A METHODOLOGICAL APPROACH TO CALLIGRAPHIC OBFUSCATION

Suliman a. Alsuhibany

Department of Computer Science, College of Computer, Qassim University, Buraydah, 51452, Saudi Arabia, THURAYA ALWEHAID

ABSTRACT

As automated optical character recognition (OCR) and deep learning-based solvers achieve near-human accuracy in breaking conventional CAPTCHAs, there is a critical need for security mechanisms that exploit the inherent limitations of machine perception. This paper proposes a novel methodological framework for "Calligraphic Obfuscation," a security-by-design approach that leverages the structural complexity and fluid entropy of traditional Arabic calligraphic styles. Unlike standard text-based challenges, our approach introduces a multi-phase generation pipeline that systematically maps linguistic strings into high-complexity visual domains. The methodology integrates a four-tier classification of calligraphic fonts—ranging from high-legibility styles like Naskh to high-entropy scripts such as Shakstah—and augments them with an adversarial layer utilizing Jacobian-based Saliency Map Attacks (JSMA). By formalizing the transition from cloud-centric generation to resource-efficient on-device architectures, this study provides a repeatable blueprint for developing robust, human-interactive proofs. The proposed framework offers a dual-benefit: significantly increasing the computational cost for adversarial machine learning models while maintaining a sustainable cognitive load for human users. This work lays the foundation for a new generation of linguistically-diverse and adversarially-hardened authentication challenges tailored for modern, resource-constrained mobile environments.

KEYWORDS

Calligraphic Obfuscation, CAPTCHA Security, Adversarial Machine Learning, Arabic Script Complexity, Human-Interactive Proofs, JSMA.

1. INTRODUCTION

The internet has become an essential platform for accessing a wide range of online services, such as email, e-commerce, blogs, and user memberships. To utilize these services, users are typically required to register by completing online forms. However, automated bots can exploit these forms by posing as legitimate users to gain unauthorized access or misuse services [1] .

The proliferation of sophisticated automated solvers driven by Large Language Models (LLMs) and advanced Convolutional Neural Networks (CNNs) has rendered traditional text-based CAPTCHAs increasingly vulnerable.

While English-based systems have dominated the field, the unique structural properties of the Arabic script offer a high-entropy alternative that remains underutilized in cybersecurity [2]. In an era marked by frequent cyber threats, accurately distinguishing between human users and automated programs—often referred to as bots—is essential for maintaining digital safety and system integrity [3].

CAPTCHA (Completely Automated Public Turing test to tell Computers and Humans Apart) is one of the most widely adopted mechanisms to prevent malicious bots from exploiting online services. It has become a standard security measure and is now integrated into nearly all modern websites. Various CAPTCHA schemes have been developed, including[4] .

Arabic handwritten CAPTCHAs are especially promising: Arabic script exhibits contextual letter shapes, diverse ligatures, and significant stylistic flexibility influenced by traditional calligraphy. Despite these advantages, limited research has examined how calligraphic styles, text semantics (meaningful vs. meaningless), and adversarial perturbations jointly influence system usability and machine resistance.

Specifically, integrating different Arabic calligraphy styles into CAPTCHA design introduces further complexity, which can enhance resistance to automated attacks, as highlighted in [5] . Figure 1 illustrates examples of handwritten Arabic in various calligraphic styles.

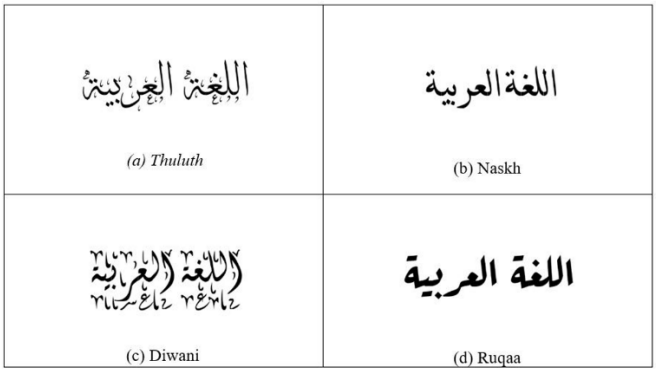

Figure 1: Examples of Arabic calligraphic styles.

This paper introduces "Calligraphic Obfuscation" as a formal methodological framework for generating human-interactive proofs. Unlike standard character distortion, calligraphic obfuscation utilizes the inherent variability, fluid ligatures, and stylistic complexity of traditional scripts to create a natural barrier against machine recognition. The primary challenge in this domain is maintaining a "usability-security equilibrium"—ensuring the script remains legible to human users while becoming computationally expensive for adversarial OCR engines.

