JOURNAL
India Publication

Journal of Computer Science and Intelligent Systems

Volume 1, Issue 1 | July - September 2026 | www.journalindiapublication.com | 2026 Journal India Publication

Beyond Text CAPTCHAs: Comparing Image and Audio Challenges for Secure, Usable, and Accessible Authentication

SANGEETHA N *

ABSTRACT

Purpose

This Study Aims to Compare Three Widely Used Captcha Techniques, Namely Text-Based, Image-Based, And Audio-Based Captchas, To Evaluate Their Effectiveness in Preventing Automated Bot Attacks While Providing a User-Friendly Experience for Legitimate Users. With The Rapid Development of Artificial Intelligence, Machine Learning, And Optical Character Recognition Technologies, Traditional Captcha Systems Face Significant Challenges in Differentiating Between Human Users and Automated Software Programs. This Research Focuses on Analyzing the Security, Accessibility, Usability, And Resistance Capabilities of Different Captcha Methods Against Modern Cyber Threats.

Design/Methodology/Approach

This Study Follows a Comparative Research Approach Based on a Review of Existing Captcha Technologies and Related Security Studies. Text, Image, And Audio Captcha Methods are Analyzed Using Various Performance Parameters Including Security Level, User Experience, Accessibility, Completion Time, Implementation Complexity, And Resistance to Automated Attacks. The Study Uses a Qualitative Analysis Framework To Identify The Advantages And Limitations Of Each Captcha Technique And To Determine Their Suitability For Different Web Application Environments.

Findings

The Study Finds That Text-Based Captchas Are Easy To Implement And Require Less Computational Resources; However, They Are Becoming Vulnerable Due To Advancements In Ocr And Machine Learning Technologies. Image-Based Captchas Provide Better Security Against Automated Attacks by Requiring Visual Analysis, But They May Create Usability Issues for Certain Users. Audio-Based Captchas Improve Accessibility for Visually Impaired Users but Can Be Challenging Due to Noise Interference And Advanced Speech Recognition Systems. The Results Indicate That No Single Captcha Method Can Provide Complete Security and Usability In All Conditions.

** Corresponding Author: Assistant Professor, School of Computing, E.S Arts and Science College, Villupuram, Tamil Nadu, India, [Email: cs@esasc.in]*

Journal of Computer Science and Intelligent Systems
Volume 1, Issue 1 | July - September 2026 | www.journalindiapublication.com | 2026 Journal India Publication

Originality/Value

This Paper Provides a Comprehensive Comparative Analysis of Text, Image, And Audio Captcha Techniques Using Multiple Evaluation Criteria. It Highlights Current Security Challenges and Provides Insights for Developing More Effective And User-Centric Captcha Systems.

Keywords: Medical image analysis; Deep learning; CNN; Vision transformer; Attention mechanism; Image classification

JEL Codes: C45; I10; O33

---

1. Introduction

The rapid growth of internet-based services has increased the need for secure and reliable user authentication mechanisms. Web applications such as online banking, e-commerce platforms, social networks, and government portals are frequently targeted by automated bots for activities including spam generation, account creation, data scraping, and password attacks. To reduce these threats, captcha (completely automated public turing test to tell computers and humans apart) systems have been widely implemented as an additional security layer.

Captcha technology was introduced to provide a simple method for differentiating human users from automated computer programs. Traditional captcha systems generally require users to identify distorted characters, select specific objects from images, or recognize audio patterns. These tasks are designed based on the assumption that humans can easily perform certain cognitive activities that are difficult for machines.

Text-based captchas were among the first and most commonly used captcha approaches. They require users to enter characters displayed in distorted form. Although they are easy to implement and require low computational resources, advances in optical character recognition (ocr) and machine learning have enabled automated systems to solve many traditional text captchas.

To improve security, image-based captchas were developed. These systems ask users to identify objects, patterns, or visual features within images. Image captchas provide better resistance against simple automated attacks because they require advanced image processing and visual understanding. However, they may create accessibility issues for users with visual impairments.

JOURNAL India Publication
Journal of Computer Science and Intelligent Systems
Volume 1, Issue 1 | July - September 2026 | www.journalindiapublication.com | 2026 Journal India Publication

Audio-based captchas were introduced to support users who experience difficulties with visual captchas. They present spoken characters or words that users must identify. While audio captchas improve accessibility, they face security challenges due to developments in speech recognition technology and audio processing methods.

