iHuman CAPTCHA: An Alternative CAPTCHA for Visually Impaired Based on Face Liveness Detection System

Nanaware Yash V., Kurhe Aishwarya K., Gade Shubhangi P., Shirode Shubham N.,

Dr. M. B. Gawali

Department of Information Technology, Sanjivani College of Engineering, Kopargaon, Savitribai Phule Pune University, Pune, MH, India.

---***---

Abstract - *With the tremendous progress in computer technology, bot attacks have become the main problem. CAPTCHA (Completely Automated Public Turing test to tell Computers and Humans Apart) has been introduced to tackle this problem. In various systems, the CAPTCHA is presented in a form of characters, numbers, symbols, images, mathematical problems, etc. that makes it challenging for the visually impaired to authenticate. In this work, we propose the Face Liveness Detection CAPTCHA (iHuman CAPTCHA) as an alternative solution, which will directly detect the face of a human. The face is the common property of the human. Most of the visually impaired use the talkback feature, which is in-build in many systems. The iHuman CAPTCHA system is trained using CNN Algorithm. To authenticate into the system, the visually impaired will click the 'Verify using the camera' button.*

Key Words: Visually Impaired, CAPTCHA, Face Liveness Detection, iHuman CAPTCHA, CNN

1. INTRODUCTION

Since 1980, people have seen tremendous development and advancement of Computer Technology. It has also observed an increase in bot attacks. To prevent this, CAPTCHA (Completely Automated Public Turing test to tell Computers and Humans Apart) differentiate whether the user is human or bot.

Today, CAPTCHA appears in numerous patterns like Text-based and Numbers- based CAPTCHA, Audio CAPTCHA, Mathematical problem CAPTCHA, reCAPTCHA (Image-based CAPTCHA, No CAPTCHA, Audio CAPTCHA), and much more [1]. The CAPTCHAs, which are dependent on reading text or any other visual perception tasks, blind or visually impaired users get barred from accessing the protected resource. Most visually impaired people use inbuilt accessibility tools like talkback and screen readers. CAPTCHAs are unreadable by machines and uninterpretable for screen readers.

According to the Census 2011 (2016 Updated) "Disabled Population of India" report, out of the 1.21 billion population, 26.8 million persons are disabled [5]. Table -1. shows the Disabled Population of India in 2016.

Table -1: Disability Population in India as per Census 2011 (2016 updated)

User Types	Identifier	Population
Normal	N	1,18,40,40,033
Blind	B	50,33,431
Deaf	D	50,72,914
Dumb	DU	19,98,692
Deaf-Blind	DB	5,00,000
Multiple Disability	M	16,16,698
Other Disability	O	1,25,93,209

Total Population: 1,21,08,54,977

To defeat this problem, audio CAPTCHA gets used as an alternative. Random numbers, letters, and words are present in the audio CAPTCHA, which is likely to be automated [2]. According to the study, the audio CAPTCHA was not audible clearly due to random numbers, letters, and words. It becomes complex to distinguish letters one from another like b and p, and j and g [3].

In this modern era, face recognition and biometric authentication systems deliver higher usability and are robust. Here we propose the Face Liveness Detection CAPTCHA (iHuman CAPTCHA) for the visually impaired. It is an alternative. This system will detect the face liveness followed by instructions like a smile, blink eyes, see left, see right, and verify the user.

In the rest of the paper, we will discuss a literature review on various CAPTCHAs in Section 2, the proposed model in Section 3, security in section 4, usability in section 5, and conclusion in Section 6.

2. LITERATURE SURVEY

In 2007, Shirali-Shahreza Mohammad et al. [6] propose a CAPTCHA for blind people. In this CAPTCHA, an easy mathematical problem gets created according to predefined patterns then it is converted into speech with the help of the Text-To-Speech (TTS) system. Then the sound of the converted problem is played for the user. Then he/she needs to enter the correct answer to the question. For example- "There are '2' cats, '4' apples, and '3' dogs on a table. How many pets are there on the table?". Then the user needs to enter the answer - 5 (2 cats + 3 dogs).