The contribution of this work is a systematic, multi-phase methodology that automates the generation of these challenges. By integrating traditional calligraphic styles with modern adversarial perturbations, such as the Jacobian-based Saliency Map Attack (JSMA), we propose a blueprint for robust, linguistically-diverse authentication. This approach is particularly relevant for the evolving landscape of Edge AI, where secure, lightweight, and on-device generation is paramount The remainder of this paper is organized as follows: Section 2 reviews related work. Section 3 details the methodology employed in this study. Section 4 concludes the paper and outlines directions for future research.

2. RELATED WORKS

This section provides a comprehensive review of prior research on Arabic CAPTCHA systems, covering both printed and handwritten formats. It critically evaluates the limitations of existing approaches and highlights advancements in usability and security.

2.2. HANDWRITTEN ARABIC CAPTCHAS

This subsection reviews research focused on handwritten Arabic text-based CAPTCHA systems. Handwritten CAPTCHAs generally offer greater resistance to automated attacks compared to printed CAPTCHAs, as they exploit the natural variation and complexity of human handwriting—an aspect that remains challenging for bots and OCR systems to interpret [6].

In 2016, Aldosari and Al-Daraiseh [7] introduced an innovative technique for generating handwritten CAPTCHAs in four languages, including Arabic. Their method leveraged human handwriting features that are difficult for OCR systems to recognize. However, a notable challenge was the nature of Arabic script, where many characters are inherently connected. This necessitated the design of CAPTCHA text that remained legible for human users despite these script complexities.

Alsuhibany and Parvez [8] were among the first to specifically address Arabic handwritten CAPTCHA generation. Their approach applied a range of transformations—commonly used in handwriting recognition systems—to pre-written Arabic word images to increase resistance to automated recognition. The study demonstrated high usability, with human users achieving strong accuracy rates. In terms of security, a comprehensive word recognition test yielded an accuracy rate of less than 0.5\%, effectively preventing lexicon-based attacks.

Building on this work, in 2020, Parvez and Alsuhibany [6] proposed a CAPTCHA scheme that incorporated complex segmentation in handwritten Arabic. Users were required to identify segmentation points in cursive words. This method not only enhanced CAPTCHA security but also maintained a high level of user-friendliness. Experimental evaluations confirmed the system's strong usability and robustness.

In 2021, Alsuhibany and Alnoshan [9] conducted an empirical study examining the feasibility of interactive handwritten CAPTCHA systems for mobile devices. Their findings revealed that interactive handwritten CAPTCHAs outperformed traditional text-based schemes in terms of both usability and security. The authors also emphasized the importance of improving segmentation algorithms and expanding usability testing with a broader participant base.

In the same year, Alsuhibany et al. [10]. introduced a methodology for generating synthetic Arabic handwritten CAPTCHAs. The study demonstrated that deep learning models trained on synthetic data could effectively recognize distorted Arabic characters, offering a scalable and efficient solution for CAPTCHA development.

Additionally, Alsuhibany and Alquraishi[11] proposed using visual cryptography tools for both printed and handwritten Arabic CAPTCHA schemes. Their experiments confirmed the method's satisfactory performance in terms of usability and security. However, they noted that improvements could be made by generating meaningless handwritten text samples from diverse writers to further enhance system robustness.

In 2022, Alsuhibany and Parvez [12] introduced an interactive CAPTCHA creation method that incorporated attack-filtering mechanisms. This approach allowed for adaptive distortion levels, sequential challenge presentation, and the implementation of additional security layers to bolster CAPTCHA system resilience.

In another study [13], the authors developed a novel CAPTCHA system based on the complexity of handwritten Arabic calligraphy. The results indicated that calligraphy-based CAPTCHAs outperformed other styles in resisting automated bot attacks. Nonetheless, the study

recommended incorporating multiple calligraphic styles to further enhance robustness.

In [14], the authors introduced an innovative CAPTCHA approach that leverages the complexity of handwritten Arabic calligraphy to enhance security against automated bot attacks. The findings demonstrated that this method was highly effective in resisting automated assaults. However, the study emphasized the need to further improve system robustness by incorporating a wider range of Arabic calligraphic styles.

Most recently, in 2025, Alrasheed and Alsuhibany [15] proposed an adversarial CAPTCHA generation framework specifically for handwritten Arabic. The method applied five perturbation techniques, with the Jacobian-based Saliency Map Attack (JSMA) yielding the best performance in terms of both security and usability. The approach significantly improved human readability while reducing the accuracy of machine-based recognition, particularly when meaningful Arabic words were used.

