Captcha Generator for Visually Impaired

(Colour-Blind)

Zara Iqbal

*Computer Science Department, Vellore Institute of Technology  
Kelambakkam, Chennai, Tamil Nadu, India*

[zara.iqbal2019@vitstudent.ac.in](mailto:zara.iqbal2019@vitstudent.ac.in)

Shreyansh Satapathy

*Computer Science Department, Vellore Institute of Technology  
Kelambakkam, Chennai, Tamil Nadu, India*

shreyansh.satapathy2019@vitstudent.ac.in

Sai Kumaresh

*Computer Science Department, Vellore Institute of Technology  
Kelambakkam, Chennai, Tamil Nadu, India*

[om.saikumaresh2017@vitstudent.ac.in](mailto:om.saikumaresh2017@vitstudent.ac.in)

Dr. Asha S

*Computer Science Department, Vellore Institute of Technology  
Kelambakkam, Chennai, Tamil Nadu, India*

*asha.s@vit.ac.in*

*Abstract***— Today's Internet users come from a variety of ages and groups. Humans enabled in other ways also use the Internet. Some websites are designed for people with special needs. Many websites offer human user services, but unfortunately some computer programs are designed to exploit this service. As a result, a specific system called captcha was introduced (Completely Automated Public Turing test to Tell Computers and Humans Apart). For security and privacy purposes websites widely use CAPTCHAs. However, for individuals with visual impairments, traditional text-based CAPTCHAs are not very suitable. With many captchas already in place to help the visually impaired no such captcha exists that specially aims to aid the colour-blinded in using captchas. In this research paper we have tried to understand and showcase how the colour-blinded can be aided in using captchas. To achieve the same, ML has been used.**

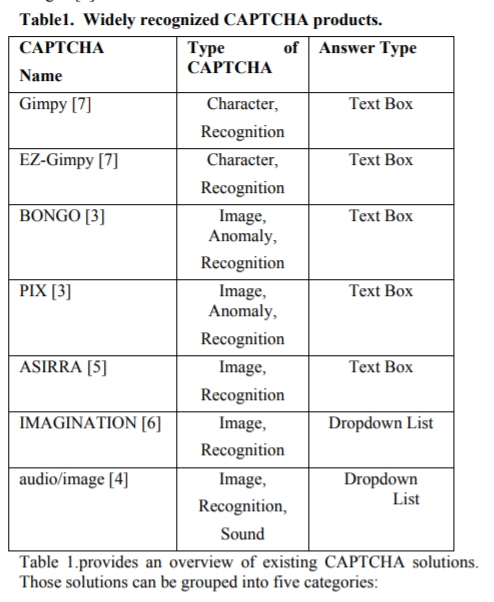
***Index Terms***— **Captcha, Visually-impaired, Colour-blindness, Colour-blinded, Captcha for colour-blind, Captcha for visually-impaired.**

1. Introduction

There are many daily activities on the Internet today, such as education, shopping, and mailing. With the rapid growth of the Internet and easy access, a lot of personal information can be found on the World Wide Web. To prevent a computer from accessing these data many technologies exist. The purpose of these technologies is to distinguish between human programs and computer programs. Several methods have been proposed to overcome this problem. This method is automatically implemented by a computer program because of the considerable time and cost of checking a large number of registration forms by manpower. This simple method is called captcha.

CAPTCHA is a fully automated public Turing test that distinguishes between computers and people, a system that automatically separates people from machines. These systems are the foundation of AI systems. They are similar to Turing, but different because the judge is a computer. The purpose of these systems is to ask questions that people can answer easily while current computer programmes cannot. Other applications are also provided by CAPTCHA systems such as spam prevention. In today's scenario, the internet is used by more than just specially enabled groups of people, but also by people of every age and groups. There are numerous internet sites for children and children that use the internet for activities such as entertainment, learning, etc.

The elderly communicates with their children and relatives on the Internet and chat with them. The disabled people also use the Internet, in addition to ordinary people. Only visual captures create a significant barrier for certain users, such as visually impaired, colour-blind, almost visible users. This means that more accessible captchas are required.

 Currently, captchas are primarily visual, and therefore rely on high human perceptions that visually impaired people cannot solve. Audio captcha, which rely on human hearing, is being introduced instead as an alternative to vision, but it is much more difficult for web users to solve. Because of the intrinsic difficulty of interpreting a sound file, the most common audio captchas achieve less than 50 percent success rates.

Besides audio based and other forms of vibration-based captchas which have proven to be very ineffective, not much research and forms of captchas exist for the colour-blind. In this paper we have tried to focus our research in efforts to make captchas more accessible specifically to the colour blind.

