SURVEY ARTICLE

A Systematic Survey on CAPTCHA Recognition: Types, Creation and Breaking Techniques

Mohinder Kumar[1] · M. K. Jindal[1] · Munish Kumar[2]

Abstract

CAPTCHA stands for Completely Automated Public Turing Test to Tell Computers and Human Apart. CAPTCHA is used for internet security. A few CAPTCHA schemes are available today like, text-based, audio-based, video/animation-based, puzzle based etc. In this paper, all these types are collaborating at single place to analyze. The main aim of this article is to present a literature to identify and recognize CAPTCHA, its types, the creation and breaking techniques. It is a systematic and complete analysis of all available CAPTCHA types. In this paper, 16 text-based CAPTCHA's generation methods are discussed with usability and security ranges from 3 to 100 and 65 to 100%, respectively. The security and usability measures are not calculated/sustained using some known English schemes. Out of 16 reviewed CAPTCHAs, 12 are based on English language, 1 on Arabic language, 1 on Chinese language, 1 on Devanagari language and 1 on Gurumukhi script. The designs are made segment proof with overlapping random shapes, overlapping characters, clasping, different colors and different shades. For making recognition proof many techniques are used like image masking, local and global warping; broken characters, random rotation, arcs, jaws, etc. Approximately 50 schemes, especially based on the English language, are successfully broken with a success rate that ranges from 2 to 100%. The techniques that are used to break these schemes include shape context matching, distortion estimation, Log Gabor 2D filter, horizontal and vertical projection (for a segment the letters) are used. For recognition CNN, KNN, DNN and MCDNN are used. Almost 15 images-based CAPTCHAs are discussed that are designed with usability and security range 90–100 and 17–100%, respectively. Out of these 5 schemes are successfully broken with a success rate ranging between 7 and 100%. The K-NN and SVM are mostly used algorithms to recognize the images. Audio based CAPTCHAs (5 designs) are discussed with usability and security range from 68.5 to 100 and 100%, respectively. The broken rate of these audio schemes is also 45–75%. These schemes are broken with SVM and K-NN algorithms. The paper also discusses 4 popular video-based designs that provide usability and security that ranges from 75 to 100 and 98 to 100, respectively. These schemes are also compromised with broken rate 16–10% using SIFT, NN and simple OCR techniques. The paper can be a benchmark to precede any specific research to dive into any one of these types.

1 Introduction

In the present era, the internet is a major interaction for every person. It does not matter what is the age, profession, gender, and sector. The availability of the rich variety of mobile devices and cheaper high-speed data plans increased the interest of the users for the Internet usage. The variety of the content on the web is also multitalented that attracts everyone. Most of the data is also available at free of cost on the Internet that adds the number of users in compounded way. As the Internet is becoming the most popular platform to provide data services the number of websites and blogs are also increasing. Today the web sites are designed for financial services, public services, entertainment services, grocery products, healthcare services, transportation services,

✉ Munish Kumar
   munishcse@gmail.com

   Mohinder Kumar
   kumarmohinderr@yahoo.co.in

   M. K. Jindal
   manishphd@rediffmail.com

1   Department of Computer Science and Applications, Panjab University Regional Centre, Muktsar, PB, India

2   Department of Computational Sciences, Maharaja Ranjit Singh Punjab Technical University, Bathinda, PB, India

hotel bookings etc. But the knowledge of the user has been also not always good for these web sites. Internet Security is always the main challenge for the web developers from the beginning. The increasing number of users also demands the high-end processing units at the deployed web site, but these high-end units are useless if the high-end machine attacks these servers. Ahn et al. [3] highlighted an instance that happened at CMU when the students developed a program for submitting ballots in online polls in their schools' favor. So, a program can be trained in such a way that can enter a website and makes the server so overloaded with requests that in turn results in crashing of the server. This program is well known as bot program. A few methods are developed by the researchers to stop such kind of attacks. But most of the methods are very expensive and demands a lot of efforts by the experts as well as like One Time Password etc.

A very effective and cheapest method is CAPTCHA. This Reverse Turing Test is also known as Completely Automated Public Turing Test to Tell Computers and Humans Apart or CAPTCHA in short. In such methods a Reverse Turing Test is given to the attacker and depending on the challenge passed, it is decided that the attacker is a human or a computer. Naor [48] explained the term Reverse Turing Test introduced in 1950, when a Turing Test was introduced to check with a human that the other side is introduced by a computer or human. In this paper we call it CAPTCHA from now to make it easy to write. The CAPTCHA is designed in such a way that is easily understandable, but very difficult for a computer program. Coates et al. [24] text-based CAPTCHA shown in Fig. 1a. This includes simple text to be recognized easily by a human being but not by a program due to some noise and distortion. In the last 20 years a lot of CAPTCHAs designs are proposed that includes a large variety of forms. In this paper, we will discuss all these types of designs. We will discuss the design features and the breaking methods of these CAPTCHA's.

2 Security and Usability Metrics of CAPTCHA

A CAPTCHA must hold the sweet spot between solvability by computers and humans as depicted in the Fig. 2. A delicate balance must be maintained by a valid CAPTCHA challenge. It must not be so easy to break by a computer program and at the same time is must not be hard to break
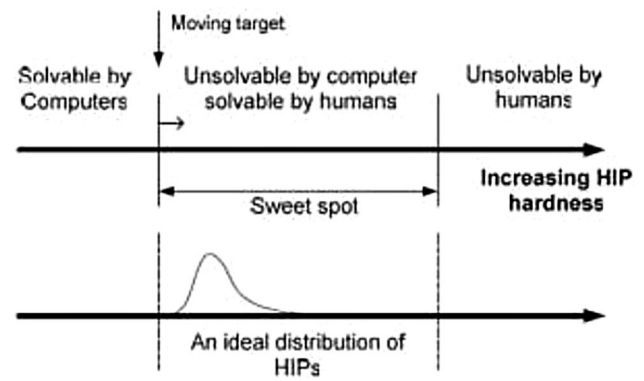

Fig. 2 Desired properties of a CAPTCHA

by a human. Although it is not an easy task to achieve this sweet spot as the history of CAPTCHA tells.

A CAPTCHA is assumed to be secure if its success rate is less than 0.0001%. It means that out of 10,000 challenges only not more than one should be broken by a computer program. The usability of a CAPTCHA should be more than 80% for users. It means that out of 100 challenges 80 times, a CAPTCHA should be easily identified by the users within a minimum time e.g. 3–5 s.

3 Motivation

CAPTCHAs schemes are of many types like text-based, image-based, audio-based, animation- based, puzzle-based, video-based, and even now invisible schemes are also introduced by Google. These all schemes are unique in one way. The classical text-based schemes use simple text as a challenge. But these letters are made segment proof and recognition proof. To make this CAPTCHA noise is added, distortion and warping are applied to the text. Sometimes letters are broken and even hidden. In the image-based CAPTCHA schemes, the challenges include images of things, persons, animals, etc. The user is asked to identify an image among the given images. In the animation-based CAPTCHA, the text is moving, and a user is asked to identify the text. In the video-based schemes the user is asked to identify the type of video based on the contents of the video. In puzzle-based schemes, the user is given a number puzzle or sometimes image-based puzzle. In the latest invisible CAPTCHA, the user is provided with a checkbox and the user just need to click in the checkbox to pass the challenge. In the previous literature, all these CAPTCHAs are never analyzed at a single platform. Only the English language text-based schemes are discussed in more details as compared to other schemes. It motivates us to present all the available CAPTCHAs in one platform and perform analysis on all these techniques. Many of the review articles pick one type of scheme most

Fig. 1 **a** and **b** Text based CAPTCHA

of the times. That does not provide the perfect review for the upcoming research. In this paper, an effort is made to include all the different types of CAPTCHAs under one tree. An effort has been made by analyzing all the text-based CAPTCHAs of different languages. All the image-based, animation-based, video-based, audio-based CAPTCHAs are discussed in detail from their generation to end, like what are the techniques used to create these schemes and what are the techniques to break these challenges? Also, the guidelines are provided to make these successfully broken schemes, stronger. In the following sections, the authors have tried to find answers to some of the very important research questions as shown in the Table 1.

4 Source the Information

The various types of available CAPTCHA schemes must be understood for this analysis. Such effort requires a big collection of the existing research. Large information about the existing CAPTCHAs must be required to cover such a broad context. The collection of this information is done from conference/symposium proceedings, journals, research articles, books, etc. The sources that are being referred for this study are as follows:

- IEEE eXplore (ieeexplore.ieee.org)
- Springer (www.springerlink.com)
- ACM Digital Library (www.acm.org/dl)
- ResearchGate Publications
- Other Reputed Research Journals
- Books and Technical Reports
- Conference / symposium Proceedings
- Ph.D. Thesis
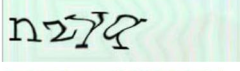- Relevant Web Articles

Table 1 Research questions

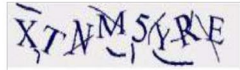|	
1	What are the existing types of CAPTCHAs available?
2	What is the national and international status of Text Based CAPTCHAs?
3	What the techniques are for generating CAPTCHA?
4	What are the tools and techniques to break CAPTCHA?
5	What tools and techniques to test security of CAPTCHA?
6	What is the usability status of current CAPTCHA?
7	What is the security status of current CAPTCHA?
8	What are the guidelines for making secure CAPTCHA?

5 Types of CAPTCHAs

It is important to know about the various types of different CAPTCHA schemes. The variety of CAPTCHA is so large that it is very beneficial to categorize these techniques. So, for the sake of simplicity, all the different types of CAPTCHA schemes are categorized under the following heads:

- Text Based
  These are the very basic type of CAPTCHA. These are the most widely used CAPTCHA. The user is presented a string of characters that are easy to recognize but not easy for a computer. The strings can vary from a clear text to very noisy text. Text-based CAPTCHA are available in various flavors like simple text CAPTCHA that is very easy to recognize. This CAPTCHA is without any kind of warping or noise. Distorted CAPTCHA in which the text is not straight forward but distorted to make is difficult to recognize by a computer program. Characters are also touched so to make the CAPTCHA segment proof. Li et al. [40] presented a little noisy with little distortion CAPTCHAs includes a few arc or random dots etc. (That may be touching or not touching the characters of the CAPTCHA to make segment proof). A few patters of noise are also presented to make it non recognizable. Very noisy, overlapped and much distorted CAPTCHAs are hard text CAPTCHA. The presence of noise is much more. Characters are rotated, distorted with greater level. Imsamai and Phimoltares [35] have presented 3D-CAPTCHA. The amount of overlapping is also more. These types of CAPTCHA are hard to recognize for both computer as well as humans. Some of the popular text based CAPTCHAs are presented in the Table 2:
- Image Based
  The next approach to design a CAPTCHA is image-based CAPTCHA. Here, the user is presented with images or images with text and asked to pick a correct picture or word that belongs to the picture/s. Computer programs are not as good as human in identifying graphics. Some of the popular image based CAPTCHAs are presented in the Table 3.
- Audio Based
  The next category of CAPTCHA is audio-based. Here, the user is presented with an audio and user must type or press the key for that word or phrase. This CAPTCHA can also be very useful for visually impaired users. Two techniques are mostly used for audio-based CAPTCHA, namely, text-to-speech, and Spoken CAPTCHA. Text-To-Speech Conversion Method- Here a voice of a recorded word/number/phrase is played to the user with some

Table 2 Examples of text based CAPTCHAs

Scheme	Database	Example
EZ gimpy	561 words	
Gimpy	411 Words	
Gimpy r	130321 (19^4) Words	
Mailblocks	10^7 Numbers	
Register	26^5 Words	
Yahoo Version 2	62^6 Words	
Ticketmaster	62^5 Words	
Google	52^6 Words	
MSN	36^8	
Holiday inn priority CAPTCHA	62^5	

noise/distortion added. Li [40] explained the user is asked to type the played word/number/phrase. Google uses this CAPTCHA for visually impaired users for its online services. Spoken CAPTCHA-This is another approach developed by Shirali-Shahreza et al. [65] for users who are unable to see or cannot respond to images. In this technique, a selected word from the system's dictionary is sent to the Text-To-Speech module where an audio file is created for that word. This word is played before the user and he is asked to say the word. Now, the user's reply is again recorded and sent as a speech file to two modules on the server, one is speech recognition and human vs. to recognize the word spoken by the user and the other is computer analysis, which analyses the reply of the user to decide whether this is a computer program or the human voice. The conclusion is derived by analyzing these two modules if the voice is from the user and provides the correct answer,the user is given access to the website. Some of the popular audio based CAPTCHAs are presented in the Table 4.

- Video-Based
  Video-based CAPTCHA is a new kind of CAPTCHA. But, it requires a much more bandwidth on the internet and requires much more attention of the user. E.g. a video-based CAPTCHA, presented by Kluever and Zanibbi [37], asks the user to describe any three words about the shown video. Some video CAPTCHA shows moving text in the running video. The user is asked to type the cheaters of color, etc. Some of the popular image based CAPTCHAs are presented in the Table 5.

