egypt, and Tunisia. Initially, they were 89 anonymous participants provided basic demographic information, including age, gender, education level, vision level, experience of using Internet, and country

Then, five samples of CAPTCHA images, each paired with five challenges in a local Saudi Arabia accent, were presented to all participants. This was followed by another five CAPTCHA samples with five challenges in a local Egyptian accent, and finally, five more CAPTCHA samples with five challenges in a local Tunisian accent. Additionally, we ensured that the selection of CAPTCHAs and audio challenges for each participant was random, avoiding repetition of the same CAPTCHA with the same audio challenge in the same accent. For example, CAPTCHA x , with a challenge to y in a Saudi Arabian local accent, can also be repeated with the challenge y in a local Tunisian accent for a participant. Snapshots of the user interface used in the experiment are provided in Figure 5.

Finally, upon completing and solving all the tests, participants were directed to a Google Form to answer the questionnaire for this study, thereby evaluating their satisfaction with the proposed CAPTCHA system adopted in this research.

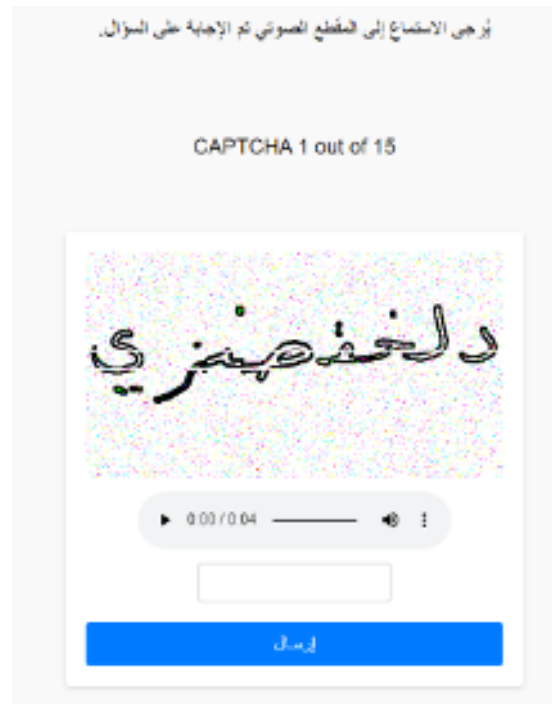
VI. RESULTS AND DISCUSSIONS

This section presents and analyses the results obtained from the experiment described in the previous section. The assessment of these results aims to evaluate the extent to which the study’s primary objective (developing a solution for Arabic text-based CAPTCHA challenges incorporating diverse local accents) has been achieved. This section is divided into two parts: the experimental results (section VI-A) and the discussion of our findings in relation to existing literature (section VI-B).

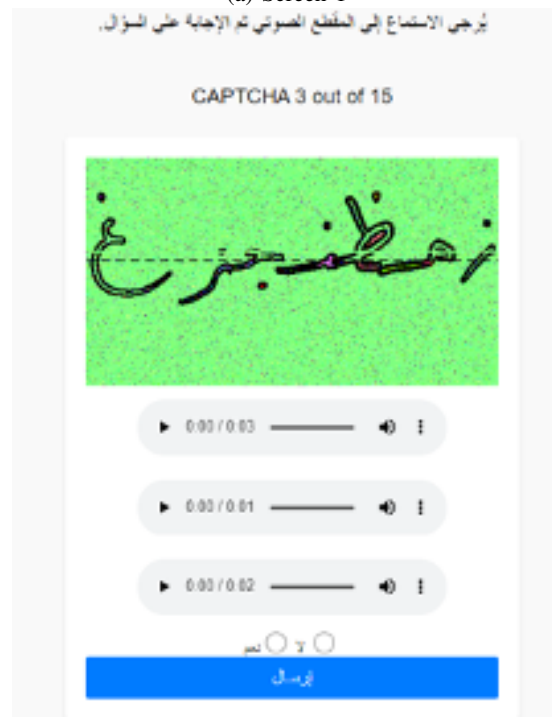
A. Experimental Results

This section presents the results of our experiment in terms of usability, variability, and accessibility, using the previously described dataset.

1) *Evaluation Parameters:* To evaluate the usability and accessibility of the proposed CAPTCHA for both normal and visually impaired individuals, two performance metrics



(a) Screen 1



(b) Screen 2

Fig. 5: Illustrations of some screens from UI used in the experiment.