In 2009, Shirali-Shahreza Sajid et al. [7] propose a new CAPTCHA called Spoken CAPTCHA for blind users, in which a random word gets generated and converted to speech. Then the audio is played to the user then he/she needs to recite that word. Then the response of the user is checked w.r.t generated word.

In 2012, Lazar Jonathan et al. [8] propose a CAPTCHA known as SoundsRight CAPTCHA. They also evaluate CAPTCHA with 20 blind users. SoundsRight CAPTCHA gets presented as a real-time audio-based challenge in which the user needs to recognize a particular sound like the sound of a piano and bell. Every time it occurs in a sequence of 10 sounds playing through the computer's audio system.

In 2014, Yamaguchi Michitomo et al. [9] propose a new type of CAPTCHA, which generates verbal tests for visually disabled people. The tests are composed of various phrases. The tests are of two types 1. distinguishing the strange meaning of the phrase from others 2. identifying the common topic of phrases. For example- Test of detecting common topic: Q. Which is the common topic to the following sentences? 1. Twitter is a social networking service. 2. Vector processors get used for parallel computing. A) Computer, B) Sports, C) Culture. Then the user needs to select the answer.

In 2020, Alnfiai Mrim [10] proposes a CAPTCHA called a HearAct. It is also a real-time audio-based CAPTCHA for visual impairments. The user listens to the audio clip of something (sound-maker), and he/she needs to recognize what the sound-maker is. Then, HearAct identifies a word and asks the user to analyze it and decide whether it has the stated letter or not. If the word has the letter, the user needs to tap. If not, they will swipe. For example- Does the sound-maker have the letter 'a'? Then the user will answer by tap otherwise swipe.

3. PROPOSED MODEL

This section specifies the broad idea of iHuman CAPTCHA. Here the system architecture, process, model, and algorithm are described. The idea of iHuman CAPTCHA gets originated by considering the following factors

A. Serviceable for the visually impaired- It has been seen that new websites, apps, and software come into the market but with some drawbacks. The problem arises especially for the visually impaired. Due to visual-based CAPTCHA Systems used by websites, apps, and software, it becomes hard for them to authorize it.

B. Systems used by the visually impaired- According to a study, many of the visually impaired use the inbuild tools to access, like screen-readers and talkback. The screen-reader and talkback tool are unable to read CAPTCHA. iHuman CAPTCHA give the user instruction in both visual and audio forms. The instructions are not the actual answer that an attacker's script can identify and answer. Instead, the user needs to perform the instructions by actions. Such as smile, see left, blink, see right, etc.

C. Learning Disabilities- Some CAPTCHA systems come with mathematical problems (also in audio format), making it hard to understand and solve the problem.

D. Language Barrier- India is a country of 22 schedule languages. Many users are from rural areas, and English is their 2nd or 3rd language. That makes it hard to understand English Phrases. But, many of them know English commands.

E. Threats and Attacks- With an increase in technology, hackers get used to new methods of attacks because of old or less secure CAPTCHA. Here, face recognition and biometric authentication delivers higher usability and are robust. As per iHuman CAPTCHA is based on Face liveness detection, the attacks are limited to some extent.
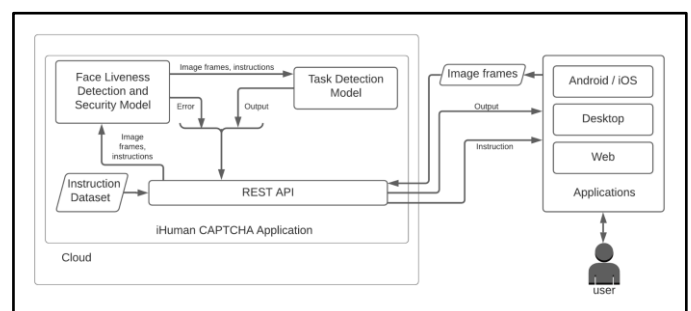
3.1 System Architecture

Fig -1: System Architecture of iHuman CAPTCHA

In Fig -1. the system architecture of iHuman CAPTCHA is described. The iHuman CAPTCHA Application has a REST API that connects to several applications on the user's end. Three random instructions get presented on the user screen. The instructions are displayed on the screen and also in audio form. The user needs to perform

the tasks instructed in provided instructions in order. The instructions set includes the instructions like - smile, see left, see right, see up, see down, etc.