With the continuous development of artificial intelligence (ai), deep learning, and automated recognition systems, captcha technologies require constant improvement to maintain security. Therefore, a comparative study of different captcha types is important to understand their security performance, user experience, accessibility, and limitations.

This paper presents a comparative analysis of text, image, and audio captcha techniques. The study evaluates these methods based on security level, implementation complexity, user friendliness, and resistance against modern automated attacks. The aim is to provide guidance for developers and security professionals in selecting appropriate captcha solutions for web applications.

2. Literature Review

CAPTCHA systems have been an important area of research in web security since their introduction. Von Ahn et al. (2003) originally introduced CAPTCHA as a method of distinguishing humans from automated computer programs by presenting problems that were difficult for machines but relatively easy for humans. Their work established the fundamental concept of CAPTCHA and introduced text-based tests that used distorted characters to differentiate human users from computer programs.

Text-based CAPTCHA methods have been widely studied because of their simple design and ease of implementation. Yan and Ahmad (2008) demonstrated that weaknesses in text-based CAPTCHA designs could be exploited through automated attacks, highlighting the security limitations of conventional CAPTCHA mechanisms. Researchers have subsequently examined different distortion techniques, including character rotation, noise addition, and background complexity, to improve resistance against automated recognition. However, advancements in Optical Character Recognition (OCR), machine learning, and artificial intelligence have reduced the effectiveness of many traditional text-based CAPTCHA systems. The increasing capability of

JOURNAL India Publication
Journal of Computer Science and Intelligent Systems
Volume 1, Issue 1 | July - September 2026 | www.journalindiapublication.com | 2026 Journal India Publication

automated recognition systems has therefore created a continuing challenge for researchers seeking to maintain an appropriate balance between security and usability.

The usability of CAPTCHA systems has also received considerable research attention. Bursztein et al. (2011) conducted a large-scale evaluation of CAPTCHA solving performance and demonstrated that human-solving accuracy varies considerably depending on the design and complexity of the CAPTCHA challenge. Their findings indicate that increasing CAPTCHA complexity does not necessarily result in better security because overly difficult challenges may reduce successful completion by legitimate users. Similarly, Kluever and Zanibbi (2009) emphasized the importance of balancing security requirements with usability, particularly because complex CAPTCHA challenges can negatively affect user experience and accessibility.

Image-based CAPTCHA systems have received significant attention as an alternative to conventional text-based approaches. These systems require users to identify objects, categorize images, or select specific visual patterns. Sivakorn et al. (2016) demonstrated that deep learning techniques could successfully attack certain semantic image CAPTCHA systems, raising important concerns regarding the assumption that image understanding necessarily provides strong protection against automated programs. Their findings indicate that although image-based CAPTCHAs may provide security advantages against some traditional bot attacks, their effectiveness depends on image quality, task design, semantic complexity, and the ability of artificial intelligence systems to interpret visual information.

The development of CAPTCHA systems has also been closely connected with advances in artificial intelligence and deep learning. Goodfellow et al. (2014) demonstrated the effectiveness of deep convolutional neural networks for multi-digit number recognition from street-view imagery. Although their study was not specifically designed as a CAPTCHA attack, it illustrates the rapid development of automated image-recognition capabilities that can potentially be applied to visual CAPTCHA challenges. These developments have increased concerns regarding the long-term security of image-based CAPTCHA mechanisms.

Audio CAPTCHA systems have been developed primarily to address accessibility requirements, particularly for visually impaired users. These systems generally require users to interpret distorted

Journal of Computer Science and Intelligent Systems
Volume 1, Issue 1 | July - September 2026 | www.journalindiapublication.com | 2026 Journal India Publication

speech, background noise, or randomly generated audio sequences. Such techniques are intended to make automated recognition more difficult while providing an alternative to visually oriented CAPTCHA challenges. However, improvements in automatic speech recognition and artificial intelligence have introduced new security challenges for audio-based CAPTCHA systems. Consequently, the design of audio CAPTCHA must consider not only resistance to automated recognition but also clarity, usability, and accessibility for legitimate users.