(accuracy and response time) were assessed and measured. Accuracy referred to the correctness of the participant's answer, capturing whether their response was correct or incorrect. Response time was recorded as the duration from the end of the participant's listening to the audio clip until they clicked the submit button. Additionally, to assess the level of variability in the proposed CAPTCHA, Google OCR, an artificial intelligence system for converting written text into digital text, was utilized. This system enabled text recognition and recorded the accuracy of the OCR's response to the proposed CAPTCHA.

2) *Usability Results:* To achieve our objective, we assessed the ease of use of the proposed system across several demographic factors, including age (under 35 years or over 35 years), gender (male or female), level of education (high school, Bachelor's degree, Master's degree, or PhD), Internet experience (low, medium, high, or expert), and the participant's native country (Saudi Arabia, Egypt, or Tunisia). Detailed explanations of each factor are provided below.

The first factor, age, refers to the measure of the ease of using the proposed CAPTCHA in relation to age and whether an individual's age influences their ability to solve the CAPTCHA tests. To explore this, Table III presents a comparison of the experimental results between two groups: participants under 35 years and those over 35 years, same to the literature classification.

TABLE III: Average of accuracy and response time by age category.

Category	Average Accuracy	Average Response Time (s)
Under 35 years	62.47%	21.9
Over 35 years	64.4%	30.0

Based on the findings presented in Table III, the accuracy rate for participants under 35 years old was 62.47%, compared to 64.4% for those over 35 years. Regarding response time, individuals under 35 took an average of 21.9 seconds, while those over 35 took 30 seconds. These results indicate that individuals over 35 years old achieved a slightly higher accuracy rate; however, they required more time to solve the CAPTCHA challenges compared to individuals under 35. The following factor is gender. This factor evaluates the ease of using the proposed CAPTCHA in relation to gender and investigates whether an individual's gender affects their ability to solve such CAPTCHA tests. To explore this, a comparison of the experimental results between two groups (male and female) is illustrated in Table IV.

TABLE IV: Distributions of accuracy and response time based on genders.

Category	Average Accuracy	Average response Time in seconds
Males	61.2%	26.5
Females	63.87%	23.5

The comparison shown in Table IV indicates that the accuracy rate for males was 61.2%, while for females, it was 63.87%. The response times for males and females were 26.5 seconds and 23.5 seconds, respectively. These results demonstrate that the male group had a lower accuracy rate than the female group despite requiring more time to solve the CAPTCHA challenges.

The third factor is educational background. This factor aims to evaluate the user-friendliness of the proposed CAPTCHA based on educational background and to determine whether one's educational background influences their ability to complete the CAPTCHA challenges. Table V addresses this question by comparing the experimental findings across four groups: high school, bachelor's, master's, and doctorate).

TABLE V: Average of accuracy and response time categorized by education level.

Category	Average Accuracy	Average Response Time (s)
High School	62.7%	26.8
Bachelor	62.1%	23.3
Masters	63.0%	17.3
Doctorate	70.0%	36.0

Table V offers the results obtained from the experiment, where the accuracy rates of solving the CAPTCHA for holders of high school, bachelor degree, master's degree, and a doctoral degree are 62.7%, 62.1%, 63%, and 70%, respectively. Although the average response time taken to solve each CAPTCHA test for high school graduates is 26.8s, it is 23.3s for participants with a bachelor's degree, 17.3s for participants with a master's degree and 36 s for participants with a doctorate degree. In fact, the master's results were the best in terms of response time, as they were the least time-consuming among the group, while the doctorate results were the best in terms of accuracy of solving the CAPTCHA challenges, but they took the longest time in solving these challenges.

Online experience is the fourth factor. Here we seek to evaluate the proposed CAPTCHA's usability in light of one's online experience and assess whether or not one's online experience affects a person's ability to finish the CAPTCHA tasks. In order to find out, Table VI compares the experiment results amongst the four groups: Low usage (less than five years), moderate usage (between six and ten years), high usage (from 11 to 15 years), and expert usage (more than 15 years).

TABLE VI: Accuracy and response time statistics categorized by level of Internet usage.