Table 1:The Summary Of Handwritten Arabic Captchas.

Study	Year	Main Contributions
[7]	2016	Presented a technique to generate handwritten multilingual CAPTCHA text using characters.
[8]	2016	Presented a different approach for generating CAPTCHAs using OCR operations to secure the CAPTCHAs.
[6]	2020	Presented a new technique for generating and validating handwritten Arabic CAPTCHAs.
[9]	2021	Presented interactive handwritten and text-based handwritten Arabic CAPTCHA schemes.
[10]	2021	Presented a method for generating synthetic CAPTCHA.
[11]	2022	Applied visual cryptography in both printed and handwritten CAPTCHA schemes.
[12]	2022	Presented a CAPTCHA generation approach that incorporates attack-filtering mechanisms.
[13]	2023	Presented an educational CAPTCHA gamification system combining learning and cybersecurity.
[14]	2024	Presented handwritten CAPTCHA using Arabic calligraphy for enhanced security.
[15]	2025	Presented a CAPTCHA generation approach using adversarial techniques with high usability.
Our study	2025	Provide the comparison between different handwritten Arabic text-based CAPTCHAs in different Arabic calligraphy styles in terms of usability and security aspects.

3. METHODOLOGY

This study adopts a structured, multi-phase approach to generate and evaluate CAPTCHA systems that leverage the inherent complexity of traditional Arabic calligraphic styles. The objective is to achieve a high level of security through machine-resistance while maintaining

human usability. The methodology is divided into four primary stages: automated generation, font classification, adversarial hardening, and evaluation.

3.1. Automated Generation of Calligraphic CAPTCHAs

The generation process utilizes a custom tool built with open-source Python libraries to automate the creation of handwritten Arabic text-based images. The workflow follows a systematic pipeline:

- **Linguistic Sourcing:** CAPTCHA strings are derived from two distinct sources: a predefined list of meaningful Arabic words and synthetically constructed random (meaningless) Arabic strings
- **Calligraphic Rendering:** These strings are rendered using diverse traditional calligraphic styles to increase visual entropy
- **Customization:** The tool allows for dynamic adjustment of dataset size and calligraphy type, facilitating rigorous experimental scalability

3.2. Structural Font Classification Model

To analyze the impact of stylistic variation, calligraphic fonts can be classified into primary and subordinate categories based on visual similarity and human legibility

- **Font Compilation:** The initial set includes prominent styles such as **Naskh, Ruqʿah, Kufic, Thuluth, Diwani, and Farsi**.

- **Style Integration:** To enrich the dataset, fifteen additional subordinate fonts (e.g., **Maghribi, Shakstah, Square Kufi, and Al Battar**) are incorporated.

- **Redundancy Filtering:** Fonts exhibiting high visual similarity or stylistic overlap with primary fonts are excluded to ensure distinctiveness in the model.

3.3. Usability-Based Selection Process

A preliminary usability study will be conducted to refine the font selection based on human perception. This can be achieved by the following:

- **Survey Design:** A user survey involving over 300 participants will be conducted using a Likert scale to measure readability across a number of CAPTCHA samples.

- **Final Selection Criteria:** Fonts can be retained only if they received a "Strongly Agree" readability rating from more than 50% of participants. However, some styles might be excluded from certain tests due to low user approval ratings.

3.4. Evaluation Framework

The final architecture can be evaluated using two primary metrics

- **Usability:** Measured by human accuracy and "time-to-solve" across different calligraphic styles.

- **Security:** Evaluated by submitting CAPTCHA images to the **Google Cloud Vision API** to simulate automated OCR-based attacks

4. CONCLUSIONS

In this paper, we have detailed a comprehensive methodological approach to calligraphic obfuscation, moving beyond simple character randomization toward a structured framework of visual complexity. By categorizing Arabic scripts into distinct stylistic tiers and applying targeted adversarial noise, we have established a repeatable pipeline for the generation of robust CAPTCHA systems.

Our findings suggest that the integration of traditional calligraphy provides a unique layer of "biological defense," where human cognitive flexibility outperforms the rigid pattern-matching of current machine learning models. The methodology presented here is not merely a solution for Arabic-speaking interfaces but serves as a universal model for script-based obfuscation in resource-constrained environments.

Future research will focus on the adaptation of this framework to "Agentic Security," exploring how calligraphic obfuscation can defend against multi-modal LLMs. Furthermore, we aim to investigate the implementation of these pipelines within Edge-computing architectures, ensuring that the next generation of authentication remains both culturally inclusive and technologically resilient.