For example - The instructions are given to the user in order as A. See left B. Smile C. Close your Eyes
Then the user needs to perform the task in the above order.

The API collects the user data in image frames and passes it to the Face Liveness Detection and Security Model along with the instruction. The Face liveness Detection and Security Model checks that the user does not violate any security rules. If the security model identifies any security threat, an error message gets returned to the user. If not, it gets forwarded to the Task Detection Model.

The Task Detection Model checks whether the user has performed the task according to instruction. If yes, the user will get authorized. Else it will throw an error. The user will get authorized after the completion of all tasks.

3.2 Models

In iHuman CAPTCHA, face liveness detection and security model works on the same aspect to prevent the attack. These models get trained with the help of the CNN Algorithm

3.2.1 Face Liveness Detection Model

Face Liveness Detection has various approaches like Frequency and Texture based analysis, Variable Focusing based analysis, Movement of the eyes based analysis, Optical Flow-based analysis, Blinking based analysis, Component Dependent Descriptor based analysis, 3D Face Shape, Binary Classification based analysis, Scenic Clues based analysis, Lip Movement-based analysis, Context-based analysis, and Combination of Standard Techniques based analysis [14].

In iHuman CAPTCHA, Movement of the eyes-based analysis [15], Blinking-based analysis [16], and Lip movement-based analysis [17] is implemented.

3.2.2 Task Detection Model

In iHuman CAPTCHA, there is more than '10' instruction to perform the task. Each task gets trained separately using CNN Algorithm. The Task Detection Model compares the user task with instructions. Based on the comparison, this model decides whether the user can be authorized.

4. SECURITY

The attacks of face detection-based systems are mainly two types- Presentation Attacks and Compromising Attacks. The presentation attacks consist of Hard Copy, Screen, and 3D Mask, while the compromising attacks consist of a 2D image, Video, and 3D face. The compromising attacks are probably more scalable than presentation attacks. It is because compromising attacks can occur entirely in the digital while presentation attacks are limited to physical attacks [4].

Here the Face Liveness Detection and Task Detection also play a significant security role. The Hard Copy, 3D Mask, 2D image, the 3D face can easily get encountered by Face Liveness Detection. The remaining presentation attack like Screen and compromising attack like Video will also get encountered by Task Detection. Here, iHuman CAPTCHA is providing three instructions in order. Currently, the system consists of 10 instructions. For a total no. of 10 instructions, if 3 random instructions are selected, then 120 combinations of instructions are created. In future work, no. of combinations will increase with the no. of instructions.

5. USABILITY

By considering, the User types, Identifiers and the Population from Table 1., the usability of CAPTCHA is predicted in Table 2.

Table -2: Usability of CAPTCHAs in terms of India's Population in 2016

CAPTCHAs	User Types	Population	Population (%)
CAPTCHA for blind people [6]	N, B, DU, M	1,19,31,88,854	98.54
Spoken CAPTCHA [7]	N, B, O	1,20,16,66,673	99.24
SoundsRight CAPTCHA [8]	N, B, DU, O	1,20,36,65,365	99.40
Accessible CAPTCHA [9]	N, B, DU	1,19,10,72,156	98.36
HearAct CAPTCHA [10]	N, B, DU	1,19,10,72,156	98.36
iHuman CAPTCHA	N, B, DU, M, O	1,20,57,82,063	99.58

Total Population: 1,21,08,54,977

Table -2. predicts that iHuman CAPTCHA is more accessible with 99.58% usability than other CAPTCHAs

followed by SoundsRight CAPTCHA with 99.40% usability and Spoken CAPTCHA with 99.24% usability.
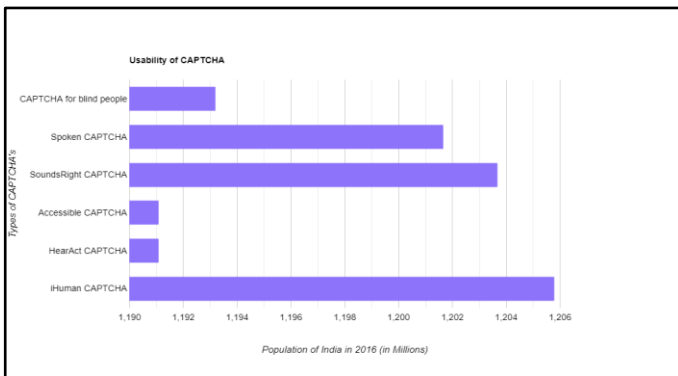

Chart -1: Usability of CAPTCHAs in India.

