Report

on

“A Spiral CAPTCHA with smart adversarial perturbation to deceive Deep Learning Networks”.

Submitted To:

Dr. C. Krishna Mohan

Prof. CSE Deptt.

Indian Institute of Technology

Hyderabad

Submitted By:

Shivani

AP CSE

Atal Bihari Vajpayee Govt. Institute of Engineering & Technology

Pragatinagar Shimla H.P.

Introduction: text based CAPTCHA

CAPTCHA: **Completed Automated Public Turing Test to Tell Computers and Humans Apart** and was coined by von Ahn et al. [1] in their paper on "**Proposing the challenge of using hard Artificial Intelligence (AI) problems for security**".

Text based CAPTCHA: Text-based CAPTCHAs are most widely used by combination of distorting characters (Alphabets and Numbers) and obfuscation techniques which can be recognized by people but may be hard for automated bots.[2]

Resistance mechanism used:

Types of CAPTCHA (as shown in Figure 1) on the bases of resistance mechanisms used [3]:

a) Character isolated
b) Rotation and warping
c) Overlapping
d) Hollow scheme
e) Varied CAPTCHA string length
f) Noise arcs
g) Complicated background
h) Two-layer structure

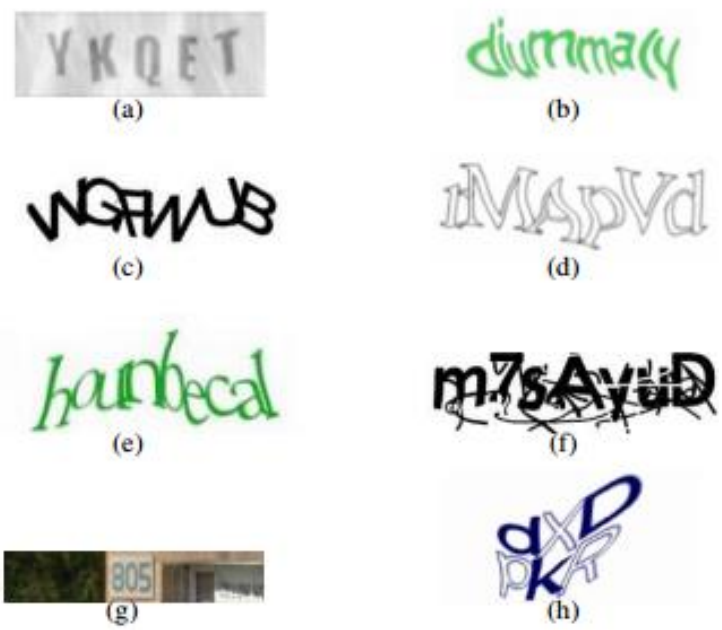

Figure1

Applications:

Various applications of CAPTCHA are [4]:

- Mitigating comment spam
- Online polling
- Web registration
- Boxbe
- Preventing dictionary attacks
- Protecting confidential web pages

- Preventing phishing
- Filtering SMS spam
- Smart cards
- Prevent spam in VoIP
- Preventing sybil attack
- Preventing bots in social network
- Online games

Attack process on text based CAPTCHA:

Attack process on text based CAPTCHA mainly comprises of three steps [2] [3] [5]:

- Pre-processing
- Segmentation
- Recognition

Which uses algorithms/mechanism like:

- K-nearest neighbours
- Support vector machine
- Convolutional neural networks (deep learning)

Limitations of text based CAPTCHA:

Limitations of text based CAPTCHA [6]:

- Constant pixel count of the same character
- Constant colors
- No perturbation/weak perturbation
- Only capital characters
- Constant font
- No rotation
- No deformation
- Constant background/ non-textured background
- Weak color variation
- Weak overlapping

Inferences:

It has been found that intensive work has been done to improve the security using Text Based CAPTCHA. These works are on the basis of design techniques, noise and arcs, complex backgrounds, connected characters, distortion etc. Though these techniques suffer from many drawbacks. The major draw backs found in current text based schemes are:

a) Simple notebook mechanism to write CAPTCHA i.e. starting from left end of image and ending at right end which enables segmentation attacking mechanism to recognize start point and end point easily.

b) Lack of use of intelligent noise to deceive pre-processing attacking mechanisms.

c) Use of one type of character i.e. alphabets only with one font style and one color which enables attacking mechanisms to recognize the pixel color intensity and contour lines easily.