- Puzzle-Based
  Puzzle-based CAPTCHA is kind of a question that can vary from very simple to very complex. A lot of versions are available in this form. It can include some games like Tic-Tac-Toe, or some mathematical operations like, addition, subtractions, or any simple reasoning problem like, series, or some visual puzzles etc. Some of the popular puzzle based CAPTCHAs are presented in the Table 6.

- Mouse-Based
  Recently, Google has developed a new CAPTCHA that does not require any text, image, audio or video data for pass the test. It requires only one click of the mouse to tell the computer, whether there is a human on program on the other side. It is known as reCAPTCHA shown in Fig. 3. The user is presented with a checkbox and he has been asked to just click on it. The checkbox is not just a checkbox, but a virtual checkbox. Google inserts an invisible text area in the form and populates it with a

Table 2 (continued)
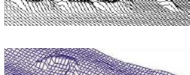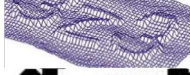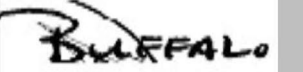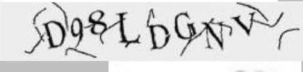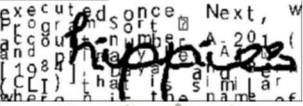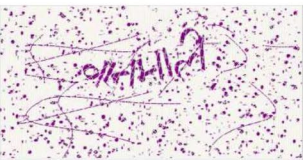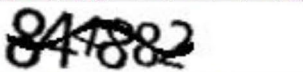
Phpcaptcha.org	Infinite	
Cryptograph	26^4 words	
FreeCap	Infinte	
LinkedIn	Infinite	
Megaupload	36^4 words	
BotDetect	36^5 words	
Yahoo	36^5 to 36^8 words	
Authorize	36^5 words	
Baidu	36^4 words	
Blizzard	26^6 words	
Captcha.net	36^6 words	
CNN	36^5 words	
Digg	62^5 words	
eBay	10^5 words	
Megaupload	26^4 words	
NIH	36^5 words	
Reddit	26^6 words	
Skyrock	36^6 words	
Slashdot	46^6 words	
Wikipedia	infinte	
Hollow styles		
Yahoo	26^4–52^6 words	
Tencent		
Sina		
CmPay		
Baidu		
Microsoft CAPTCHA	6 upper alpha	
Chinese CAPTCHA	1–5 characters	
Super CAPTCHA	Alpha and fdigits	
3D CAPTCHA		
TeaBag CAPTCHA 1.2		
Baffle Text	26^8 Words	

Table 2 (continued)
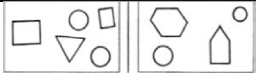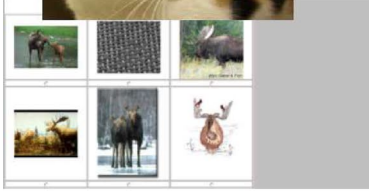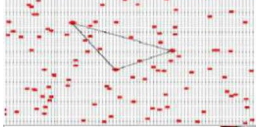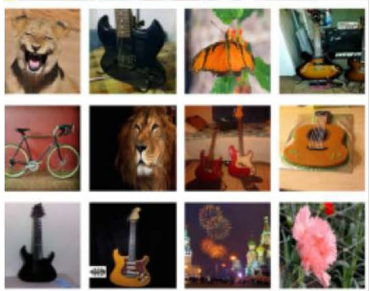
Handwritten CAPTCHA	$20,000^{10}$ words	
	4000 words	
MSN CAPTCHA	36^8	
Clickable CAPTCHA	Infinite	
Mixed text synthetic handwritten CAPTCHA	Infinite	
3D CAPTCHA	45^6	
STE3D-CAP	Infinite	
Synthetic handwritten CAPTCHA	4000 handwritten city names and infintie words from 20,000 handwritten characters	
Sigma-Lognormal CAPTCHA	15000 charaters images	
DevaCAPTCHA	44 Devanagari vowels and consonants	
Google CAPTCHA	4-15 letters (a-z,A-Z,0-9)	
Gurmukhi CAPTCHA	5 letters	
Arabic handwritten CAPTCHA	KHATT databse	

value that is unique. This value is working as an indicator that the user is a bot or not. The value can be true or false for the test. An online article explained that Google not only depends on the checkbox, but it also relies on the pattern of movements of the mouse that differentiate humans and programs. It also uses user time on page algorithms; bots IP addresses database, HTTP referrer, number of requests, etc. It also uses algorithms for Google Analytic's (to prevent bots from increasing page view) and Google AdSense (to prevent fraud clicks on ads) to detect bots. If the reCAPTCHA is not sure about the user, then it displays the old-style image based Pix CAPTCHA. There is a very important thing to note that Google has not disclosed the algorithm that how it works. So, according to Ahn et al. [3] it is not a CAPTCHA because it is not PUBLIC that the definition of CAPTCHA says.

- Invisible CAPTCHA

Table 3 Examples of image based CAPTCHAs

Scheme	Database	Example
BONGO	Finite 2D shapes	
ASIRRA CAPTCHA	3,000,000 images (10,000 images addition daily)	
Anomalies Image CAPTCHA	10,116 images	
PIX CAPTCHA	70 image classes	
Implicit CAPTCHA	finitie images	
Drawing CAPTCHA	Infinite images	
Google Image CAPTCHA	Infinte images	
Facebook CAPTCHA	Infinte images	

Google improved the reCAPTCHA and developed a totally new concept, known as No CAPTCHA or Invisible reCAPTCHA. This CAPTCHA does not require any kind of user interaction that means even the clicking of check box is not required. Ana article at developers.google.com explained that, this CAPTCHA is invoked directly when the user clicks on an existing button on the web site or it can be invoked via a JavaScript API call.

Table 4 Examples of audio based CAPTCHAs

Scheme	Database	Example
CAPTCHA for blind users	4 image clasess	
HIPUU	Fintine images	
Google reCaptcha	Random 0–9 digits and letters	

Table 5 Examples of video/ animation based CAPTCHAs
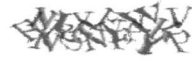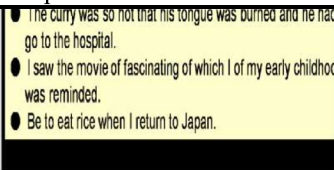
Scheme	Database	Example
3D animation CAPTCHA	Infinite	
AniCAP	Infinte	
Motion CAPTCHA	4500 videos	
New video CAPTCHA	Infinite	
NuCaptcha	Combination of 3 reduced alphabets	

Table 6 Examples of puzzle based CAPTCHAs

Scheme	Database	Example
SS CAPTCHA	Infinite	
Qestion based CAPTCHA	NM	
Four panel CAPTCHA	NM	
DCG CAPTCHA	4 games	
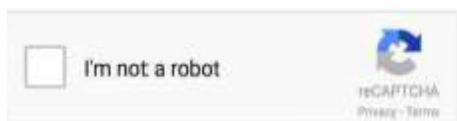
Fig. 3 Google mouse CAPTCHA

6 Applications of CAPTCHA

Some of the applications are highlighted in the Fig. 3.

- *Restriction on Getting e-mail Address* Ahn et al. [3] explained that web scrapers (Program that extracts data from web sites) can be restricted from getting user's e-mail addresses by presenting a checkpoint to solve a CAPTCHA before displaying e-mail address information.

- *Restriction on Fake E-mail Registration* In the modern digital life most of the companies provide free e-mail services. These are the target of bot attacks. A lot of companies use this technique to get free e-mail accounts from which they can send junk mail. The best and easy method to avoid these bot attacks is the use of CAPTCHA. Pope and Kaur [51] identified that most of the e-mail providers have adopted this solution.

Prevent Dictionary Attacks

Restricting on Getting E-mail Address

Stop Worms and Spams

A Barrier for Search Engine Bots

Applications

Restrict Generation of Fake E-mail Address

Rejecting Shpping agents

Restrict Program for Online Game

Restrict Automatic Online Polls

Better Password Management

Defend Phishing Attacks
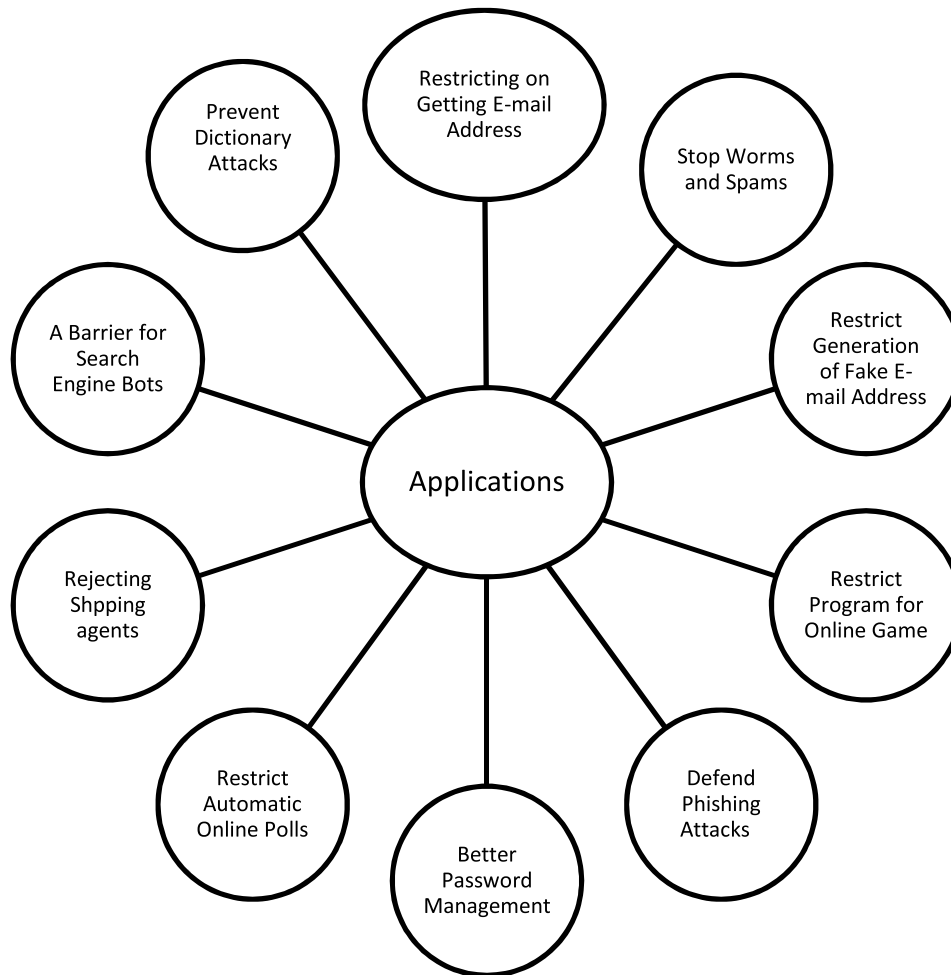
- *Defending Phishing Attacks* Phishing is used by non-ethical hackers (crackers) to crack into online banking accounts, social site accounts and e-mail accounts by presenting a similar fake web page to befool the account holders. For example a link is send by these crackers to some targeted account holders into their e-mail accounts and as they click on the link along with submitting some critical information (usernames, passwords etc.). In this way they lose their personal information like office related documents, project tenders etc. In modern days bank details like usernames, passwords and credit/debit card details are collected by phishing attackers. Some websites are duplicates of some original financial web sites and user is grabbed by giving secret and confidential details of these fake sites. These Crackers do not use their own system to crack the online accounts that makes it very difficult to find these phishing attackers. Cyber Law is also not so strong to punish such non-ethical hackers. In such situations a web site must adopt some

countermeasures with sound web site authentication system. CAPTCHA is also beneficial in this concern that provides a string barrier for bot protection and phishing attack. Figure 2 depicts the various applications of CAPTCHAs.

- *Restricting Robots from Playing Online Games* Golle and Ducheneaut [30], Hilaire et al. [31] discovered that using

Table 7 Text based CAPTCHA development in different countries

Language	Country
Arabic	Saudi Arabia, USA, Iran, Vietnam
Chinese	China, UK, USA
Devanagari	India, USA, China, USA, Vietnam
English	USA, Iraw, UK, China, Vietnam,India
Persian	Iran
Punjabi	India
Urdu	India

by the CAPTCHA, it is very easy to restrict robots or computer from playing online games. This is a fair play for fraud players.

- *Preventing Dictionary Attacks* An online article explained that dictionary attack is a technique for defeating an authentication mechanism by trying to guess its secret password or passphrase by retrieving likely possibilities. Chakrabarti and Singhal [16], Pinkas and Sander [49] explained that in such situations CAPTCHA can be very effective for defending against such dictionary attacks.

- *A Solution for Worms and Spam* CAPTCHA is used provides a solution against worms and spam to receive mail only if there is not a computer program but human behind it.

- *Better Password Management* After a fixed number of wrong attempts of password, an account gets locked, but it is not a better solution. Ahn et al. [3] explained that if the attempts are made by a computer program then it can replace by entering a CAPTCHA to prove that there is a human on the other side and not a computer program.

- *A Barrier for Search Engine Bots* Pope and Kaur [51] also explained that if a company wants that its web pages are not be indexed then CAPTCHA can play a very important barrier for rejecting any computer program that try to index a web page.