Chart -1 represents the graph of usability of CAPTCHAs in India in terms of the 2016 population.

6. CONCLUSIONS

In this paper, we propose a new CAPTCHA for the visually impaired. It is known as iHuman CAPTCHA. The iHuman CAPTCHA provides three random instructions to the user. The instructions get display on the screen and also in audio form. Based on provided instructions, the user needs to perform the tasks. Then the system will detect the liveness of the user. Simultaneously it will detect threats. The user tasks get compared with the instructions. If all three instructions get matched, then the user will get authorized. This system prevents presentation attacks and compromising attacks by default using Face Liveness Detection and Task Detection. After comparing the CAPTCHA usabilities, It has been predicted that 99.58% of India's population can easily access it. We are working on new features and new tasks.

REFERENCES

[1] Singh, Ved Prakash, and Preet Pal. "Survey of different types of CAPTCHA." *International Journal of Computer Science and Information Technologies* 5.2 (2014): 2242-2245.

[2] Fanelle, Valerie, et al. "Blind and Human: Exploring More Usable Audio {CAPTCHA} Designs." *Sixteenth Symposium on Usable Privacy and Security ({SOUPS} 2020)*. 2020.

[3] Aiswarya, K., and K. S. Kuppusamy. "A study of audio captcha and their limitations." *International Journal of Science and Research* 4.5895 (2015): 918-923.

[4] Uzun, Erkam, et al. "rtCaptcha: A Real-Time CAPTCHA Based Liveness Detection System." NDSS. 2018.

[5] Editor. "Disabled Population in India as per Census 2011 (2016 Updated)." Enabled.In, 29 Aug. 2017, enabled.in/wp/disabled-population-in-india-as-per-census-2011-2016-updated.

[6] Shirali-Shahreza, Mohammad, and Sajad Shirali-Shahreza. "CAPTCHA for blind people." 2007 IEEE International Symposium on Signal Processing and Information Technology. IEEE, 2007.

[7] Shirali-Shahreza, Sajad, et al. "Spoken captcha: A captcha system for blind users." 2009 ISECS International Colloquium on Computing, Communication, Control, and Management. Vol. 1. IEEE, 2009.

[8] Lazar, Jonathan, et al. "The SoundsRight CAPTCHA: an improved approach to audio human interaction proofs for blind users." Proceedings of the SIGCHI conference on human factors in computing systems. 2012.

[9] Yamaguchi, Michitomo, et al. "An Accessible CAPTCHA System for People with Visual Disability–Generation of Human/Computer Distinguish Test with Documents on the Net." International Conference on Universal Access in Human-Computer Interaction. Springer, Cham, 2014.

[10] Alnfiai, Mrim. "A Novel Design of Audio CAPTCHA for Visually Impaired Users." *International Journal of Communication Networks and Information Security* 12.2 (2020): 168-179.

[11] Chakraborty, Saptarshi, and Dhrubajyoti Das. "An overview of face liveness detection." arXiv preprint arXiv:1405.2227 (2014).

[12] Jee, Hyung-Keun, Sung-Uk Jung, and Jang-Hee Yoo. "Liveness detection for embedded face recognition system." *International Journal of Computer and Information Engineering* 2.6 (2008): 2142-2145.

[13] Sun, Lin, et al. "Blinking-based live face detection using conditional random fields." International Conference on Biometrics. Springer, Berlin, Heidelberg, 2007.

[14] Kollreider, Klaus, et al. "Real-time face detection and motion analysis with application in "liveness" assessment." IEEE Transactions on Information Forensics and Security 2.3 (2007): 548-558.