Various research studies show that about one in every twelve males and one in every two-hundred female suffers from colour-blindness. Even if colour blind men can see stuff as clearly as any other person, red, green or blue light cannot be fully discerned. Colour blindness is different; the most common are deuteranopia and protanopia. People with blindness in red-green (deuteranopia) face difficulties in differentiating red from green. Likewise, all red colours look dull for people with protanopia colour-blindness.

To tackle this and make captchas easier to be read by the colour-blind, our UI provides captcha according to the type of colour-blindness of the user.

1. Related literature

Article [1] explains and explains the concept and history of captchas and the classification of modern captcha methods based on texts, photos, videos, and puzzles. Various methodologies have been introduced and described in each classification. The authors have further discussed the pros and cons of each category. In this study [2], a novel CAPTCHA method for the visually impaired has been presented by the authors. A small math problem is generated using a text-to-speech system based on established patterns. The following uses text-to-speech to play the user's sound, prompting the user to enter an answer to a question. As an alternative to typical auditory CAPTCHA, the article [3] proposes an accessible type of CAPTCHA for visually impaired users that uses gestures to answer CAPTCHA tasks. The user listens to the sound of something (the "sound-maker") and then has to figure out what it is. Hear Act then recognises a word and asks the user to examine it to see if it contains the specified letter. The user must tap if the word contains the letter, and swipe if it does not.

In article [4] they designed, implemented and evaluated four novel audio CAPTCHAs. The Math prototype required users to calculate a running total; the Character prototype required users to give the number of times a character appeared in an alphanumeric series; the Pauses prototype required users to transcribe the alphanumeric characters they hear; and the Categories prototype required users to get the number of sounds in a series that fit into a specific category. This article [5] discusses the pedestrian wireless navigation system. It integrates a range of technology, such as GIS and GPS, wearable computers, wireless networks and voice recognition. It increases the visually impaired context information and calculates almost optimised routes for the user, temporal restrictions (e.g., traffic congestion) and dynamic obstacles. Based on dynamic and static data, the system guides the visually-impaired user to find their way.

In this paper [6], the authors propose to engage in visual analysis of the environment of the visually-impaired. The analysis is based on the extraction of GIST and K-Nearest Neighbour methods. The results of the analysis help visually-impaired people understand places and events. The authors conducted a test on a pseudo-visually-impaired person walking six scenes on a university campus. The scene recognition rate is about 85 per cent. This paper [7] presents a portable system to support people with visual impairments indoors and outdoors. It uses various sensors, using GPS and compass, to detect and guide obstacles in movement. A multi-core android smartphone is the main component of the system. Other sensory modules detect obstacles and provide the main part with relevant information. For remote monitoring the system can optionally communicate remotely.

This paper [8] examines how crowd recognition is one of the main problems when the blind people navigate the fields they recognise or don't know. The person with visual impairment is equipped with sensors to detect and warn the human being accordingly. In this paper, all R&D to serve visually impaired people is thoroughly examined. A wide range of strategies are analysed and compared, based on their accuracy, compactness, the performance of the calculations and their cost.

In the paper [9], CAPTCHA can usually be divided into three groups, namely, Non-Visual, Visual non-OCR and OCR. OCRs are used to read texts automatically, but they are difficult to read texts printed in poor quality and they can only identify high-quality text types using common standard formats. They are also used to read texts automatically. However, if the image of a word is changed so that only humans and not any OCR system may recognize it, flaws in the OCR system are could be exploited.

This research article [10] shows that in an extensive study of more than 150 participants it is clearly difficult to complete conventional audio captchas over visual captchas for the visually impaired. The authors instead developed and evaluated a new non-visual CAPTCHA resolution interface that could be added to an existing audio captcha. The most useful audio interfaces are often not direct translations of visual interfaces. Optimized interface uses voice CAPTCHA to improve success rate of the visually-impaired by 59 per cent.

1. Proposed methodology

A Web server can store both public and private resources, such as web pages, data saved in a database or files, or any other type of service.

Human users on the client side will be able to use it. The client computer sends a resource to the server when the user asks for it.

If the resource is not protected, it is allowed this right. If a resource is CAPTCHA protected, access is restricted. It is only given to it after passing the CAPTCHA test. as seen in the diagram