- *Restrict Automatic Online Polls* CAPTCHA can also be utilized to restrict a computer program to cast a vote for online polls. Ahn et al. [3] and Pope and Kaur [51] presented the idea that, however, it cannot be used to restrict a human from voting more than one time.

- *A Solution for Rejecting Shopping Agent* It is very common today to develop a software that gives you a complete comparative analysis of prices from various similar web-sites, e.g. *makemytrip*, *trivago* etc. The online stores have a loss because the user is not able to see all the advertisements from these online stores. To reject such software from revealing price details the CAPTCHA can be used very effectively.

7 Reported Work on Text Based CAPTCHAs

First of all it is found that Text based CAPTCHAs are developed all around the world. Table 7 shows the various countries in which the Text Based CAPTCHAs are developed that in short it tells the National and International status of these Text Based schemes:

7.1 Creation of Text Based CAPTCHA's

In this section, the authors have presented CAPTCHA generation techniques. Chen et al. [19] discussed the most popular CAPTCHA scheme is text-based CAPTCHA that starts with Gimpy CAPTCHA. Thereafter, a lot of text-based CAPTCHA comes into the world of internet. Baffel text-based CAPTCHA was developed by Chew and Baird [20] at the Palo Alto Research centre. They designed their CAPTCHA after studying the psychophysics of human reading. The Baffle CAPTCHA is an enhanced version of PessimalPrint CAPTCHA because PessimalPrint is a dictionary-based challenge and it has only 70 words, so it is very easy to break with the probability of 1/70 success. On the other hand, Baffle text is not based on dictionary words. Rusu and Govindaraju [53] proposed a handwritten CAPTCHA. It was the first proposal of handwritten CAPTCHA. The scheme uses two kinds of challenges: actual city names of US (4000) and with challenges generated by random handwritten letters (20,000). The author tested these challenges with advanced word recognizers Word Model Recognizer and HMM Recognizer. Testing was performed on 4127 city names and 3000 words of random letters. They also reported 18% error rate by humans as well, so they also decided to remove some of the confusing letters for enhancing the usability of the handwritten CAPTCHA. Chellapilla et al. [17] proposed a new kind of segment proof CAPTCHA. The author proposed a baseline for designing new challenge that is made up of level 20 of translation, level 20 of rotation, level 20 of scaling and level 75 of global warming. The author proposed a few schemes like the local warp + baseline, Thin Arcs that Intersect + Baseline, Thick Arcs that Intersect + Baseline and Arcs that don't intersect + baseline etc. The author also showed that the human usability is 90% and above. Chow et al. [22] proposed a Clickable CAPTCHA. This CAPTCHA is also based on text CAPTCHA, but it is designed for mobile devices where a keyboard is not presented, or the keyboard takes a lot of time that makes the user frustrated. The user is given a grid of 3 × 4 or 3 × 5 of distorted text images as shown in Fig. 2 and asked to click on the English words because not all the words are true English words. This CAPTCHA is like Microsoft's image based ASIRRA CAPTCHA developed by Elson et al. [26] but it has some advantages. Initially, ASIRRA CAPTCHA is based on a hypothesis that images of animals are not recognized by the computer. But, Golle [29] shows that the problem of telling cats from dogs automatically is significantly easier than hypothesized by the ASIRRA designers on the other hand Clickable CAPTCHA has the security features of text-based CAPTCHA. The second advantage is that text images can be generated algorithmically while the images of cats and dogs are taken from the database that is very small and more vulnerable to attack. The optimal desired success rate of a CAPTCHA is 1 out of 1000. So Clickable CAPTCHAs can be solved 30% faster with a cell phone screen and keypad than Google CAPTCHAs. Thomas et al. [72] proposed a new type of

text-based CAPTCHA that is made of two type's handwritten and printed text. They achieved 77% accuracy of humans and less than 0.0001 accuracy of machine in recognizing this new CAPTCHA. The testing is done with OCR by giving a segmented challenge by assuming that these CAPTCHA can be broken with segmentation attack. The challenge is made segmented proof by perturbation of text. Imsamai and Phimoltares [35] designed a new 3D CAPTCHA in 2010. In April 2014, with some improvement they used it as in an article by Parc's Captchas (2014). Six random alphanumeric letters are given to the user with a lot of distortion like rotation, overlapping, straight line in the middle, salt and pepper noise, background with color/patterns, character color variation, scaling of characters, font variation, use of special characters are used to make it attack proof. The author has shown that the scheme is resistant to pre-processing, vertical segmentation, Color Filling Segmentation, pixel count attack, OCR and Dictionary attack. Susilo et al. [69] proposed a Stereoscopic 3D CAPTCHA. The proposed scheme is text-based CAPTCHA built from stereoscopic 3D images. However, a clear limitation is that it is usable only with the help of 3D display glasses. It is very secure as the 3D images produce a lot of noise. The variable length also strengthens it. Tomas and Govindaraju proposed a synthetic handwritten CAPTHCA. 20,000-character images were taken from US mail and each character image is processed for making it real handwritten word by following 7 steps like a character auto scaling, automatic base line determination, ligature endpoint detection, ligature parameterization, ligature joining, skeleton perturbation and skeleton thickening. The generated CAPTCHA is tested by the program as well as humans. The machine accuracy was 2.6% and the human accuracy was 84%. Rusu et al. [54] extended this idea for generating more verities of this synthetic handwritten CAPTCHA. The idea was based on the concept of Gestalt Laws of Perception and Geon Theory that says that humans are very comfortable in recognizing an object even with its incomplete image. To design a more secure CAPTCHA noise addition, segmentation errors, variable stroke, variable slope and random rotation and stretching is used. Overlapping of characters, occlusion by small circles, rectangle, occlusions by waves where more foreground pixels are present, occlusion by same pixels color arcs, lines with various thickness as foreground pixels colors, use of empty, broken and letter fragmentations, splitting of images in horizontal, vertical or diagonal parts, adding extra strokes of foreground color and changing orientation of words or letters are used for making string design. The author also highlighted some transformations that can be reversed like gaps in image but not successful reversed if handwriting is slant, mosaic effects (separation in parts) but not successfully reversed if discontinuity in strokes that remove parts of image, waves but not successfully reversed if wave thickness is same as character thickness, overlapping

is reversed successfully but not if random overlapping is used within characters, arcs/jaws but not successively reversed if thickness is same as of characters, fragmentation is not reversed, more complex background noise is not easily reversed. The accuracy of the design for humans is 82% and computer is 0. Yalamanchili and Rao [74] proposed a Devanagari script-based CAPTCHA. The word is a random number of letter words. The word contains "*Matra*" as well. To add the noise in a few patterns like mosaic, arcs/jaws, vertically overlapped patterns are used. The headline is also removed to make it unrecognizable for the bots. The author has not given any accuracy rate of the humans as well as error rate of the bot for the new CAPTCHA. Alsuhibany [6] presented an optimization technique to enhance the usability of CAPTCHA that are based on the crowding character together like Google CAPTCHA. The author has designed a 3-stage algorithm to optimize the usability while retaining the security of the CAPTCHA. The algorithm is based on the optimization rules that tells the system to decide what should be the optimized characters for a given confusing character or set of characters, e.g. "vv" should be replaced as "vk", then refining the optimized text that tells the system to refine the optimized text so that after picking the non-confusing characters still the problem is remained, e.g. in "cl" replace with *m* can further create confining as "am" and finally position the optimized characters. Taal et al. [70] proposed an reCAPTCHA assisted OCR for Devanagari text. The design achieved 99.53% accuracy for OCR only and 99.61% accuracy for OCR integrated with reCAPTCHA. The authors used 2-layer Convolution Neural Network for feature learning integrated with 2 Layer Support Vector Machine (SVM) for classification. Saini and Bala [55] presented a new Gurumukhi script-based CAPTCHA for bot protection on the internet. A few advantages are discussed of Gurumukhi CAPTCHA like that can be good for Punjabi users, no need to learn Punjabi as the on-screen keyboard is given in the design as given in the screen keyboard in the design. 41 letters are used to design CAPTCHA random, so the number of possible CAPTCHAs is very large, with the inclusion of audio. The authors reported the average time to solve CAPTCHA 14.07 s and the success rate for humans is 75%. Bursztein et al. [12, 15] proposed an enhancement in Google CAPTCHA scheme by making it more usable scheme. The usability is increased 6.7% from the previous design and it is reported 95.3%. The author applied a lot of changes in visual features (character set, change counts of characters, font size, font families, foreground colors, background colors), anti-segmentation features (character overlapping, random dot size, random dot counts, line types, line counts, line widths, line position, similar foreground/background colors), anti-recognition features (rotated character counts, rotated character degrees, vertical shifting sizes, character size variations, character distortions). Ramaiah and

Govindaraju [52] proposed asigma lognormal model for character level CAPTCHA. The author designed a new character level CAPTCHA because word level CAPTCHA is vulnerable to dictionary attack. The novel idea is to use the accents of different persons' handwritten characters. All the characters are then concatenated with each other by curve fitting. The use of accents makes a single word with a few styles. Human accuracy is 90.49% and machine accuracy is almost nil. Alsuhibany [7] made efforts to define a benchmark for designing usable and secure text-based CAPTCHAs. The author proposed core features and distortion features to make a CAPTCHA scheme more usable and secure. Yu et al. [80] proposed Chinese language-based CAPTCHA along with a comparison with Roman based English CAPTCHA. The main objective is to do an usability analyses on the Chinese characters. The usability is evaluated on terms of 3 independent variables effectiveness (average solving time), efficiency (correction rate) and satisfaction (online questionnaire and face to face interview). They highlighted the major difference between cognition process of solving English and Chinese CAPTCHA's as the Chinese character set is much large and not as simple to recognize as English letters are. The study includes random English characters, frequent English words, random Chinese characters, frequent Chinese words, less frequents Chinese characters, similar form Chinese characters, similar initial consonants and similar simple or compound vowels Chinese characters. The four types of fonts *Yahei*, *Songti*, *Heiti* and *Caoshu* are used in the Chinese CAPTCHA schemes. The last font has the least readability. No trouble is reported in identifying the character by the humans in both languages. Kumari et al. [38] proposed a new CAPTCHA design based on new character locations for enhancing the security of the existing CAPTCHA. The major goal of this new design is to reduce the time for users spending to solve the CAPTCHA. The study has shown that a person can take almost 1–20 s to solve a CAPTCHA around the world every day. So, 200 million CAPTCHA is being solved taking 150,000 h of each day. The solution is to save this time by applying time validations between random CAPTCHA images which comes by clicking reset button. So, when a user clicks in the text box for CAPTCHA entry a counter starts, and a new CAPTCHA image is automatically coming after the counter finishes in the case of CAPTCHA is unsolved by the user. Alsuhibany [8] developed an Arabic handwritten CAPTCHA for cyber security. It is easier for Arab users to solve Arabic CAPTCHA. The images are generated from KHATT database of offline Arabic handwritten text. Various distortions are applied to 28 Arabic characters. Arabic characters are special features as they have different shapes of a single character when it used in beginning, middle and end of word. The overlapping of characters in writing style of Arabic language also makes it secure scheme. Also, the author

has used the less mature Arabic OCR as an advantage for this new CAPTCHA. The algorithm contains 6 steps. In the first step pseudo word generation a 3–8 random letters word is generated, that are converted to Unicode. In the second step "Converting Characters to Unicode". In the third step, selecting a writer the algorithm selects a specific or random writer to generate CAPTCHA. In the fourth step, *selecting a Character Image*, the image of each character is returned to make handwritten Arabic word. The background of height 200 pixels and random width including several handwritten character images is also prepared in this step. In fifth step, Detecting *Joint Points and Connection* the algorithm connects the characters to make a complete word. In the final step *Binarization* the image is converted to B&W. Distortion of black's arcs, white arcs or combination of black and white arcs of different width and length are used. For testing Tesseract, ABBYYY and Newocr.com engine issued, and they are failed in recognition the characters. The robustness of CAPTCHA is 98.3% and Table 8, show the various parameters for these text-based CAPTCHA designs:

7.2 Breaking of Text Based CAPTCHAs

In this section, the authors have discussed various breaking techniques for CAPTCHA proposed by different scholars. Gimpy and EZ Gimpy CAPTCHA were mainly text-based CAPTCHAs. Gimpy CAPTCHA was difficult as compare to EZ Gimpy that is why it is known as EZ (easy). As in Fig. 3 [A] Gimpy CAPTCHA displays approximate 10 words, and the user is asked to correctly type 3 words, but the EZ Gimpy CAPTCHA displays only a single word with cluttered background. Mori and Malik [44] found that Gimpy CAPTCHA uses 411 words and EZ Gimpy uses 561 words. Mori and Malik [44] have proposed two algorithms to break these CAPTCHAs. They optimized the shape context matching technique for matching EZ Gimpy words. They developed novel algorithms for breaking these CAPTCHAs. For EZ Gimpy CAPTCHA, they proposed 3 steps algorithm. The first step performs quick tests to hypothesize locations of letters in the range. In the second step, they extract strings of these hypothesized letters with a direct acyclic graph that form candidate words. They use tri-grams for further minimize the data set at this stage. And in the third step, that choose most likely word(s) by evaluating matching score for each of these words. The success rate for EZ Gimpy is 92%. For Gimpy CAPTCHA they proposed an algorithm that extracts the whole word at once not by individual letters. The algorithm success rate for finding ≥ 1 word(s) in Gimpy CAPTCHA is 92%, ≥ 2 words is 75% and 3 words is 33%. Moy et al. [45] again broke EZ Gimpy and a newer version of Gimpy CAPTCHA that is Gimpy *r* with success rate of 99% and 78%, respectively. EZ-Gimpy is a collection of dictionaries based on 561 words and Gimpy-r produces