Category	Average Accuracy	Average Response Time (s)
Low Usage	57.1%	18.9
Moderate Usage	63.6%	23.0
High Usage	63.75%	23.2
Expert Usage	63.3%	28.6

The findings indicate that participants with less Internet experience exhibit a lower accuracy rate compared to those with moderate, high, and expert levels of Internet use, as shown in Table VI. The response times for participants with low, moderate, high, and expert Internet usage were 18.9 seconds, 23 seconds, 23.2 seconds, and 28.6 seconds, respectively. Participants with less Internet experience demonstrated the lowest accuracy in both their answers and response time. In contrast, participants with moderate, high, and expert Internet usage had comparable answer accuracy. Regarding response time, expert users took the longest to provide their answers, with a difference of 5.4 seconds.

The final factor examined is the participants' country of origin. To enhance CAPTCHA usability for humans, we analyzed the results in relation to participants' native countries. Specifically, we investigated participants' accuracy rates and compared them with others from the same country, regardless of their vision level or whether they were sighted or visually impaired. The aim of this analysis is to determine whether a participant's native country influences their ability to solve CAPTCHA challenges. Table VII below presents the results of this factor.

TABLE VII: Accuracy distributions categorized by countries.

Participant Group	Accent	Average Accuracy
Saudi Arabia Participants	Saudi Arabia Accent	67.5%
	Egyptian Accent	62.5%
	Tunisian Accent	60.2%
Egyptian Participants	Saudi Arabia Accent	64.0%
	Egyptian Accent	68.0%
	Tunisian Accent	64.0%
Tunisian Participants	Saudi Arabia Accent	56.6%
	Egypt Accent	58.3%
	Tunisia Accent	65.0%

The results presented in Table VII demonstrate that accuracy rates are higher when the CAPTCHA challenge corresponds to the participant's country. For example, when Saudi Arabian participants encountered a CAPTCHA challenge with a Saudi Arabian accent, their accuracy rate was 67.5%, which was higher than that of participants from other countries completing the challenge with a Saudi Arabian accent. Similarly, Egyptian participants achieved an accuracy rate of 68% when solving a challenge with an Egyptian accent, while Tunisian participants had an accuracy rate of 65% when completing a challenge with a Tunisian accent. Moreover, our findings indicate that participants' responses were relatively similar in terms of accuracy. As shown in the table above, participants' accuracy rates were higher when the CAPTCHA challenge featured the accent of their native country compared to other accents. This suggests that a participant's native country may influence their ability to solve CAPTCHA challenges.

3) *Variability Results:* To evaluate the variability of the proposed CAPTCHA and its robustness against automated attacks and hacking attempts, Google Cloud OCR was used. A total of 1,440 CAPTCHA samples were tested in Google Cloud

OCR, and the results showed an accuracy rate of 0.56%. This percentage is notably low due to the distortions applied to the CAPTCHA images, which enhance security. Therefore, this approach facilitates their use on sensitive websites, improving overall protection.

4) *Accessibility Results:* To assess the accessibility and ease of use of the proposed CAPTCHA, 23 individuals with varying degrees of visual impairment affecting their daily lives participated in the experiment. Table VIII presents a comparison of the results between visually impaired and sighted participants.

TABLE VIII: Results of visually impairment and sighted participants.

Category	Average Accuracy	Average Response Time (s)
Visually impaired	56.2%	25.7
Sighted	65.45%	23.9

The findings reveal that the accuracy rate for visually impaired individuals was 56.2%, while for sighted participants, it was 65.45%. The average response time for visually impaired individuals was 25.7 seconds, compared to 23.9 seconds for sighted participants. The results indicate that individuals with visual impairments exhibit a lower accuracy rate than sighted participants and require more time to complete the CAPTCHA challenges. This may be attributed to their difficulty in perceiving the CAPTCHA images, which influenced the findings of this study.

B. Discussions

In this section, we discuss the results of the experiment, which aimed to provide up-to-date statistics on the usability of the proposed CAPTCHA within the adopted system while introducing potential improvements to enhance its use. The experiment results showed that the average accuracy rate for solving the CAPTCHA challenges across all participant categories was 63%, while the average response time was 24.4 seconds. From our perspective, these results are promising, as no other study closely replicates our proposed system.

To evaluate the robustness of the proposed CAPTCHA, we compare our findings with those of a previous study by (5), which used random Arabic words in CAPTCHAs. As mentioned earlier, we tested our CAPTCHA samples using Google Cloud OCR, and the results revealed an accuracy rate of 0.56% for our system. In contrast, the accuracy rate in (5)'s study was 17%. This comparison indicates that our proposed CAPTCHA is highly secure, given its significantly low accuracy rate. Additionally, the system demonstrates exceptional resistance to cyber-attacks, making it an ideal choice for multi-level authentication systems that require maximum security.