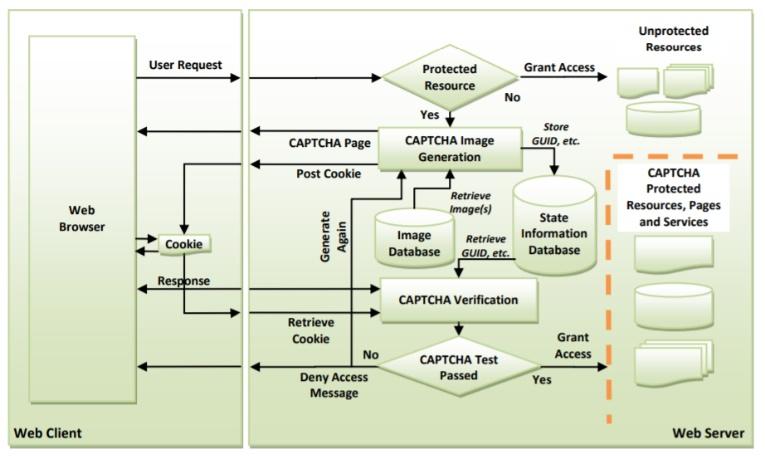
In the dataset that we are following, Inception v3 is a commonly used image recognition model[13] that has been proven to achieve better than 78.1 percent accuracy [11]. The model represents the result of numerous concepts explored over time by a number of scholars.

“

The model contains multiple elements such as convolutions, pooling and fully linked layers.[13]

Batchnorm is applied to activation inputs and is utilised extensively throughout the model. Softmax program is used to calculate loss.

The CNN module extracts the different features present in a test image.

1) The pre-trained CNN system can easily identify the different objects or actions present in the test image.

2) All the features and objects are passed on to the LSTM module.

3) The Flicker8K dataset has 1200 images with all the objects clearly defined and named. The

LSTM matches the objects in the test image with those already present in the dataset. The

similar object is chosen by comparing the probability, and the one with the highest matching

statistic is chosen.

Figure 1: Flowchart of a captcha system

4) The entire process is repeated several times with the same image to increase the accuracy and

train the dataset in recognizing the objects or actions with different filters.

The Flickr30k dataset has established a gold standard for picture description using sentences.

For continuing advancement in automated picture description and grounded language interpretation, annotations are required.

Annotations allow us to set a new standard for picture localization. We have tried to provide a baseline for this task in the form of image-text embedding. Colour classifiers etc. Our basic prototype is obviously not great in terms of image-sentence retrieval, we do find ourselves in a better situation in terms of accuracy for a smaller database.

• It is modest in size, which is why we picked the Flicker8K dataset. As a result, “the model may be readily trained on low-end laptops and PCs.”

• The data has been correctly labelled. There are five captions for each image.

• The dataset is freely accessible.

“Within this collection, there are primarily two sorts of data:

* images
* captions (Text)



**Figure 2: Normal Image**



**Figure 3: Red-Blind/Protanopia**

On our prototype website (Figure 5), the user gets an option to input whether they have any sort of colour-blindness, and if the answer is yes, then from a drop-down menu they can specify which form of colour-blindness affects them. Once chosen, our system will show the pictures only pertaining to that so that the user can answer without any difficulty. The entered sentence should be be an 80% match with the system description or else it shows as captcha failed (Figure 4).

1. Results

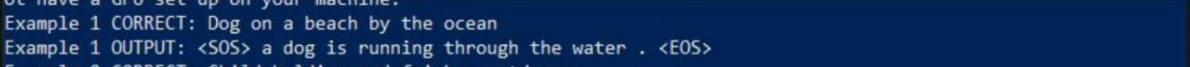
In this research paper, we have implemented a solution that aids the colour-blinded in using CAPTCHAs. Using images that convey a proper image to the color blind and Machine learning we were able to figure out a way to help the color-blinded people complete the CAPTCHA process with ease and at the same time maintain the security of the website.

Figure 4: System Verification of Captcha

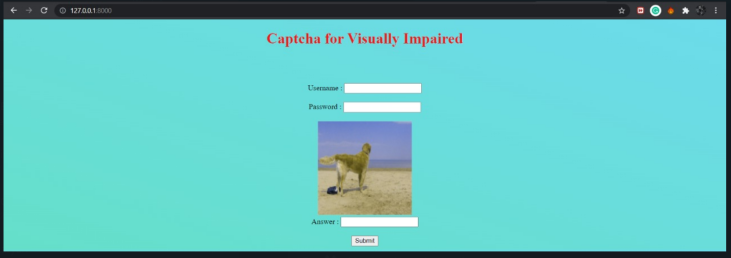


Figure 5: Screen

1. discussion

In the study [12], the authors suggested “two new Captchas”, “Farett-Gender” and “Farett-Gender&Age face recognition”. Analytic authors have shown that Captchas proposed for brute-force and IA-based attacks are secure unless the rates of age and gender are above 80 percent, and users are not willing to classify more than 25 images. One of the conclusions drawn from this study was that more images certainly provide the captcha recognition system against brutal and AI attacks with more security. The major limitation in this paper is that to secure the captcha for the user, the user has to upload too many images and this system is hardly usable for any sort of visually impaired human.

Our paper, on the other hand, focuses on the different types of colour blindness and how to help color-blinded people use captcha more easily than ever before. We have shown limitations of other researches through our related literature, in addition to the colorblindness research and its appropriate captcha aid.