Table 8 Generation of Text Based CAPTCHAs

Authors	Scheme	Usability/security in %	Usability enhancements	Security enhancments	Language	Generation technique
Chew and Baird [20]	Baffle text	89/89	Optimal presentation of text	Noise: occlusion by random shapes, Oblitrate parts of image Anti Recognition: Image Masking with AND, OR Anti Segmentation: Overlapping of Random Shapes	English	Gestalt-motivated image-masking and Degradation
Rusu and Govindaraju [53]	Handwritten CAPTCHA	82/88–97	Removal of g,q,r,n,e and c letters	Noise: lines, grids, arcs,circles, Antisegmenation: touching and overlapping characters Anti-Noramlization: Variable stroke, widh, slope and rotation Dictionary attack: confusing and complex characters w and m	English	Handwritten Characters Based
Chellapilla et al. [17]	MSN CAPTCHA	90/65–100	Optimum level of local and global warping with thin and thick arcs interaction	Anti Recognition: Local and Global Warping with Anti Segmentation: Thin and Thick Arcs	English	Random Noise and Warping
Chow et al. [22]	Clickable CAPTCHA	95/ 99.98	Just clikcing is required	Anti Recognition: Rotation Anti Segmentation: Overlapping and Touching Characters Dictionary attack: non English and English words are mixed	English	Multiple text CAPTCHAs
Thomas et al. [72]	Mixed Text Synthetic Hand-written CAPTCHA	77/99,99	Simple English Words	Noise: Background Filled with Letters Anti Recognition: Broken Handwritten Letters Anti Segmentation: Touching Characters	English	Character Level Perturbation
Imsamai and Phimoltares [35]	3D CAPTCHA	100/NM	Simle English Letters with some special symbols	Noise: Lines, Dots, Color background, Scaling, Texture, Anti-Recognition: Random Rotation, Special Charactres, Anti Segmentation: Overlapping, Different Shades and Colors Dictionary Attack: Random Characters	English	3D Rendering

Table 8 (continued)

Authors	Scheme	Usability/security in %	Usability enhancements	Security enhancments	Language	Generation technique
Imsamai and Phimoltares [35]	STE3D-CAP	100/100	Easy to see with 3D Stereoscopic Display Glasses	Noise: 3D Noise Anti Recognition: 3D Letters Anti Segmenation: 3D Image	English	Stereoscopic 3D Images
Rusu et al. [54]	Synthetic Handwritten CAPTCHA	82/NM	US City names and meaningful words	Noise: arcs, waves Anti Recognition: Random handwrittten characters Anti Segmentation: Verticle/Horizontal Overlapping, Displacement and Fragmentation	English	Getault and Geon Theory trnasformations Handwritten images
Ramaiah and Govindaraju [52]	Sigma-Lognormal CAPTCHA	90.40/NM	Simple English letters	Anti-Recognition: Unsupervised selection of character for more variety of same letters Anti-Segmentation: Supervised selection of characters including slant, direction, and curvature	English	Curve Fitting and Use of Accents
Yalamanchili and Rao [74]	DevaCAPTCHA	NM	Devanagari Words	Noise: arcs, jaws, Anti-recognition: different font size Anti-segmentation: vertical overlapping	Devanagari	Simple Image Combination
Alsuhibany [6]	Crowding Character Together CAPTCHA`	NM	English words	Anti-Recognition: Random rotation and warping Anti-Segmentation: Overlapping and touching characters	English	Text Optimizer Optimize Usability Replacing confusing characters with non-confusing characters
Bursztein et al. [12, 15]	Google CAPTCHA	NM/95.3	English/Pseudo words Or Random numbers	Anti-Segmentation: Overlapping, Dots, Lines, Similar foreground/background color) Anti-Recognition: Rotation, Vertical shift, Font size, character distortion)	English	Random Distortions
Alsuhibany [7]	Easy and Secure Fast Hard Annoying	95/100 95/98 3/100 NM/100	Alpha and digits	Noise: Color background Anti-Recognition: Rotation, Shifting, Font Size Anti-segmentation: Overlapping, Lines, Collapsing	English	Core Features and Distortions

Table 8 (continued)

Authors	Scheme	Usability/security in %	Usability enhancements	Security enhancments	Language	Generation technique
Yu et al. [80]	Chinese CAPTCHA	98.44/NM	304Chinese letters	Noise: Dots and Arcs Anti-Recognition: Warping Anti-Segmentation: Chinese Letters Features	Chinese	Chinese Letters Scalability
Saini and Bala [55]	Gurumukhi CAPTCHA	75%	41 Gurmukhi letters	Noise: Dots, Lines, Color Anti-Recognition: Similar Background and Foreground Color Anti-segmentation: Touching	Punjabi	Random Distortions
Alsuhibany [8]	Arabic Handwritten CAPTCHA	83.88%	28 Arabic letters	Anti-recognition: Handwritten Characters Anti-segmentation: Arcs (white, black)	Arabic	Random Handwritten Letters

challenges that are a combination of 4 random letters among 19 letters. The challenges are hard to break with clutter background and distorted letters. The EZ-gimpy is handled with a template image because of very small database and Gimpy-r is handled with the letter segmentation approach due to large possible combinations. The author used a concept of cores (3 most distinct circular of radius 16-pixel areas in the image and 24 mini patches for breaking EZ-Gimpy. To break the Gimpy-*r* challenges the author proposed a novel technique of mesh generation and global/local distortion removal. The author claimed that these techniques can be used to remove all kinds of distortions in the text images. Chellapilla and Simard [18] developed a method to break some popular CAPTCHAs like Mailblocks (88.8%), Register (95.4%), Yahoo/EZ Gimpy (90.3%), Yahoo version2(95.2%), Ticketmaster (82.3%), Goole/Gmail (89.3%). They followed a segmentation step and then a recognition step for break all these HIPs. For Mailblocks the red channel is selected, then binarizes the image and then after erosion extract the largest connected components. Too large components were divided into 2 or 3 components. Vertically overlapped half characters are merged in the last. For Register.com the images are smoothed, binarized and 5 largest connected components are identified. Yahoo/EZ Gimpy (No mesh) images are converted to grayscale images, threshold to black and white and select large connected components. Image is converted to grayscale, threshold to black and white, remove vertical and horizontal lines that don't have neighboring pixels and finally select Connected Components. In case of White mesh, the image is converted to grayscale, threshold into black and white, add black pixels in white line locations if there exist neighboring pixels and finally select large CCs. For Ticketmaster, the image is converted to grayscale, then it is threshold to B&W, dilate and then erode the image and finally select the large Connected Components. Chellapilla and Simard [18] used the same attack to break Yahoo version 2 and Google/Gmail HIP. Aboufadel et al. [1] developed a new method to break a popular CAPTCHA that was known as Holiday Inn Priority CAPTCHA. It is a combination of 5 letters contain lower and upper alphabets and digits. First, they rotated the challenge image to make it horizontally straight. In the next step they segment the characters in 5 images of 30X30 pixels. Finally, they used the Haar Wavelet filters to recognize the characters and achieved 100% accuracy in identifying the characters. Their method was also applied on General Electric CAPTCHA and a CAPTCHA used by Chicago Cubs. Yan and Ahmad [101] designed algorithms to break four CAPTCHA design of CAPTCHAservice.org. These four schemes were word_image (six letter dictionary word), random_letter_images (six random letter image) user_string_image (max 15 distorted letters of the user supplied string). The author segmented the challenges with vertical segmentation and novel Snake Segmentation

Algorithm. The recognition is done by performing statistical analysis on the characters. The CAPTCHA breaking method not only applied to CAPTCHAservice.org scenes, but also BoBlock, BotCheck and HumanVerfiy schemes are also vulnerable to this attack. The success rate of this attack is 100% for all the schemes. The countermeasures are also discussed in the paper. Ahmad and Yan [2] discussed the role of colors in CAPTCHA design with respect to usability and security. A few color CAPTCHAs are broken using their novel algorithms. Gimpy-r (81%), EZ-Gimpy (100%), LinkedIn and FreeCap (100%), BotBlock (100%), Megaupload (63.7%), BotDetect (100%), phpCAPTCHA.org (95–100%) and Cryptograph (100%). The author used a new Color Filling Segmentation (CFS) technique for segmentation for solving color CAPTCHAs. The algorithm fills every large connected component with different color. They also proposed some guidelines for designing color CAPTCHAs. Yan and Ahmad [77] broke Microsoft CAPTCHA with a novel technique with accuracy of 60% in just 80 ms. The author binarizes the image, fixing the broken characters, segmented the image vertically. After this the Color filling Segmentation is applied to fill each segment with different color. A lot of thick arcs remain in the challenge image that are removed with some observations like pixel count, location, shape (without circle). Before recognition some of the connected characters are present in the challenges that are removed by pixel counts. The author also discussed the cases where the algorithm fills to segment the challenge and strength and weaknesses of the Microsoft CAPTCHA. Huang et al. [33] developed an enhanced segmentation algorithm to segment Yahoo and MSN CAPTCHA schemes, and they achieved a segmentation rate of 79% and 76% on MSN and Yahoo CAPTCHA, respectively. Two new algorithms were developed, namely, projection profiles and middle-axis point separation for segmentation of line clutters that are as thick as the characters and they are of varying lengths. These clutters can be intersected or not intersected with characters. With projection profiles the pixel counts are projected of each column on the *x*-axis and using a middle axis point separation method the cutting lines are generated that cut segment the characters vertically. Bursztein [14] developed a new DeCAPTCHA tool for breaking 13 most popular CAPTCHAs like Baidu, Blizzard, Ebay, Google, reCAPTCHA, etc. and they found that most of the schemes are easily broken with their novel attack. A lot of guidelines are proposed to design an anti-recognition (small char set, not use distortion, use rotation in conjunction, etc.) and anti-segmentation (not use complex background, use large lines, match line slope, match line color with text, use collapsing, etc.) scheme. The accuracy ranges from 2 to 73% is achieved in breaking of all these 13 schemes. Gao et al. [27] developed an algorithm to break 5 Hollow CAPTCHAs schemes used by Yahoo, Tencent, Sina, CmPay and Baidu. Gao et al.

[27] achieved 36%, 89%, 59%, 66% and 51% accuracy in breaking these CAPTCHAs. The attack contains image binarization, repair contour lines, CFS to fill hollow parts, noise component removal. The contour repairing is a novel technique explained by the authors. The author also suggested guidelines to make the design stronger. Bursztein et al. [12, 15] developed a generic algorithm to break 6 popular CAPTCHAs schemes (Baidu 54.38%,CNN 48.54%; eBay 48.61%; reCAPTCHA 19.74; Wikipedia 26.36% and Yahoo 5.33%. The algorithm is not based on classical segmentation followed by recognition method, but it does both these tasks simultaneously. The algorithm has four major components Cut-Point detector, Slicer, Scorer and Arbiter. The Cut-Point detector generates a few possible cuts that are the basis for segmentation. The cuts are derived from inflection points of two curves. The curves are made up of examining second derivative of curves generated by following the top and bottom pixels of the CAPTCHA. The slicer applies heuristics to extract the meaning full segments on the cut points and build a graph. The Scorer traverses the graph and applies OCR to each meaningful segment with the help of KNN. Finally, the Arbitrator using the Ensemble Learning Approach selects the final value for the CAPTCHA. The algorithm decreases a lot of cut points by applying some filters and handles the occluding lines with by adding a new character class to the KNN. The author also pinpoints some areas of improvements, but they think that some design improvements can make the CAPTCHA stronger like adding an occluding line of sine wave shapes or blobs. The characterpositions are linear in all CAPTHCAs, somaking the position of each character at random location both vertically and horizontally a defeat the algorithm. Bansal et al. [9] presented a novel approach to break visual CAPTCHA (EZ-GIMPY's four schemes simple background, back mesh background, white mesh background, and loosely connected characters). They reported 97.90% success rate of recognition. The model uses preprocessing (image binarization, low pass filtering, dilation, etc.), character segmentation, feature extraction (number of holes, height of characters, maximum number of white-black transitions, light-fall and nature of vertical stroke) a finally HMM (Hidden Morkov Model) for recognition characters. Hussain et al. [34] proposed a novel CAPTCHA breaking technique for recognition of merged characters. The proposed algorithm to recognize merged characters are having 6 steps. The algorithm starts by taking an input image that is preprocessed (no noise is present) and in the 1st step set Left and Right Margins the algorithm finds the left and right margin of the CAPTCHA. In the 2nd step place window on initial point a window of size *t* (minimum character size in CAPTCHA) is placed on the top left of the window. In the 3rd step (*Move Window Vertically Downwards*) the window is moved vertically downward one by one pixel until it reaches the background pixel means it starts