Recent developments have further shifted CAPTCHA research toward AI-based, adaptive, and risk-oriented security mechanisms. Modern approaches increasingly combine multiple security features rather than depending exclusively on a single CAPTCHA modality. Adaptive CAPTCHA models, behavior-based analysis, and risk-based authentication approaches have been proposed to provide stronger protection while minimizing inconvenience to legitimate users. Ometov et al. (2018) highlighted the increasing importance of adaptive and multi-factor authentication approaches in addressing evolving cybersecurity threats. Such approaches demonstrate the broader movement toward authentication mechanisms that dynamically assess risk rather than relying solely on conventional challenge-response tests.

Overall, the existing literature demonstrates that each CAPTCHA modality has specific advantages and limitations. Von Ahn et al. (2003) established the foundational concept of CAPTCHA for distinguishing humans from automated programs, while Yan and Ahmad (2008) demonstrated vulnerabilities associated with traditional text-based CAPTCHA mechanisms. Bursztein et al. (2011) highlighted variations in human-solving performance and usability, and Kluever and Zanibbi (2009) emphasized the need to balance security with user accessibility. Sivakorn et al. (2016) demonstrated the vulnerability of semantic image CAPTCHAs to deep learning attacks, while Goodfellow et al. (2014) illustrated the growing capabilities of automated image recognition. Finally, Ometov et al. (2018) highlighted the growing relevance of adaptive and multi-factor approaches to modern authentication.

Therefore, text CAPTCHAs offer simplicity and ease of implementation, image CAPTCHAs provide visual challenge mechanisms, and audio CAPTCHAs enhance accessibility. Nevertheless, no single CAPTCHA approach provides complete protection against all forms of automated

attacks while simultaneously ensuring high usability and accessibility. This gap creates the need for a comprehensive comparative evaluation of text, image, and audio CAPTCHAs, particularly in terms of their security, usability, and accessibility. Such a comparison can provide useful insights into the relative strengths and limitations of different CAPTCHA modalities and help identify suitable approaches for modern web security requirements.

3. Methodology

This study adopts A comparative research methodology to analyze and evaluate three major types of captcha systems: text-based captcha, image-based captcha, and audio-based captcha. The methodology is designed to identify the strengths, weaknesses, and security characteristics of each captcha approach under modern web security conditions.

The research is based on A systematic review of existing studies, technical reports, and security analysis papers related to captcha technologies. Different evaluation parameters are selected to compare the performance of each captcha type. The main parameters include security against bot attacks, resistance to artificial intelligence-based attacks, user accessibility, complexity of implementation, response time, and user experience.

Text-based captcha analysis focuses on character distortion techniques, ocr resistance, and machine learning-based recognition attacks. Image-based captcha analysis examines visual recognition requirements, image classification challenges, and computer vision-based attacks. Audio-based captcha analysis concentrates on speech recognition issues, audio distortion methods, and accessibility features for visually impaired users.

A comparative framework is developed to assess the overall effectiveness of each captcha technique. The results are interpreted based on security performance and practical usability requirements. This approach helps to identify the most suitable captcha solution for different types of online applications.

JOURNAL India Publication
Journal of Computer Science and Intelligent Systems
Volume 1, Issue 1 | July - September 2026 | www.journalindiapublication.com | 2026 Journal India Publication

4. Research Results and Discussion

The comparative analysis of text, image, and audio captcha systems is presented based on security, usability, accessibility, and implementation factors.

the analysis shows that text-based captchas provide basic security and are easy to deploy; however, their security has decreased due to advances in ocr and deep learning technologies. Automated systems can now identify many distorted characters with high accuracy.

Image-based captchas show better security performance because they require visual understanding and object identification. Current ai systems can still attack image captchas, but the complexity of these attacks is higher compared with text captchas.

Audio-based captchas provide important accessibility benefits for users with visual disabilities. However, improvements in automatic speech recognition technology have created new security challenges. Audio quality, background noise, and language differences also affect user performance.

Based on the comparison, image-based captchas provide relatively strong security, while audio captchas offer better accessibility. Text captchas remain popular because of their simple implementation but require additional security enhancements.