The following Table IX summarizes the key findings of our study compared to relevant previous literature in term of variability.

TABLE IX: Comparing variability results: Current study vs. Alsuhibany & Alquraishi (2021).

Category	Current Study	Alsuhibany & Alquraishi (2021)'s Study
Accuracy	0.56%	17%

Furthermore, to assess the accessibility of the proposed CAPTCHA for individuals with disabilities, we compared its accuracy rate with that of a prior study (3). Our study showed a slightly higher accuracy rate (56.2%) compared to (3)'s findings (48.18%) when testing an audio CAPTCHA for visually impaired individuals. Moreover, the response time in our study was 25.7 seconds, whereas in (3)'s study, it was 138.75 seconds. Although the results obtained are not highly promising, they are considered favorable compared to those of the audio CAPTCHA.

The following Table X presents the key findings of our study compared to relevant previous literature in term of accessibility of visually impaired people.

TABLE X: Comparing accessibility results: Current study vs. Alnfai et al. (2022).

Category	Current Study	Alnfai et al. (2022)'s Study
Accuracy	56.2%	48.18%
Response Time (s)	25.7	138.75

VII. CONCLUSIONS

This study aimed to develop an improved CAPTCHA system that balances usability for sighted individuals, accessibility for the visually impaired, and security through varying levels of image distortion. The proposed system, a novel Arabic text-based CAPTCHA incorporating diverse local dialects, was evaluated for usability, variability, and accessibility. Usability was assessed through solution accuracy and response time across demographic factors such as age, gender, education level, and internet experience.

A key finding was the significant impact of dialect familiarity on accuracy, with participants performing better when presented with CAPTCHA challenges in their native accents. Accessibility analysis revealed that visually impaired participants exhibited lower accuracy and longer response times than sighted users. Additionally, security analysis using Google OCR demonstrated a low accuracy rate, indicating strong resistance to automated attacks. The proposed CAPTCHA outperformed previous systems in robustness, further enhancing security.

Although the study involved 89 participants from three nationalities, the experiments were conducted outside a controlled laboratory setting, where potential distractions may have influenced performance. Despite this limitation, the findings provide valuable insights into the usability, accessibility, and security of CAPTCHA systems, particularly in the context of Arabic dialects.

As for future research, utilizing artificial intelligence technology to present challenges in the user's tone and voice. Moreover, incorporating specific local accents for each region within a country might give new insight.

REFERENCES

- [1] R. Agarwal and A. E. Day, "The impact of the Internet on economic education," *The Journal of Economic Education*, vol. 29, no. 2, pp. 99–110, 1998. [Online]. Available: <https://doi.org/10.1080/00220489809597943>
- [2] M. Alnfai, "Evaluating the accessibility and usability of a universal CAPTCHA based on gestures for smart-phones," *Universal Access in the Information Society*, vol. 20, no. 4, pp. 817–831, 2021. [Online]. Available: <https://doi.org/10.1007/s10209-020-00730-x>
- [3] M. Alnfai, S. Altalhi, and D. Alawfi, "IRemember: Memorable CAPTCHA Method for Sighted and Visually Impaired Users," *International Journal of Advanced Computer Science and Applications*, vol. 13, no. 11, 2022.
- [4] S. A. Alsuhibany, "Evaluating the usability of optimizing text-based CAPTCHA generation," *International Journal of Advanced Computer Science and Applications*, vol. 7, no. 8, 2016.
- [5] S. A. Alsuhibany and M. Alquraishi, "Usability and Security of Arabic Text-based CAPTCHA Using Visual Cryptography," *Computer Systems Science and Engineering*, vol. 40, no. 2, pp. 421–440, 2022. [Online]. Available: <https://doi.org/10.32604/csse.2022.018929>
- [6] S. A. Alsuhibany and M. T. Parvez, "Attack-filtered interactive Arabic CAPTCHAs," *Journal of Information Security and Applications*, vol. 70, p. 103318, 2022. [Online]. Available: <https://doi.org/10.1016/j.jisa.2022.103318>
- [7] D. Brodic and A. Amelio, "Association rule mining for the usability of the CAPTCHA interfaces: a new study of multimedia systems," *Multimedia Systems*, vol. 24, no. 6, pp. 625–644, 2018. [Online]. Available: <https://doi.org/10.1007/s00530-018-0589-6>
- [8] D. Brodic and A. Amelio, "Exploring the usability of the text-based CAPTCHA on tablet computers," *Connection Science*, vol. 31, no. 4, pp. 430–444, 2019. [Online]. Available: <https://doi.org/10.1080/09540091.2019.1609417>
- [9] H. M. M. Eriksen, L. Hoel, C. S. Jetteberg, A. C. Leiknes, and F. E. Sandnes, "Can't You Tell I am a Human? A Comparison of Common Text and Image CAPTCHAs Using a Low-Fidelity Methodology," in *Proc. 10th Int. Conf. Software Development and Technologies for Enhancing Accessibility and Fighting Info-exclusion (DSAI 2022)*, Lisbon, Portugal, Aug.–Sep. 2022, pp. 155–159. [Online]. Available: <https://doi.org/10.1145/3563137.3563179>
- [10] J. Holman, J. Lazar, J. Feng, and J. D'Arcy, "Developing usable CAPTCHAs for blind users," in *Proc. 9th Int. ACM SIGACCESS Conf. on Computers and Accessibility (ASSETS 2007)*, Tempe, AZ,