crossing the foreground character. In the 4th step (*Increase Window Size*) the size of the window is increasing in length one by one pixel until it reaches the character max length and the final window is sent to the classifier for recognition. In the 5th step (*Store Results*) the recognized character is stored and if not then null is stored. In the last step the above steps are repeated until the right margin is reached. Algwil et al. [4] made an analysis of the viability of Chinese language-based CAPTCHA. The authors are motivated towards this new design due to the scalability of the Chinese character set. The Chinese language character set is very large in number (3375 classes) as compare to Roman counterparts (26 classes). It is proven that a computer can recognize single Roman character with a 100% success rate. Algwil et al. [4] reported that this success rate drops 60% in the worst case, but the human success rate also drops to 10%. The recognition of the Chinese character set is not tested so the author has tried to try this. For this, 3 types of Character design are prepared G1 (geometric transformation and random warping), G2 (geometric transformation, random warping, 8 arcs of 1 pixel wide that are placed on the top of each character) and G3 (all distortions of G2 with the arcs are 16). To test the designs for anti-recognition Deep Neural Network (DNN) and Multi Column Deep Neural Network (MCDNN) are used. The error rate of DNN/MCDNN on G1, G2 and G3 are 0.407%, 7.805%, 15.517% and 10.086%, respectively. A popular Chinese CAPTCHA, CCAPTCHA is also tested with DNN/MCDNN and 0% error rate is reported. So, it is concluded that although single Chinese character is successfully recognized by the classifier, but it is possible to make a better CAPTCHA design with Chinese language due to its scalability features. Nguyen [47] developed an algorithm to break 3 most popular, but similar 3D CAPTCHA schemes (Super CAPTCHA, 3dCAPTCHA and TeaBag3D 1.2). The authors found very easy weaknesses in the design and then attacked these 3 CAPTCHA schemes with a success rate of (32% in 3 s, 58% in 4 s and 31% in 4 s) for Super CAPTCHA, 3dCAPTCHA and TeaBag3D 1.2, respectively. Gao et al. [28] presented a generic algorithm for breaking almost every type of text CAPTCHA Microsoft and Wikipedia (Isolated Schemes), Yahoo and QQ (hollow Schemes) and reCAPTCHA and Baidu (CCT Schemes). The toolbox uses Log-Gabor 2D filter for directly extracting the characters from the CAPTCHA without preprocessing the challenge image. The algorithm contains two major steps extracting components and partition and recognition. In the first step 2D Log Gabor Filter is used to extract character component along four directions (0, PI/4, PI/2, 3 × PI/4). In the second step the extracted components are recognized by K-NN. In this step, the components are numbered and sorted with color filling segmentation. A graph is made of the choices for which two components can be grouped as a single character or not. The authors also reported that the attack

is also applicable on broken character schemes like Yandex CAPTCHA. In the last, some counter measures are also suggested like overlapping can enhance the security of text CAPTCHA. Warping is the best option to make the most secure CAPTCHA. The rotation can be useful to make a CAPTCHA stronger. Recently Tang et al. [41] presented a new model for breaking text based CAPTCHA using deep learning. They reported a high success rate in breaking Roman characters and Chinese CAPTCHAs. The study includes 50 web sites CAPTCHAs. Their work is Deep CNN based. Breaking techniques of text based CAPTCHAs are highlighted in Table 9.

8 Reported Work on Image Based CAPTCHAs

8.1 Creation of Image Based CAPTCHAs

The concept of image-based CAPTCHA like BONGO method was first proposed by Bongard [11]. It is a visual pattern recognition problem. The user is given two blocks of some random shapes and finally a random shape is to be identified by the user by telling which block it belongs. No language requirements are needed to pass the test, but a program can be easily trained to solve this kind of visual puzzle. Chew and Tygar [21] described using labeled photographs to generate a CAPTCHA. Elson et al. [26] generated a database of labeled images by feeding a list of easily illustrated words to Google image search. For making it simpler, so that bad images are not shown, they make its database small that resulted in a weak CAPTCHA. Also, labels are also vulnerable to attack on image-based CAPTCHAs. In 2004 PIX CAPTCHA was proposed by Li [40] as a solution to this problem. In this CAPTCHA some label images are picked from the database and a choice of 70 options are given to the user from which he must pick one. The CAPTCHA is presented with four pictures related to a common class. But, the PIX CAPTCHA had a lot of problems like small database only 70 classes are used. If the database is larger than more than 70 classes can be handled that make it time consuming. Finally, some images are related to abstract class, so the user is frustrated. HotCAPTCHA is also introduced in that was based on images. 9 images of humans are presented to the user and the user is asked to select a "hot" one. The "Beauty" is a subjective concept, so the choice is very confusing for a few cases. Li [40] explained that the CAPTCHA is not very popular for most of the websites. Elson et al. [26] at Microsoft developed ASIRRA CAPTCHA. It is an image-based CAPTCHA. The user is given 12 images and is asked to find only cat among 12 images. Image size is 250 × 250. The images are collected from a large image database (3,000,000 images and 10,000 added everyday) of

Table 9 Breaking Techniques of Text Based CAPTCHAs

Authors	Scheme	Breaking Rate in %	Language	Method
Mori and Malik [44]	EZ Gimpy	92	English	Noise Removal:
				Shape Context Matching
	Gimpy	33		Extract whole word Using Bigrams
Moy et al. [45]	EZ Gimpy	99	English	Matching Whole Word by Correlation
	Gimpy r	78		Matching sub-Objects by Distortion Estimation
Chellapilla and Simard [18]	Mailblocks	88.8	English	Noise Removal: Dilation and Erosion
	Register	95.4		Segmentation: Largest Connected Components
	EZ Gimpy	90.3		Recognition: (CNN)
	Yahoo	95.2		
	Version 2	82.3		
	Ticketmaster			
	Google	89.3		
Yan and Ahmad [77]	MSN	60	English	Segmentation: Color Filling and Projection Recognition: (CNN)
Aboufadel et al. [1]	Holiday Inn Priority CAPTCHA	100	English	Segmentation: Vertical Recognition: Bivariate Haar Wavelet Filter
Yan and Ahmad [78]	CAPTCHAservice.org	100	English	Segmentation: Vertical Projection
	BotCheck			Recognition: Pixel Count and Dictionary Attack
	BotBlock			
	Human Verify			
Ahmad and Yan [2]	Gimpy-r	81	English	Segmentation: CFS
	EZ-Gimpy	100		Recognition: CNN
	PhpCAPTCHA.org	95–100		
	Cryptograph	100		
	FreeCap	100		
	LinkedIn	100		
	Megaupload	63.7		
	BotDetect	100		
Huang et al. [33]	Yahoo	76	English	Segmentation: Projection and Point Separation
	MSN	79		
Bursztein et al. [14]	Authorize	66	English	Noise Removal Line removal
	Baidu	5		Segmentation: CFS
	Blizzard	70		Recognition kNN and SVM
	CAPTCHA.net	73		
	CNN	16		
	Digg	20		
	eBay	43		
	Megaupload	93		
	NIH	72		
	Reddit	42		
	Skyrock	2		
	Slashdot	35		
	Wikipedia	25		
Gao et al. [27]	Yahoo	36	English	Noise Removal: Image Binarization
	Tencent	89		
	Sina	59		Segmentation: CFS
	CmPay	66		
	Baidu	51		Recognition: CNN

Table 9 (continued)

Authors	Scheme	Breaking Rate in %	Language	Method
Bursztein et al. [12, 15]	Baidu	36.58	English	Cut point generation
	CNN	48.54		
	eBay	48.61		Recognition: OCR
	reCAPTCHA	19.74		
	Wikipedia	26.36		
	Yahoo	5.33		
Bansal and Gupta	EZ Gimpy	97	English	Noise Removal: Erosion, dilation, AND operation
	Simple			
	Black Mesh			Segmentation: connected components
	White Mesh			Feature Extraction: height of characters, Light fall, number of holes, vertical strokes
	Loosely Connected Characters			Recognition: Hidden Morkov Model
Hussain et al. [34]	Microsoft CAPTCHA	NM	English	Segmentation: windows scaling Recognition: SVM
Algwil et al. [4]	Chinese CAPTCHA	100	Chinese	Noise Removal: binarization and blurring Recognition: DNN and MCDNN
Gao et al. [28]	reCPTCHA	77.2	English	Segmentation: Log Gabor 2D Filter Recognition: KNN
	Yahoo	5.0		
	Baidu	44.2		
	Wikipedia	23.8		
	QQ	56.0		
	Microsoft	16.2		
	Amazon	25.8		
	Taobao	23.4		
	Sina	9.4		
	Ebay	58.8		
Nguyen et al. [47]	Super CAPTCHA	32	English	Segmentation: vertical and diagonal segmentation, Pixel Processing, horizontal and vertical scanning, line distance and line gradient, connected pixels
	3D CAPTCHA	58		
	TeaBagCAPTCHA 1.2	31		
Tang et al. [41]	Roman CAPTCHA	NM	Roman	Deep CNN
	Chinese CAPTCHA		Chiense	

petfindder.com. The author reported a quick response time 15 s to solve it. ASIRRA is secure CAPTCHA from a few aspects. The design of ASIRRA is secured from the Brute Force attack by using the use of the token bucket algorithm. Further, ASIRRA also maintains a two-token bucket, one for per IP and one for per session. This technique saves it in the sharing IP address environment. The designers of ASIRRA also made it more useful and more accessible using their novel partial credit algorithm. That, they use the concept of intermediate verified users. This algorithm made a scope of minor errors for the humans so that even after a wrong selection they are not treated as bots. Also, with PCA and Tokens the success rate of bots is also controlled almost zero. Shirali-Shahreza [58-60] and Shirali-Shahreza [61-65] have presented a few new image-based CAPTCHAs like CAPTCHA for children in which the user is given some images of object like vehicles, animals etc. The user is given audio based (in English language) question to select an object to pass the test. The images are downloaded online from Yahoo etc. This is very effective for such systems that does not support keyboard. In Collage CAPTCHA, the images are presented to the user, but this time the question is rendered on the screen rather than in audio. In the advanced Collage CAPTCHA, the images are given in two sets. Left side displays the goal image and the user are asked to pick all the images on the right side that belong to the left side image. So, the numbers of clicks are more as compared to the previous CAPTCHA schemes. In online collage CAPTCHA, the picture is retrieved online and otherwise it is same as collage CAPTCHA. A photo-based authentication was proposed by Yardi et al. [79]. It is known as Lineup CAPTCHA. The user is given a photo and asked to identify the person in the photo. The images of persons are taken from the Facebook account of the user. The persons that are being asked to belong the

group of the user. The security of the scheme is enhanced by setting the level of the questions being asked. Like what is happening in the photograph, etc. The authors assume that to solve such kind of authentication requires a lot of effort. Almazyad et al. [5] utilized Multi Model CAPTCHA. In this scheme the user is given an image with four text labels on it. The labels are associated with the image, but only one is a perfect match to the image. This perfect label must input by the user to pass the challenge. For security, the labels are embedded in the background of the image, but these are very easily extracted from the image as the color of the text is same for all the labels. Font of the labels is same but not the straight as Times New Roman. Segmentation of the font is straight forward. Another image-based CAPTCHA's that is known as Implicit CAPTCHA and Drawing CAPTCHA was proposed by Shirali-Shahreza [58-60] and Shirali-Shahreza [61-65]. Implicit CAPTCHA is questions-based CAPTCHA. The user is given a question in the image, e.g. to click on the top of the hill, etc. the scheme requires the use of the English language. On the other hand, Drawing CAPTCHA does not require any language proficiency. In this scheme, the user is given a screen with noisy background with a few dots. The user is asked to connect certain dost on the screen. So, it is a mixture of image based and puzzle-based CAPTCHA. Various image based schemes are highlighted in Table 10.