To improve our project, we could:

• Use a bigger dataset.

• Change the model architecture, that is, include an attention module.

• Use a BLEU Score to evaluate and measure the performance of our model.

1. conclusion

The existing CAPTCHAs are not supportive for the visually impaired. The prototype we have implemented in this paper helps color blinded people in solving CAPTCHAs. This prototype makes the process easier than making it more complex by including additional elements like a TTS system or an additional device making this a better alternative for the existing systems. With a more robust data set the Machine Learning algorithm can function with more accuracy making it a better choice.

1. references

[1] Walid Khalifa Abdullah Hasan, A SURVEY OF CURRENT RESEARCH ON CAPTCHA, International Journal of Computer Science & Engineering Survey (IJCSES) Vol.7, No.3, June 2016

[2] Mohammad Shirali-Shahrezal, Sajad Shirali-Shahreza, CAPTCHA for Blind People, 2007 IEEE International Symposium on Signal Processing and Information Technology

[3] Mrim Alnfiai, A Novel Design of Audio CAPTCHA for Visually Impaired Users, International Journal of Communication Networks and Information Security (IJCNIS), Vol. 12, No. 2, August 2020

[4] Valerie Fanelle, Sepideh Karimi, Aditi Shah, Bharath Subramanian, and

Sauvik Das, Georgia Institute of Technology, Blind and Human: Exploring More Usable Audio CAPTCHA Designs, August 10–11, 2020

# [5] [A. Helal](https://ieeexplore.ieee.org/author/37298540900), [S.E. Moore](https://ieeexplore.ieee.org/author/37732848900), [B. Ramachandran](https://ieeexplore.ieee.org/author/37739290000), Drishti: an integrated navigation system for visually impaired and disabled, [Proceedings Fifth International Symposium on Wearable Computers](https://ieeexplore.ieee.org/xpl/conhome/7613/proceeding), IEEE Xplore: 07 August 2002, ISBN:0-7695-1318-2

[6] Quoc-Hung Nguyen, Thanh-Hai Tran, "Scene description for visually impaired in outdoor environment", Advanced Technologies for Communications (ATC) 2013 International Conference on, pp. 398-403, 2013.

[7] Laviniu Ţepelea, Virgil Tiponuţ, Péter Szolgay, Alexandru Gacsádi, "Multicore portable system for assisting visually impaired people", Cellular Nanoscale Networks and their Applications (CNNA) 2014 14th International Workshop on, pp. 1-2, 2014

[8] Orangzaib Rajput, Samiya Qureshi, Ali Raza Solangi, Zuneera Aziz, Faisal Karim Shaikh, "Applicable operational mechanisms to assist visually impaired people-A WSN perspective", Information and Communication Technology for the Muslim World (ICT4M) 2013 5th International Conference on, pp. 1-6, 2013.

[9] S. Shirali-Shahreza, H. Abolhassani, H. Sameti and M. Hassan, "Spoken captcha: A captcha system for blind users," 2009 ISECS International Colloquium on Computing, Communication, Control, and Management, Sanya, 2009, pp. 221-224, doi: 10.1109/CCCM.2009.5270464.

[10] J. P. Bigham and A. C. Cavender, "Evaluating Existing Audio CAPTCHAs and an Interface Optimized for Non-visual Use", *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 1829-1838, 2009.

[11] [Christian Szegedy](https://arxiv.org/search/cs?searchtype=author&query=Szegedy%2C+C), [Vincent Vanhoucke](https://arxiv.org/search/cs?searchtype=author&query=Vanhoucke%2C+V), [Sergey Ioffe](https://arxiv.org/search/cs?searchtype=author&query=Ioffe%2C+S), [Jonathon Shlens](https://arxiv.org/search/cs?searchtype=author&query=Shlens%2C+J), [Zbigniew Wojna](https://arxiv.org/search/cs?searchtype=author&query=Wojna%2C+Z), "Rethinking the Inception Architecture for Computer Vision", arXiv:1512.00567, Submitted on 2 Dec 2015 ([v1](https://arxiv.org/abs/1512.00567v1)), last revised 11 Dec 2015 (this version, v3).

[12] Bursztein, E. Moscicki, A. Fabry, C. Bethard, S. Mitchell, J. C. and Jurafsky, D. (2014). “Easy does it: More usable captchas”, Proceedings of the 32nd annual ACM conference on Human factors incomputing systems - CHI '14.

[13] Shahzia Siddiqua, Naveena Chikkaguddaiah, Sunilkumar S. Manvi, Manjunath Aradhya. "AksharaNet: A GPU Accelerated Modified Depth-Wise Separable Convolution for Kannada Text Classification", Revue d'Intelligence Artificielle, 2021