- USA, Oct. 2007, pp. 245–246. [Online]. Available: <https://doi.org/10.1145/1296843.1296894>
- [11] K. Kaur and S. Behal, “Designing a secure text-based CAPTCHA,” *Procedia Computer Science*, vol. 57, pp. 122–125, 2015. [Online]. Available: <https://doi.org/10.1016/j.procs.2015.07.381>
 - [12] Y.-L. Lee and C.-H. Hsu, “Usability study of text-based CAPTCHAs,” *Displays*, vol. 32, no. 2, pp. 81–86, 2011. [Online]. Available: <https://doi.org/10.1016/j.displa.2010.12.004>
 - [13] R. Lin, S.-Y. Huang, G. Bell, and Y.-K. Lee, “A New CAPTCHA Interface Design for Mobile Devices,” in *Proc. 12th Australasian User Interface Conf. (AUIC 2011)*, Perth, Australia, Jan. 2011, pp. 3–8.
 - [14] M. Meiryani, L. Aliffiyah, L. Lusianah, L. Lidiyawati, T. Lesmana, and K. Kriswanto, “Benefits of Information Technology in Improving the Quality of Health Services,” in *Proc. 6th Int. Conf. on E-Education, E-Business and E-Technology (ICEBT 2022)*, Beijing, China, Jun. 2022, pp. 97–101. [Online]. Available: <https://doi.org/10.1145/3549843.3549858>
 - [15] V. D. Nguyen, Y.-W. Chow, and W. Susilo, “On the security of text-based 3D CAPTCHAs,” *Computers & Security*, vol. 45, pp. 84–99, 2014. [Online]. Available: <https://doi.org/10.1016/j.cose.2014.05.004>
 - [16] G. Reynaga, S. Chiasson, and P. C. van Oorschot, “Exploring the usability of CAPTCHAs on smartphones: Comparisons and recommendations,” in *NDSS Workshop on Usable Security (USEC)*, vol. 112, 2015. [Online]. Available: <https://doi.org/10.14722/usec.2015.23006>
 - [17] X. Song, G. Baltra, and J. S. Heidemann, “Inferring Changes in Daily Human Activity from Internet Response,” in *Proc. ACM Internet Measurement Conf. (IMC 2023)*, Montreal, QC, Canada, Oct. 2023, pp. 627–644. [Online]. Available: <https://doi.org/10.1145/3618257.3624796>
 - [18] C. Tangmanee, “Effects of text rotation, string length, and letter format on text-based CAPTCHA robustness,” *Journal of Applied Security Research*, vol. 11, no. 3, pp. 349–361, 2016. [Online]. Available: <https://doi.org/10.1080/19361610.2016.1178553>
 - [19] X. Wang, Z. Zhao, L. Yi, Z. Ning, L. Guo, F. R. Yu, and S. Guo, “A Survey on Security of UAV Swarm Networks: Attacks and Countermeasures,” *ACM Computing Surveys*, vol. 57, no. 3, pp. 74:1–74:37, 2025. [Online]. Available: <https://doi.org/10.1145/3703625>