8.2 Breaking Techniques of Image CAPTCHAs

The ASIRRA CAPTCHA is broken with the success are about 82.7% by Golle [29]. The author used a SVM classifier that extracts the color features (F1, F2, F3) and texture features (G1). For the color features, the author uses HSV model and images are divided into cells. The author said that as the number of images in the classifier is increasing the accuracy is also increasing. The author also pointed that the PCA is also helpful in breaking the CAPTCHA. But the Token Buck et al. gorithm is decreasing the success rate of the classifier. Even the use of PCA and Token Bucket does not stop the classifier to break the ASIRRA CAPTCHA. Polakis et al. [49] designed an attack on the photo-based CAPTCHA. They reported that a determined attacker can achieve 100% success rate of breaking such Social Authentication like Facebook. Even a casual attacker can break these challenges with success rate of 22%. They also highlighted a lot of weaknesses of Facebook's social authentication like it requires at least 50 friends to present the challenge, the user's friend must be tagged (this tagging is not very accurate most of the time, so the usability is also very low for a legitimate user). The time of solving this challenge is 5 min that is very long as compare to another CAPTCHA's. 80% of the challenges contain such photos in which the person is not clearly visible. Also, some of the challenges do not contain even a single face to be identified. The challenge is presented if the user is logging from a different geographical location or a new device is used to access the account. Finally, the challenge presents the user with an option to bypass the test by providing their date of birth. The data is obtained very easily by the attacker because the profile contain date of birth, etc. The author tried to break Social Authentication in two ways: casual attack and determined attack. In casual attack the attacker can gain victim's friend list that is publicly available (47% of the users) and a determined attacker can gain 84% by issuing a friend request to the victim's friends. The user ID and names of the friends are retrieved by the Facebook database. The next photos are stored with a user ID and names by accessing the albums of the target's friends. Then by using face detection software the faces are detected. The faces are labeled with the tag

Table 10 Generation of image based CAPTCHAs

Authors	Schemeb	Usability Enchancements	Usability (%)	Security (%)	Language
Bongard [11]	BONGO	Simple Shapes	100	100	English
Elson et al. [26]	ASIRRA CAPTCHA	Only Cat and Dog Images	99.96	100	English
Chew and Tygar [21]	Anomalies Image CAPTCHA	Lableed Images	95	100	Independent
Li [40]	PIX CAPTCHA	Only 70 Choices	90	100	English
Li [40]	HOT CAPTCHA	User friendly choices(beauty)	95	100	English
Shirali-Shahreza [61-65]	Implicit CAPTCHA	Simple clicking is required	100	100	English
Shirali-Shahreza [61-65]	Drawing CAPTCHA	Simple drawing is required	100	100	Independent
Shirali-Shahreza [61-65]	CAPTCHA for Children	Simple images of toys	100	100	English
Shirali-Shahreza [58-60]	Collage CAPTCHA	Larger but easy images are given	100	100	English
Shirali-Shahreza 61-65]	Advanced collage CAPTCHA	Simple matching of images is required	100	100	English
Shirali-Shahreza [58-60]	Online collage CAPTCHA	Online simple images	100	100	English
Mehrnejad et al. [42]	Multiple SEIMCHA	Simple drawings	90	17.5	English
Almazyad et al. [5]	Multi modal CAPTCHA	Clear object and label are given	100	100	English
Yardi et al. [79]	Lineup CAPTCHA	Facebook friends images are used	100	100	English

information. Finally, the names are classified with k-NN classifier with k = 3. The latest indivisible CAPTCHA is announced by Google that is free from every text, image, video and puzzle. It just requires a click of the mouse by the user. It is very famous across the websites. Sivakorn et al. [67, 68] proposed the first successful attack on Google Invisible CAPTCHA is reported. Google reCAPTCHA is based on Google's advanced risk analysis system. Based on the level of confidence assigned to a specific request this system will select which type of challenge to present the user. The collage is presented from simple to hard as follows: First the new user-friendly CAPTCHA is presented to the user. After clicking the checkbox in the widget if the advanced risk analysis system has high confidence, then the checkbox is changed to a tick. If the confidence is not high, then according to the level of confidence one of the following versions of reCAPTCHA is represented as shown below:

- Image reCAPTCHA
- Distorted one word
- Scanned word
- Distorted two word
- Fallback CAPTCHA

The author tried to do a fake click on the Google reCAPTCHA and they succeeded in that. They highlighted a number of weaknesses in the Google's advanced risk analyses in term of browsing history (repeated on 9th day), genuine account and fake accounts (by bots), Geo locations (fraud countries), no detection of Automation of browser actives, mouse behavior (java script based click events), no reputation of cookies files, no site restriction of (number of requested per IP address), etc. The author also proposed a semantic based attack on the Google Image CAPTCHA. They used Google Reverse Image Search for image guess. They also strengthen their algorithm using a lot of Image annotation generator like Clarifai, NeuralTalk, etc. They developed their tag classifier and then attacked image CAPTCHA with the success rate of 70.78%. Another kind of image CAPTCHA is also proposed by Facebook that is

like Google reCAPTCHA. The same technique is also used to break Facebook image CAPTCHA with a success rate of 83.5%. Breaking techniques of image based schemes are highlighted in Table 11.

9 Reported Work on Puzzle Based CAPTCHA

9.1 Creation of Puzzle Based CAPTCHAs

Shirali-Shahreza [59] have also proposed a question–based CAPTCHA. It is a mixture of text based and image-based CAPTCHA. A simple arithmetic problem is given to the user that contains text and images. For example, "There are 5 bananas and 7 apples, then how many total fruits are there?" Again, the user is required to proficient in English language. The challenges are not made of large database so the probability to solve this scheme is very high. Yamamoto et al. [76] proposed SS-CAPTCHA. A few sentences are presented to the users that are made of natural language sentences to the machine-translated sentences. The user is asked to select the one which is created by the human not by the machine. The user must be very proficient in the English. A machine translation program is used to translate natural sentences from the non-mother tongue into a mother-tongue language. Yamamoto et al. [75] developed a Four Panel CAPTCHA. It is a funny CAPTCHA because the user must be able to understand the humor. The challenge is made of four-panel cartoon images, but not arranged in order. To pass the challenge the user must arrange these panels in an order. With 4 panels there are a chance to pass the test by a bot with 1/24 probability. It is time consuming as well as space consuming CAPTCHA scheme. It also requires a lot of dynamic database that is not possible. Mohamed et al. [43] a new kind of CAPTCHA is developed at the University of Alabama at Birmingham. Dynamic Cognitive Game CAPTCHA. It is the solution of relay attacks on the human interactive proofs. It is categorized as a puzzle-based CAPTCHA. It is a set of 4 games Animal Game, Parking Game, Shape Game and Ships Game. In all the games user has a target object on which

Table 11 Breaking Techniques of Image Based CAPTCHAs

Authors	Scheme	Breaking rate	Language	Technique
Golle [29]	ASIRRA CAPTCHA	82.7%	English	Feature Extraction (Color and Texture) Classification: SVM
Polakis et al. [49]	Facebook CAPTCHA	100%	English	Face detection:face.com Classification:kNN
Sivakorn et al.67 [68,]	Google Image CAPTCHA	70.78%	English	Image Collection:Google reverse Classification: tag classification
Sivakorn et al. [67, 68]	Facebook CAPTCHA	83.5%	English	Image Collection:Google reverse Classification: tag classification
Sivakorn et al. [67, 68]	Invisible CAPTCHA	77%	Independent	Fake click implementation

answer object must be dragged. The answer object can be moved on the activity area. The user can move the answer object in the 8 possible directions. The user is presented with only one game at a time to pass the test. The size of the game is 360×130 so it does not consume much space. The author has tested its game in three different conditions i.e. high latency relay, small game relay and low latency relay. The author has succeeded in finding the difference between normal conditions and these three attacking conditions. Although there are some limitations of DCG CAPTCHA like these are not designed for touch screen devices, etc. yet these are very effective in guessing relay attacks. Various puzzle based schemes are highlighted in Table 12.

9.2 Breaking Techniques Puzzle Bases CAPTCHAs

Although no work have been reported for breaking these puzzle based CAPTCHAs. It seems that it is not required at all. It is clear that designing of these puzzles and that makes the database very small. Such schemes can be compromised very easily. These schemes become very difficult to solve when images are used with them. Also these make the user very confused. On the other hand these schemes are not being used by most of the web sites due to the long solving time by the user.

10 Reported Work on Audio CAPTCHA

10.1 Creation of Audio Based CAPTCHAs

Holman et al. [32] proposed a new version of CAPTCHA that are more accessible for visually impaired users and it is equally usable for both the users with and without visually impairments. The CAPTCHA is considering the mix of both pictures and their relating sounds. The user is given a picture and a sound that represents the picture object. The user must select the correct list item from the drop-down list that matches the image or sound to pass the test. The images belong to 4 classes' musical instruments, animals, transportations and weather. Sauer et al. [56, 57] proposed as HIPUU (Human Interaction Proof Universally Usable). In HIPUU the usability is 90% as the images and sounds are very easy to understand by the users. But from the security

of this scheme is not very high as the images are very easily identified by the programs. Tam et al. [71] designed a new CAPTCHA to make the progress rate at least 70% and that sound is not simply broken by the ASR framework. The author uses audio that is made of phrases rather than text, sound or any single sound object. The author believes that human is more comfortable in understanding phrases rather than isolated words. Additionally, the phrases are rendered as old radio program sounds that are of poor quality and difficult to transcribe by ASR systems. Bigham and Cavender [10] proposed a new version of audio-based CAPTCHA that combined playback controls and answer box control into one control for optimizing the performance of the classical audio-based CAPTCHA. The success rate is 68.5% for this new design. The design is an old one but with more control of the playing audio. So, this new audio CAPTCHA is for enhancing the usability of the audio CAPTCHA. Shirali-Shahreza et al. [65] designed Spoken CAPTCHA. The Scheme is designed for blind users who cannot respond to images or text. The scheme works as follows: a word from the database is given as audio sound to the user and user is asked to say the word. The user response is then sent to the server for analyses of file. The server decides that the sound is from a human or machine, if the voice is from the user and correct then the challenge is passed. Lazar et al. [39] proposed an improved audio CAPTCHA that combined with reCAPTCHA. In this new scheme, the sounds are not spoken text, but audio clips of environmental sounds are used. The sounds of objects like train, animals, birds etc. are used. The usability is more than 90% of this design. The design of does not address the problem of distorted audio but is solved in this new scheme. It is real time audio-based challenge. Various audio based schemes are highlighted in Table 13.

The user is asked to identify a specific sound, e.g. a piano from a series of sounds. The scheme is resistant to all three kinds of attack massive guessing (as the wrong input restarts the whole CAPTCHA), a human proxy (there is no time for an intruding process to collect audio data) and algorithmic soling (use of noon textual sounds and less mature audio recognition programs).

Table 12 Generation of Puzzle Based CAPTCHAs

Authors	Scheme	Usability enhancements	Usability (%)	Security (%)	Language
Shirali-Shahreza et al. [59]	Question based CAPTCHA	Simple math questions	100	100	English
Yamamoto et al [76]	SS CAPTCHA	Natural Sentences	90–100	0	English
Yamamoto et al [76]	Four Panel CAPTCHA	Simple Arrangements of 4 Images	100	1/24	English
Manar Mohamed et al. [43]	DCG CAPTCHA	Games are used	92–100	100	English

10.2 Breaking Techniques Audio CAPTCHA

Tam et al. [71] designed a deCAPTCHA tool to break audio CAPTCHA. They broke three popular audio CAPTCHAs at that time like Google audio CAPTCHA (one speaker saying random digits 0–9 with background noise of human voices at varying volumes. The author achieved 67% success rate in breaking Google audio CAPTCHA. Digg CAPTCHA (consists of one speaker saying a random combination of letter and digits. There is also a background noise consists of static sounds like tricking water, etc. For the simplicity 0, 1, 2, 5, 9, I, o and z are not present in the challenge. The author achieved 71% accuracy in breaking this CAPTCHA. Finally, the older version of reCAPTCHA audio CAPTCHA (consists of several speakers speaking random digits with background noise of human voices at varying volumes. The author achieved 45% accuracy in breaking this CAPTCHA. The author used SVM and k-NN classifier for breaking these CAPTCHA. Bursztein and Bethard [13] also proposed new audio CAPTCHA that will contain long phrases from old radio audio files with the assumption that the poor quality will make it harder to recognize by Automated Sound Recognizers at the same time not very difficult to the humans. eBay audio CAPTCHA is also broken successfully at the rate of 75%. The author used the classical steps of breaking text-based CAPTCHA to break this audio CAPTCHA like preprocessing, segmentation and classification. Although, eBay restricted 20 to 40 CAPTCHAs per IP but the author has succeeded in getting approximately 100,000 IP using IP botnet pool to make a large corpus to breaking CAPTCHA. The author used DFT to recognize high energy spikes by applying DFT to the wave files. It helps to isolate energy spikes in the audio file. The author also uses the redundancy of eBay CAPTCHA for improving the performance of deCAPTCHA. Breaking techniques of audio based schemes are highlighted in Table 14.