Problem definition:

- On the bases of inferences drawn from literature survey it has been found that text based CAPTCHA are highly used security mechanism against automated bot attacks. Still they lag in many security features and make them vulnerable to various attacks.
- CAPTCHAs used by Microsoft, Google, Yahoo, PayPal, Baidu etc. all has been attacked successfully. Though Text based CAPTCHA are used widely because Image, audio, video, animation, and games based CAPTCHA uses large databases for their storage.
- To overcome these problems a new text based CAPTCHA technique has been proposed "**A Spiral CAPTCHA with smart adversarial perturbation to deceive Deep Learning Networks**".

Objectives:

- To study existing text based CAPTCHA techniques and their limitations.
- To study various attacking mechanisms on text based CAPTCHAs.
- To build a novel text based CAPTCHA "Spiral CAPTCHA" with intelligent perturbation.
- To perform deep learning attack on proposed CAPTCHA design.
- To compare and analyze the obtained results with existing text based CAPTCHA techniques.

Methodology:

To achieve the objectives of the proposed work a new text based CAPTCHA system is to be designed which divides the overall task into following steps:

Step 1	Design of spiral CAPTCHA with randomized fonts, angle, position, alphabets, numerals and intelligent perturbations. **Tools used:** Web page creation using PHP to display CAPTCHA image and verify it by user. XAMPP server to host the web page locally.
Step 2	Creation of datasets of CAPTCHA images. **Tools used:** Creation and storage of image datasets for training and testing purpose.
Step 3	Training a Deep Learning network with datasets and testing it with test datasets. **Tools used:** Creation of conventional neural network using Python and train it using CAPTCHA image dataset.
Step 4	Comparing the results with existing approaches.

Expected outcome:

- It is expected that proposed CAPTCHA design will successfully deceive deep learning techniques and increase the security aspects of text based CAPTCHA over existing text based CAPTCHA techniques.

References:

1. Von Ahn L, Blum M, Hopper N J and Langford J,"CAPTCHA: Using hard AI problems for security" Lect. Notes Comp. Sc. 2656, 2003, pp. 294–311.
2. Ye Wang and Mi Lu, "A self-adaptive algorithm to defeat text-based CAPTCHA" 2016 IEEE International Conference on Industrial Technology (ICIT), Taipei, Taiwan, May 2016, pp. 720-725.
3. Mengyun Tang, Haichang Gao, Yang Zhang, Yi Liu, Ping Zhang and Ping Wang, "Research on Deep Learning Techniques in Breaking Text-based CAPTCHAs and Designing Image-based CAPTCHA", IEEE Transactions on Information Forensics and Security, Vol. 14, No. 8, pp. 2522-2537 , 2016
4. RAGAVI V & G GEETHA, " CAPTCHA AND ITS", APPLICATIONS", Journal of Computer Science Engineering and Information Technology Research, Vol. 4, Issue 1, pp. 11-16, 2014.
5. Jun Chen, Xiangyang Luo, Yanqing Guo, Yi Zhang and Daofu Gong, "A Survey on Breaking Technique of Text-Based CAPTCHA", Security and Communication Networks, Vol. 4, Article ID 6898617, 15 pages, 2017.
6. Haichang Gao, Wei Wang, Jiao Qi, Xuqin Wang, Xiyang Liu, Jeff Yan, "The Robustness of Hollow CAPTCHAs", CCS '13 Proceedings of the 2013 ACM SIGSAC conference on Computer & communications security, Berlin, Germany, November 2013, pp. 1075-1086.