REFERENCES

[1] B. Khan, K. Alghathbar, M. K. Khan, A. M. AlKelabi, and A. Alajaji, “Cyber security using arabic captcha scheme.,” Int. Arab J. Inf. Technol., vol. 10, no. 1, pp. 76–84, 2013.

[2] Y. Yang, “Understanding of the Cyber Security and the Development of CAPTCHA,” in Journal of Physics: Conference Series, IOP Publishing, 2018, p. 12008.

[3] R. Gafni and I. Nagar, “CAPTCHA–Security affecting user experience,” Issues Informing Sci. Inf. Technol., vol. 13, pp. 63–77, 2016.

[4] A. M. Algwil, “a Survey on Captcha: Origin, Applications and Classification,” J. Basic Sci., vol. 36, no. 1, pp. 1–37, 2023.

[5] M. T. Parvez and S. A. Alsuhibany, “Challenges and opportunities for Arabic CAPTCHAs,” Multimed. Tools Appl., vol. 83, no. 5, pp. 14047–14062, 2024, doi: 10.1007/s11042-023-16166-3.

[6] M. T. Parvez and S. A. Alsuhibany, “Segmentation-validation based handwritten Arabic CAPTCHA generation,” Comput. Secur., vol. 95, p. 101829, 2020.

[7] M. H. Aldosari and A. A. Al-Daraiseh, “Strong multilingual CAPTCHA based on handwritten characters,” 2016 7th Int. Conf. Inf. Commun. Syst. ICICS 2016, pp. 239–245, 2016, doi: 10.1109/IACS.2016.7476118.

[8] S. A. Alsuhibany and M. T. Parvez, “Secure Arabic handwritten CAPTCHA generation using OCR operations,” in 2016 15th International Conference on Frontiers in Handwriting Recognition (ICFHR), IEEE, 2016, pp. 126–131.

[9] S. A. Alsuhibany and A. A. Alnoshan, “Interactive Handwritten and Text-Based Handwritten Arabic CAPTCHA Schemes for Mobile Devices: A Comparative Study,” IEEE Access, vol. 9, pp. 140991–141001, 2021, doi: 10.1109/ACCESS.2021.3119571.

[10] S. A. Alsuhibany, F. N. Almohaimeed, and N. A. Alrobah, “Synthetic Arabic handwritten CAPTCHA,” Int. J. Inf. Comput. Secur., vol. 16, no. 3–4, pp. 385–398, 2021.

[11] S. A. Alsuhibany and M. Alquraishi, “Usability and Security of Arabic Text-based CAPTCHA Using Visual Cryptography.,” Comput. Syst. Sci. Eng., vol. 40, no. 2, 2022.

[12] S. A. Alsuhibany and M. T. Parvez, “Attack-filtered interactive arabic CAPTCHAs,” J. Inf. Secur. Appl., vol. 70, p. 103318, 2022.

[13] M. T. Parvez, A. M. Alsuhibani, and A. H. Alamri, “Educational and Cybersecurity Applications of an Arabic CAPTCHA Gamification System,” Ing. des Syst. d’Information, vol. 28, no. 5, pp. 1275–1285, 2023, doi: 10.18280/isi.280516.

[14] H. Lajmi, F. Idoudi, H. Njah, H. M. Kammoun, and I. Njah, “Strengthening Applications’ Security with Handwritten Arabic Calligraphy Captcha,” in 2024 IEEE 8th Forum on Research and Technologies for Society and Industry Innovation (RTSI), IEEE, 2024, pp. 548–553.

[15] G. Alrasheed and S. A. Alsuhibany, “Enhancing Security of Online Interfaces: Adversarial Handwritten Arabic CAPTCHA Generation,” Appl. Sci., vol. 15, no. 6, 2025, doi: 10.3390/app15062972.

Authors

Suliman A. Alsuhibany, PhD, is a Professor in the Computer Science Department, College of Computer, Qassim University, Saudi Arabia. He received his PhD in Information Security from Newcastle University, UK, and an MSc in Computer Security and Resilience from Newcastle University, UK. He has published in some of the most reputed journals and conferences. His research interests include human aspects of security (e.g. the so-called “usable security”), CAPTCHAs, spam-filter, keystroke dynamics, information security.

Thuraya Alwehaid received the B.Sc. degree in Computer Science from Qassim University, Saudi Arabia. She is currently an M.Sc. student in the Department of Computer Science, College of Computer, Qassim University, Saudi Arabia. Her research interests include information security, especially CAPTCHA.