11 Reported Work on Animation/Video CAPTCHAs

11.1 Creation of Animation/Video CAPTCHAs

One of the video CAPTCHA known as Motion CAPTCHA was developed by Shirali-Shahreza [58-60] and Shirali-Shahreza [61-65]. In this CAPTCHA a video clip is presented to the user in which a person performs some action. Now, a list of actions is given to the user and the user is required to select the correct description according to the actions in the clip. This scheme also requires an English language proficiency. Kluever and Zanibbi [37] proposed a secure and usable video CAPTCHA. The author proposed a new Video CAPTCHA that collects videos from YouTube database. The new design offers 75% usability and 98% security. Cui et al. [25] developed a CAPTCHA based on moving objects. The scheme contains English alphabets and Arabic numerals. Only 3 letters are used to display in a frame. A few attributes are used to make the scheme, dynamic like initial position, color, shape, size and moving orbit of each object. The author has reported that the scheme is safe from Frame Difference attack, Optical Flow based attack, Temporal Difference based attack and Background Subtraction based attack. The usability results are not discussed in detail. Chow and Susilo [23] developed a new Animated 3D

Table 13 Generation of Audio Based CAPTCHAs

Authors	Scheme	Usability enhancements	Usability (%)	Security (%)	Language
Shirali-shahreza et al. [66]	Spoken CAPTCHA	Only respond to a sound clips	92.5	100	English
Bigham and Cavender [10]	New Audio CAPTCHA	More control on sound playback	68.5	100	English
Lazar et al. [39]	Sound Right CAPTCHA	Simple identification of a sound (bell)	> 90	100	English
Holman et al. [32]	CAPTCHA for Blind Users	Just selection of an object is required	100	100	English
Tam et al. [71]	Robust Audio CAPTCHA	Sound phrases are used	> 70	100	English
Sauer et al. [56, 57]	HIPUU	Simple images and sounds are combined	90	100	English

Table 14 Breaking Techniques of Audio Based CAPTCHAs

Authors	Scheme	Breaking rate (%)	Language	Technique
Tam et al. [71]	Google	67	English	Segmentation: Vertical fixed size
	Digg	71		Recognition: SVM and k-NN
	reCAPTCHA	45		
Bursztein and Bethard [13]	eBay	75	English	Segmentation: Vertical
				Recognition: DFT

CAPTCHA based on Motion Parallax. It is also a text-based CAPTCHA but in 3D. The main characters are rendered over background characters. The main characters are also overlapped and crowded. Foreground and Background characters occupy different depths. All characters have the same font and color to make it more secure. But with motion, parallal humans can recognize main characters. The concept of perspective projection is used, but the characters have similar size for making is attack proof. Due to random characters the dictionary attack is also not workable. The local and global distortions are applied. In this Scheme 3 rows are used. The middle row is the challenge. The camera movements and rotations are randomized for movements at varying degrees. The limitation is that the user must tell very specifically that what is to do for this challenge. The authors has reported that the scheme is resisted to various computer vision attacks like Edge Detection (all foreground and background edges are selected), Image Difference (too many overlapping characters), 3D Reconstruction (works only real world scenarios), Optical Flow (fail with noise, texture less regions, and non-rigid objects), Brute Force Attack (random characters), Machine learning attacks (after training is conducted new random challenge is given). Although, the security analyses are given yet usability analysis is not given in the scheme. Various animation/video based schems are highlighted in the following Table 15.

11.2 Breaking Techniques of Animation/Video CAPTCHA

Nguyen et al. [46] made efforts to break HelloCAPTCHA (84 variations). The author has tried to break animated CAPTCHA with simple methods. The 84 variations are grouped in 12 broad categories in HelloCAPTCHA. The algorithm first identifies the scheme by number of frames, maximum frame delay, a few blank frames and background colors. Then, for single image extraction, various techniques are successfully applied. One of the major techniques is PDM (Pixel Delay Map) that uses the concept of pixel display timing. As in most of the target schemes the main challenge is displayed more time as compare to noise for making it more usable. Another method is Extraction by Calculating Line (CL) in which full character is extracted as it touches the vertical line at 1/10 from the upper boundary. A third method is Extraction by color Selection (CS) in which characters are extracted of distinct color in the frame that has maximum pixels. The fourth method is Extraction by Frame Selection (FS) in which vertical segmentation is applied. The maximum number of pixels of each character is also useful to make the selection by so by counting the number of connected pixels the characters are extracted. The fifth method, to extract characters is Extraction by Roller Selection (RS) in which flood fill to extract characters. Finally, character recognition is done by simple attacks. The success rate is reported 16–100%. Xu et al. [73] published their novel attack on moving CAPTCHA. The attack was designed for NuCAPTCHA but it is applicable to all kind of moving text object CAPTCHAs. The scheme has 4 features: (1) The letters are rendered as rigid objects to make it more usable: (2) The background video and foreground character colors are constant while maintaining high contrast. It is also for making more usable shame; (3) The codeword (actual text to be recognized) are random and each have independent and overlapping trajectories for enabling users to distinguish adjacent characters;(4) Codewords are made up of less alphabets for avoiding confusing. The design also has a lot of weaknesses like (1) The color is given in the written that is be recognized in the text; (2) The length of the text not much changed so the number of characters are guessed; (3) The trajectories of codewords is constant etc. The attack algorithm has 4 major steps: detecting salient features and motion, motion trajectories clustering, segmentation and code word extraction and classification. For segmentation, K-means clustering is used on trajectories and for the classification neural network is used. For optimize classification SIFT features are also exploited. The success rate for 3 characters is 77%. The study was very comprehensive. The author first developed an attack for NuCAPTCHA, then they tried to make is more secure and more usable by changing the design. But the changed design did not show much security. The usability was also not much improved. Breaking techniques of animation/video-based schemes are highlighted in Table 16.

Table 15 Generation of Animation/Video Based CAPTCHAs

Authors	Scheme	Usability enhancements	Usability (%)	Security (%)	Language
Cui et al. [25]	3D Animation CAPTCHA	NM Removal C,G,I,O,S,W,Z and 0,2,5,	100	100	English
Chow and Susilo [23]	AniCAP	Easy to understand with Motion Parallax	100	100	English
Shirali-Shahreza [58-60] and Shirali-Shahreza 61[-65]	Motion CAPTCHA	Easy to get the human movements	100	100	English
Kluever and Zanibbi [37]	New video CAPTCHA	Tag related videos are used	75	98	English

12 Breaking of Mouse CAPTCHA and Invisible CAPTCHA

Sivakorn et al. [67, 68] proposed a method to falsify the popular Google latest reCAPTCHA (Mouse CAPTCHA). They conducted a comprehensive study of reCaptcha, and explored that how the advance risk analysis process of Google new scheme is influenced by each aspect of the request.

Through extensive experimentation, they identified serious flaws in this scheme that allows adversaries to effortlessly influence the risk analysis, bypass restrictions, and deploy large-scale attacks. They designed a novel low-cost attack that leverages deep learning technologies for the semantic annotation of images. Their attack is extremely effective, automatically solving 70.78% of the image reCaptcha challenges, while requiring only 19 s per challenge. The Google Invisible CAPTCHA is also the advanced version of Mouse CAPTCHA. After passing some challenges on a particular site the Mouse CAPTCHA is converted to Invisible CAPTCHA and the users actions are analyzed in the backend automatically. Kevin et al. [36] also developed a method known as unCAPTCHA to attack Google reCaptcha by attacking its audio challenge. They reported 85.15% success rate over testing of 450 challenges. Their method solved successfully broke 450 challenges in just 5.42 s. They used speech to text for text generation and speech recognition systems (Google Cloud, Bing Speech Recognition, IBM Bluemix, Google Speech API, Wit-AI, and Sphinx) for recognition. They also used the concept of phonetic mapping and ensembling. The failure of Mouse CAPTCHA indirectly makes the Invisible unusable CAPTCHA also. Breaking techniques of mouse-based scheme is highlighted in Table 17.

13 Tools and Techniques for Security Testing of Various CAPTCHAS

It is observed that breaking of a CAPTCHA is not trivial task. A number of methods and tools are used to break such complex challenges in the following Table 18 the various tools and techniques are highlighted that are being used by the research comminutes to test the security these schemes.

14 Guidelines to Make a Strong CAPTCHA

It is not an easy task to design a secure and usable CAPTCHA. It is always a challenge for the research community to design new ideas for CAPTCHA system. Some of the guidelines are discussed to make a string CAPTCHA (along with its types) in the following Table 19.

15 Conclusions

In this paper, the authors have surveyed most of the CAPTCHA schemes that have been developed around the world. The authors have assessed the work done for various types like text-based CAPTCHAs (i.e. Arabic,

Table 16 Breaking techniques of animation/video based CAPTCHAs

Authors	Scheme	Breaking Rate	Database	Language	Technique
Nguyen et al. [46]	Hello CAPTCHA (84 varients)	16–100%	Alpha and digits	English	Segmentation: pixel display timing, color selection, frame selection, vertical segmentation, connected pixels, flood fill Recognition: OCR
Xu et al. [73]	NuCAPTCHA	77%	Combination of 3 Reduced Alphabets	English	Segmentation: k-means Clustering Feature Extraction (SIFT) Classification: NN

Table 17 Breaking Techniques of Mouse CAPTCHA

Authors	Scheme	Breaking rate (%)	Database	Language	Technique
Sivakorn et al. [67, 68]	Mouse CAPTCHA	70.78	Infinite images	English	Image annotation services and tag classifiers
Kevin et al. [36]	unCAPTCHA	85.15	Digits (0–9)	English	Speech to text Phonetic maping Ensembling Speech recognition

Table 18 Security Testing Techniques Various CAPTCHAs

Scheme	Segmentation, classification tools and techniques	Security testing and techniques	Tools for security testing
Text based CAPTCHAs	Bivariate Harr Wavelet Filter, Dictionary Attacks, Pixel Counts, Log Gabor 2D Filter, Connected Components,Color Filling Segmentation, Horizontal/Vertical Projections Shape Context Matching, Correlations and Distortion Estimations,	Hidden Morkov Mode, kNN, SVM, DNN,MCDNN,CNN	Tesserect, WMR (Word Model Recognizer), Accuscript, CMR, HMM Recognizer,
Image Based CAPTCHAs	Partial Credit Algorithm, Image annotation generator like Clarifai, NeuralTalk, Fake Click Implementation	kNN, SVM (using Color and texture features), Tag Classification, Google Reverse Classification, HSV Model	OpenCV toolkit, Sklearn Face.com API
Audio Based CAPTCHAs	DFT, Vertical Projection	Ada Boost, SVM, kNN,	Sphinx, Decaptcha
Video and Animation Based CAPTCHAs	Pixel display timing, color selection, frame selection, vertical segmentation, connected pixels, flood fill, SIFT	OCR, NN, k-means clustering	ABBYY Fine Reader 11 Professional Edition

Chinese, Devanagari, Gurumukhi and Roman), image-based CAPTCHAs, animation-based CAPTCHAs, puzzle-based CAPTCHAs etc. Also, the authors have presented the work done for breaking of these CAPTCHAs schemes. The authors have presented recognition accuracies achieved for images, characters and numeral of different CAPTCHA schemes. The authors have seen that most of the breaking methods are based on recognition techniques that are based on supervised machine learning techniques. It is also observed that recognition is not big hurdle these days due to the advancement in the classification algorithms, but the schemes that is segment proof is still safer in the modern times. If the work is done to improve the anti-segmentation of these CAPTCHA schemes, specially text-based schemes, then this technique can be used even for upcoming years.

Table 19 Guidelines to make a strong CAPTCHA

Scheme	Guidelines for making secure CAPTCHA
Text based CAPTCHAs	Letters should be perceptually connected and physically disconnected (using arcs)
	Uniform FG and BG provides more security
	More overlapping is effective if random overlapping is used within characters
	Juxtaposing characters in any direction is better
	Noise color should be used
	In case of slant text gaps are useful
	Mosaic effects are useful if discontinuity in strokes that remove parts of text
	Waves and arcs/jaws are useful if wave thickness is same as character thickness
	Randomize line length and match its color with text
	Fragmentation provides more security
	More complex background noise is not easily reversed
	Random number, size, width and positions of character is more effective
	Loops and shared white components less distinguishable
	Let characters connect more randomly
	Cluttering is not effective as projection finds them very easily. Also middle axis separation is very effective in finding these clutters
	Use all layers of security not only one like segmentation (background, lines, collapsing), recognition (multi fonts, font size, distortion, blurring, tilting, waving, rotation) and all Combine more than one distortions
	Use large lines, not use strange slope but match is with text slope
	Use unpredictable collapsing but not use aggressive collapsing
	Use larger character set
	Broken contour are good for more security
	Thick interference arcs should cut through characters
	Crowded characters provides better anti segmentation
	Occluding lines shapes should be sine or blob
	Both type of warping (global and local) are best for security
Image based CAPTCHAs	Social web sites must improve the authentication mechanism
	Dynamically increase the image database
	Deployment of scheme in conjunction with IP monitoring scheme
	Intelligent Tagging of images
Audio based CAPTCHAs	Use of phrase instead of isolated words
	Make it universally usable by inclusion of multiple languages
	Combine images with audio sounds
Video and animation based CAPTCHAs	The number and positions of characters in animations should not be fixed
	Intelligent use of colors
	Use of variable number of frames
	Optimum delay in frames
	Usage of GIF must be avoided
	Lengthy video stream must be preferred
	Random changes in color

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Research Involving Human Participants and/or Animals No human and animal participants were used.

References

1. Aboufadel E, Olsen J, Windle J (2005) Breaking the holiday inn priority club CAPTCHA. College Math J 36(2):101–108
2. Ahmad A, Yan J (2012) CAPTCHA color, usability and security. IEEE Internet Comput 16(2):1089–7801
3. Ahn L, Blum M, Langford J (2004) Telling humans and computers apart automatically. Commun ACM 47(2):56–60
4. Algwil A, Ciresanand D, Liu B (2016) A security analysis of automated chinese turing tests. In: Proceedings of the 32nd annual conference on computer security applications, pp 520–532
5. Almazyad A, Ahmad Y, Kouchay S (2011) Multi-modal CAPTCHA: a user verification scheme. In: Proceedings of

international conference on information science and applications (ICISA), pp 1–7

6. Alsuhibany S (2011) Optimizing CAPTCHA generation. In: Proceedings of 6th international conference on availability, reliability and security (ARES), pp 740–745

7. Alsuhibany S (2016) A benchmark for designing usable and secure text-based CAPTCHAs. Int J Netw Sec Its Appl 8(4):41–54

8. Alsuhibany S (2018) Generating Arabic handwritten CAPTCHA for cyber security. Int J Comput Sci Netw Sec 18(3):41–47

9. Bansal A, Garg A, Gupta A, Gupta A (2008) Breaking A Visual CAPTCHA: A Novel Approach Using HMM, pp 1–6.

10. Bigham J, Cavender A (2009) Evaluating existing audio CAPTCHAs and an interface optimized for non-visual use. In Proceedings of the SIGCHI conference on human factors in computing systems, pp 1829–1838.