The results of this study show that captcha systems must maintain a balance between security and user experience. A captcha that provides high security but creates difficulties for users may reduce the effectiveness of a web application. Similarly, a simple captcha with low security may not provide sufficient protection against automated attacks.

Text captchas are still widely used because they require minimal resources and are easy to integrate into applications. However, the development of ocr and ai technologies has reduced their reliability. Image captchas provide stronger protection but require more computational resources and may affect users with visual limitations.

Audio captchas improve inclusiveness by supporting users who cannot access visual captchas. However, they need improvements to handle advanced speech recognition attacks and environmental noise conditions.

JOURNAL India Publication
Journal of Computer Science and Intelligent Systems
Volume 1, Issue 1 | July - September 2026 | www.journalindiapublication.com | 2026 Journal India Publication

Table 1. Comparative Evaluation of Text, Image, and Audio CAPTCHA Systems

Evaluation Parameters	Text CAPTCHA	Image CAPTCHA	Audio CAPTCHA
Security Level	Low	High	Moderate
Resistance Against AI Attacks	Low to Moderate due to Optical Character Recognition (OCR) advancements	High due to image processing complexity and deep learning challenges	Moderate due to advancements in speech recognition technologies
Implementation Complexity	Low	Medium	High
User Friendliness	Easy for most users but may create difficulties due to distorted text recognition	Requires visual understanding and image interpretation skills	Suitable for visually impaired users but may be challenging in noisy environments
Accessibility	Limited	Limited	High
Processing Requirement	Low	Medium to High	High
Common Attack Methods	OCR, pattern recognition, automated text solving	Image classification, deep learning-based attacks	Speech recognition, audio analysis

5. Conclusions, Proposals and Recommendations

The study suggests that future captcha systems should use adaptive security approaches. Combining multiple captcha types with ai-based risk analysis can provide better protection while maintaining user convenience.

JOURNAL
India Publication
Journal of Computer Science and Intelligent Systems
Volume 1, Issue 1 | July - September 2026 | www.journalindiapublication.com | 2026 Journal India Publication

This paper presented a comparative study of text, image, and audio captcha techniques used for web security applications. Each captcha method has unique advantages and limitations. Text captchas provide simple implementation but are increasingly vulnerable to ocr and machine learning attacks. Image captchas provide higher security through visual analysis but may create accessibility challenges. Audio captchas improve accessibility but face threats from advanced speech recognition systems.

The study concludes that no single captcha method can completely solve the challenge of bot detection. Future captcha systems should focus on hybrid approaches that combine different captcha types, user behavior analysis, artificial intelligence, and adaptive security techniques.

Future research can explore ai-based captcha generation, privacy-preserving authentication methods, and intelligent systems that can automatically adjust security levels based on user behavior and threat conditions.

6. Theoretical Implications

This study contributes to the theoretical understanding of CAPTCHA systems by providing a comparative analysis of text-based, image-based, and audio-based CAPTCHA techniques in the context of web security. The findings highlight the relationship between technological advancement and the effectiveness of human–computer differentiation mechanisms. The study extends existing cybersecurity theories by showing that improvements in Artificial Intelligence (AI), Machine Learning (ML), Optical Character Recognition (OCR), and speech recognition technologies continuously influence CAPTCHA reliability. It also supports the theoretical view that adaptive and multi-layered security approaches are required to address evolving automated threats.

7. Practical Implications

The practical implications of this research help web developers, cybersecurity professionals, and organizations select appropriate CAPTCHA techniques based on security requirements, accessibility needs, and implementation resources. The findings suggest that text CAPTCHA solutions may be suitable for low-risk applications due to their simplicity, while image and audio

JOURNAL India Publication
Journal of Computer Science and Intelligent Systems
Volume 1, Issue 1 | July - September 2026 | www.journalindiapublication.com | 2026 Journal India Publication

CAPTCHA methods may provide stronger protection in higher-risk environments. The study recommends adopting hybrid CAPTCHA systems that combine multiple verification methods, AI-based threat detection, and user behavior analysis to improve bot prevention while maintaining usability.