11. Bongard MM (1970) Pattern recognition. Hyden Book Co., New York

12. Bursztein E, Aigrain J, Moscicki A (2014) The end is nigh: generic solving of text-based CAPTCHAs. In: Proceedings of 8th USENIX workshop on offensive technologies (WOOT 14), pp 1–15

13. Bursztein E, Bethard S (2009) DeCAPTCHA: breaking 75% of eBay audio CAPTCHAs. In: Proceedings of 3rd USENIX workshop on offensive technologies, pp 1–7.

14. Bursztein E, Martin M, Mitchell J (2011) Text-based CAPTCHA strengths and weaknesses. In: Proceedings of the 18th ACM conference on computer and communications security, pp 125–138

15. Bursztein E, Moscicki A, Fabry C (2014) Easy does it: more usable CAPTCHAs. In: Proceedings of the 32nd annual ACM conference on human factors in computing systems, pp 2637–2646.

16. Chakrabarti S, Singhal M (2007) Password-based authentication: preventing dictionary attacks. Computer 40(6):68–74

17. Chellapilla K, Larson K, Simard P, Czerwinski M (2005) Designing human friendly human interaction proofs (HIPs). In: Proceedings of the SIGCHI conference on human factors in computing systems, pp 711–720

18. Chellapilla K, Simard P (2004) Using machine learning to break visual human interaction proofs (HIPs). In: Proceedings of the advances in neural information processing systems, pp 265–272

19. Chen J, Luo X, Guo Y (2017) A survey on breaking technique of text-based CAPTCHA. In: Security and communication networks, pp 1–15

20. Chew M, Baird HS (2003) Baffle text: a human interactive proof. Proc SPIE 5010:305–316

21. Chew M, Tygar J (2004) Image recognition CAPTCHAs. In: Proceedings of the 7th international information security conference (ISC 2004), pp 268–279

22. Chow R, Golle P, Jakobsson M, Wang L, Wang X (2008) Making CAPTCHA clickable. In: Proceedings of the 9th workshop on mobile computing systems and applications, pp 91–94

23. Chow W, Susilo W (2011) AniCAP: an animated 3D CAPTCHA scheme based on motion parallax. In: Proceedings of 10th international conference on cryptology and network security, pp 255–271

24. Coates A, Baird H, Fateman R (2001) Pessimal print: a reverse turing test. In: Proceedings of international conference on document analysis and recognition, pp 1154–1158

25. Cui S, Mei J, Zhang W (2010) A CAPTCHA implementation based on moving objects recognition problem. In: International conference on E-business and E-government, pp 1277–1280

26. Elson J, Douceur J, Howell J (2007) Asirra: a CAPTCHA that exploits interest-aligned manual image categorization. In: Proceedings of CCS, pp 366–374

27. Gao H, Wang W, Qi J (2013) The robustness of hollow CAPTCHAs. In: Proceedings of the 2013 ACM SIGSAC conference on computer and communications security, pp 1075–1086

28. Gao H, Yan J, Cao F (2016) A simple generic attack on text CAPTCHAs. In: Proceedings of network and distributed system security symposium (NDSS), pp 1–14

29. Golle P (2008) Machine learning attacks against the asirra CAPTCHA. In: Proceedings of the 15th ACM conference on computer and communications security, pp 535–542

30. Golle P, Ducheneaut N (2005) Preventing bots from playing online games. ACM Comput Entertain 3(3):1–10

31. Hilaire S, Kim H, Kim C (2010) How to deal with bot scum in MMORPGs. In: Proceedings of IEEE international workshop technical committee on communications quality and reliability (CQR), pp 1–6

32. Holman J, Lazar J, Feng J (2007) Developing usable CAPTCHAs for blind users. In: Proceedings of the 9th international ACM SIGACCESS conference on computers and accessibility, pp 245–246

33. Huang S, Lee Y, Bell G (2010) An efficient segmentation algorithm for CAPTCHAs with line cluttering and character warping. Multimedia Tools Appl 48(2):267–289

34. Hussain R, Gao H, Kumar K (2016) Recognition of merged characters in text based CAPTCHAs. In: Proceedings of 3rd international conference on computing for sustainable global development, pp 16–18

35. Imsamai M, Phimoltares S (2010) 3D CAPTCHA: a next generation of the CAPTCHA. In: Proceedings of international conference on information science and applications (ICISA), pp 1–8

36. Kevin B, Daven P, George H, Dave L (2017) University of Maryland unCaptcha: a low-resource defeat of reCaptcha's audio challenge. In: WOOT 17-proceedings of the 11th USENIX conference on offensive technology

37. Kluever K, Zanibbi R (2009) Balancing usability and security in a video CAPTCHA. In: Proceedings of the 5th symposium on usable privacy and security (SOUPS), pp 1–11

38. Kumari B, Kumawat A, Gaur H (2017) Enhancing the security of CAPTCHA based on the new character locations. In: Proceedings of 4th international conference on computing for sustainable global development, pp 6997–7001

39. Lazar J, Feng J, Brooks T (2012) The sounds right CAPTCHA: an improved approach to audio human interaction proofs for blind users. In: Proceedings of the SIGCHI conference on human factors in computing systems, pp 2267–2276

40. Li C, Sudani W, Wang J, Liu F, Gill A (2010) Protection through multimedia CAPTCHA. In: Proceedings of 8th international conference on advances in mobile computing and multimedia, pp 63–68

41. Tang M, Gao H, Zhang Y (2019) Research on deep learning techniques in breaking text-based captchas and designing image-based captcha. IEEE Trans Inf Forensics Secur 5(10):2522–2537

42. Mehrnejad M, Bafghi A, Harati A, Toreini E (2011) Multiple SEIMCHA: multiple semantic image CAPTCHA. In: International conference on internet technology and secured transactions (ICITST), pp 196–201

43. Mohamed M, Gao S, Saxena N (2014) Dynamic cognitive game CAPTCHA usability and detection of streaming-based farming. In: Proceedings of the workshop on usable security (USEC), pp 1–10

44. Mori G, Malik J (2003) Recognizing objects in adversarial clutter: breaking a visual CAPTCHA. In: Proceedings of IEEE computer society conference on computer vision and pattern recognition, pp 1–134

45. Moy G, Jones N, Harkless C (2004) Distortion estimation techniques in solving visual CAPTCHAs. In: Proceedings of the IEEE computer society conference on computer vision and pattern recognition, pp 1–6

46. Nguyen V, Chow Y, Susilo W (2012) Breaking an animated CAPTCHA scheme. In: International conference on applied cryptography and network security, pp 12–29.

47. Nguyen V, Chow Y, Susilo W (2014) On the security of text-based 3D CAPTCHAs. Comput Sec 45:84–99

48. Naor M. (1998) Verification of a human in the loop, or identification via the turing test. http://www.wisdom.weizmann.ac.il/~naor/PAPERS/human.pdf

49. Pinkas B, Sander T (2002) Securing passwords against dictionary attacks. In: Proceedings of 9th conference on computer and communications security, pp 161–170

50. Polakis L, Lanciniand M, Kontaxis G (2012) All your face are belong to us: breaking Facebook's social authentication. In: Proceedings of the 28th annual computer security applications conference, pp 399–408

51. Pope C, Kaur K (2005) Is it human or computer? Defending e-commerce with CAPTCHAs. IT Profess 7(2):43–49

52. Ramaiah C, Govindaraju V (2015) A sigma-lognormal model for character level CAPTCHA generation. In: Proceedings of 13th international conference on document analysis and recognition, pp 966–970

53. Rusu A, Govindaraju V (2004) Handwritten CAPTCHA: using the difference in the abilities of humans and machines to read handwritten words. In: Proceedings of 9th IAPR international workshop on frontiers of handwriting recognition, pp 226–231

54. Rusu A, Thomas A, Govindaraju V (2010) Generation and use of handwritten CAPTCHAs. Int J Doc Anal Recogn 13:49–64

55. Saini B, Bala A (2013) Bot protection using CAPTCHA: Gurumukhi script. Int J Appl Innov Eng Manag 2(5):267–275

56. Sauer G, Holman J, Lazar J (2010) Accessible privacy and security: a universally usable human-interaction proof tool. Univ Access Inf Soc 9(3):239–248

57. Sauer G, Lazar J, Hochheiseret H (2010) Towards a universally usable human interaction proof: evaluation of task completion strategies. ACM Trans Access Comput 2(4):15:1-15:32

58. Shirali-Shahreza M, Shirali-Shahreza S (2007) Collage CAPTCHA. In: Proceedings of 9th international symposium on signal processing and its applications, 1–4.

59. Shirali-Shahreza M, Shirali-Shahreza S (2007) Online collage CAPTCHA. In: Proceedings of 8th international workshop on image analysis for multimedia interactive services, pp 58–58

60. Shirali-Shahreza M, Shirali-Shahreza S (2007) Question-based CAPTCHA. In: Proceedings of the international conference on computational intelligence and multimedia applications (ICCIMA 2007), vol 4, pp 54–58

61. Shirali-Shahreza M, Shirali-Shahreza S (2008) Advanced collage CAPTCHA. In: Proceedings of 5th international conference on information technology: new generations, pp 1234–1235

62. Shirali-Shahreza M, Shirali-Shahreza S (2008) Dynamic CAPTCHA. In: International symposium on communications and information technologies, pp 436–440

63. Shirali-Shahreza M, Shirali-Shahreza S (2008) A CAPTCHA system for Nintendo DS. In: Proceedings of the 7th ACM SIGCOMM workshop on network and system support for games, pp 104–105

64. Shirali-Shahreza M, Shirali-Shahreza S (2008) Motion CAPTCHA. In: Proceedings of conference on human system interactions, pp 142–144

65. Shirali-Shahreza M, Shirali-Shahreza S (2008) CAPTCHA for children. In: Proceedings of international conference on system of systems engineering, pp 1–6

66. Shirali-shahreza S, Abolhassani H, Sameti H, Shirali-shahreza M (2009) Spoken CAPTCHA: a CAPTCHA system for blind users. Int Colloq Comput Commun Control Manag 1:221–224

67. Sivakorn S, Polakis I, Angelos D (2016) I am robot deep learning to break semantic image CAPTCHAs. In: Proceedings of IEEE European symposium on security and privacy, pp 388–403

68. Sivakorn S, Polakis I, Keromytis A (2016) I'm not a human: breaking the Google reCAPTCHA. In: Proceedings of annual computer security applications conference (ACSAC), pp 399–408

69. Susilo W, Chow Y, Zhou H (2010) STE3D-CAP: stereoscopic 3D CAPTCHA. In: International conference on cryptology and network security CANS 2010: lecture notes in computer science, vol 6467, pp 221–240

70. Taal K, Atal A, Singh D (2013) reCAPTCHA assisted OCR for Devanagari texts. In: Proceedings of the 1st Indian workshop on machine learning, pp 1–2

71. Tam J, Simsa J, Huggins-Daines D (2008) Improving audio CAPTCHAs. In: Proceedings of international symposium on usable privacy and security (*SOUPS*), pp 1–2

72. Thomas A, Chaudhury S, Govindaraju V (2010) Leveraging the mixed-text segmentation problem to design secure handwritten CAPTCHAs. In: Proceedings of 12th IAPR international conference on handwriting recognition, pp 13–18

73. Xu Y, Reynaga G, Chiasson S (2012) Security and usability challenges of moving-object CAPTCHAs: decoding codewords in motion. In: Proceedings of 21st USENIX security symposium, pp 49–64

74. Yalamanchili S, Rao K (2011) A framework for Devanagari script-based CAPTCHA. Int J Adv Inf Technol 1(4):47–57

75. Yamamoto T, Suzuki T, Nishigaki M (2011) A proposal of four-panel cartoon CAPTCHA: the concept. In: Proceedings of international conference on advanced information networking and applications (AINA), pp 159–166

76. Yamamoto T, Tygar J, Nishigaki M (2010) CAPTCHA using strangeness in machine translation. In: Proceedings of 24th IEEE international conference on advanced information networking and applications (AINA), pp 430–437

77. Yan J, Ahmad A (2008) A low-cost attack on a Microsoft CAPTCHA. In: Proceedings of the 15th ACM conference on computer and communications security, pp 543–554

78. Yan Y, Ahmad A (2007) Breaking visual CAPTCHAs with naive pattern recognition algorithms. In: Proceedings of the 23rd annual computer security applications conference, pp 279–291

79. Yardi S, Feamster N, Bruckman A (2008) Photo-Based Authentication using Social Networks. In: Proceedings of the first workshop on Online social networks, pp 55–60

80. Yu J, Ma X, Han T (2016) Usability investigation on the localization of text CAPTCHAs: take chinese characters as a case study. School of Media and Design, Shanghai Jiao Tong University, Shanghai, China, ResearchGate, pp 1–23