8. Social Implications

The social implications of this study emphasize the importance of creating secure and inclusive online environments. While CAPTCHA systems protect digital platforms from malicious bots, poorly designed CAPTCHAs can create barriers for users, especially individuals with visual, hearing, or cognitive disabilities. The research highlights the need for accessible authentication solutions that balance cybersecurity protection with equal access for all users. Future CAPTCHA developments should consider privacy protection, accessibility standards, and user-centered security designs to promote safer and more inclusive digital services.

9. Industry Implications

This study has significant implications for the cybersecurity and technology industries by providing insights into the strengths and limitations of different CAPTCHA techniques used for online security. The findings can help software companies, web service providers, and cybersecurity organizations develop more effective bot detection and authentication solutions. Industries can use the results to design adaptive CAPTCHA systems that combine text, image, and audio verification with Artificial Intelligence, machine learning, and user behavior analysis to improve security against evolving automated attacks.

The research also highlights the importance of balancing security, usability, and accessibility in digital services. Organizations should invest in advanced CAPTCHA technologies that reduce user inconvenience while maintaining strong protection against malicious bots. Future industry developments may focus on AI-driven CAPTCHA generation, risk-based authentication, and privacy-preserving security frameworks to enhance trust and reliability in online platforms.

JOURNAL India Publication
Journal of Computer Science and Intelligent Systems
Volume 1, Issue 1 | July - September 2026 | www.journalindiapublication.com | 2026 Journal India Publication

10. Limitations & Scope for Future Research

This study has certain limitations. First, it relies on comparative analysis of existing CAPTCHA techniques rather than large-scale user experiments or real-world bot attack simulations. Second, CAPTCHA effectiveness may vary according to website design, user behavior, attack conditions, and technological developments. Third, rapid advances in Artificial Intelligence, Machine Learning, Optical Character Recognition, and speech recognition may alter the security performance of existing CAPTCHA systems. The study also focuses on text, image, and audio CAPTCHAs without examining biometric, behavioral, or passwordless authentication. Future research should investigate adaptive and AI-driven CAPTCHA systems, hybrid CAPTCHA models, privacy-preserving authentication, and large-scale real-world evaluations to assess security, usability, accessibility, and resistance to evolving cyber threats.

Conflicts of Interest Declaration

- The author declares that there are no conflicts of interest regarding the publication of this research.

Informed Consent Declaration

- This study does not involve human participants, personal data collection, or direct interaction with individuals. Therefore, informed consent was not required. The research is based on a comparative analysis of existing CAPTCHA techniques, including text-based, image-based, and audio-based CAPTCHA methods, using publicly available information and previously published research findings. All sources of information have been appropriately acknowledged and referenced.

References

- Bursztein, E., Bethard, S., Fabry, C., Mitchell, J. C., & Jurafsky, D. (2011). How good are humans at solving CAPTCHAs? A large-scale evaluation. 2010 IEEE Symposium on Security and Privacy, 399–413. https://doi.org/10.1109/SP.2010.31

- Goodfellow, I. J., Bulatov, Y., Ibarz, J., Arnoud, S., & Shet, V. (2014). Multi-digit number recognition from street view imagery using deep convolutional neural networks. arXiv. https://doi.org/10.48550/arXiv.1312.6082
- Kluever, K. A., & Zanibbi, R. (2009). Balancing usability and security in CAPTCHA design. Proceedings of the 5th Symposium on Usable Privacy and Security.
- Ometov, A., Bezzateev, S., Mäkitalo, N., Andreev, S., Mikkonen, T., & Koucheryavy, Y. (2018). Multi-factor authentication: A survey. Cryptography, 2(1), 1–22.
- Sivakorn, S., Polakis, I., & Keromytis, A. D. (2016). I am robot: (Deep) learning to break semantic image CAPTCHAs. 2016 IEEE European Symposium on Security and Privacy (EuroS&P), 388–403. https://doi.org/10.1109/EuroSP.2016.40
- von Ahn, L., Blum, M., Hopper, N. J., & Langford, J. (2003). CAPTCHA: Using hard AI problems for security. Advances in Cryptology—EUROCRYPT 2003, 294–311. https://doi.org/10.1007/3-540-39200-9_18
- Yan, J., & Ahmad, A. S. E. (2008). A low-cost attack on a Microsoft CAPTCHA. Proceedings of the 15th ACM Conference on Computer and Communications Security, 543–554. https://doi.org/10.1145/1455770.1455840